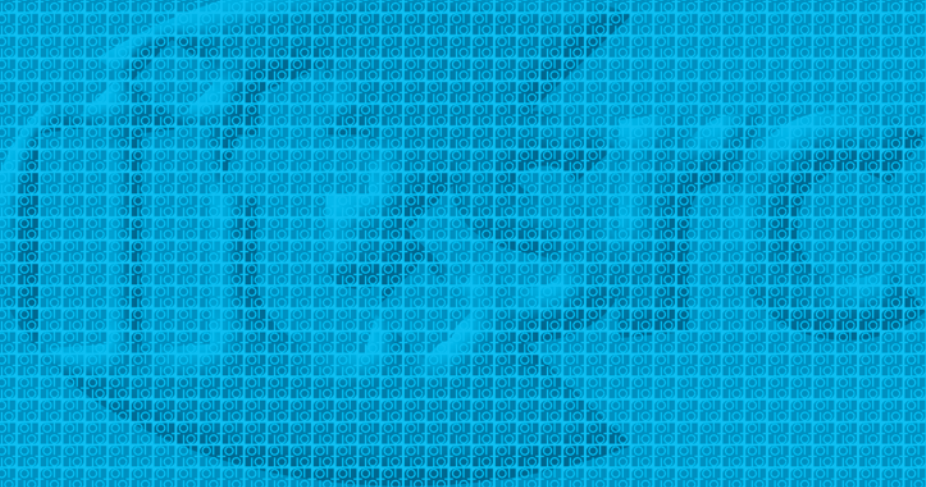
Proceedings in Informatics and Information Technologies

Student Research Conference 2015

Mária Bieliková (Ed.)

Keynote by Jan van Leeuwen



STU FIIT

Proceedings in
Informatics and Information Technologies

IIT.SRC 2015
Student Research Conference

Mária Bieliková (Ed.)

IIT.SRC 2015: Student Research Conference

11th Student Research Conference
in Informatics and Information Technologies
Bratislava, April 23, 2015
Proceedings

Czechoslovakia Section

Slovakia Chapter

SLOVAK UNIVERSITY OF TECHNOLOGY IN BRATISLAVA
Faculty of Informatics and Information Technologies

Proceedings in
Informatics and Information Technologies

IIT.SRC 2015
Student Research Conference

Editor

Mária Bieliková
Faculty of Informatics and Information Technologies
Slovak University of Technology in Bratislava
Ilkovičova 2
842 16 Bratislava, Slovakia

Visit IIT.SRC on the Web: http://iit-src.stuba.sk/2015/

Executive Editors: Katarína Mršková, Anton Andrejko
Copy Editors: Jozef Filipek, Andrej Fogelton, Peter Krátky,
František Kudlačák, Ján Laštinec, Peter Laurinec,
Ondrej Perešíni, Márius Šajgalík, Martin Vojtko,
Petra Vráblecová
FIIT STU
Cover Designer: Peter Kaminský

Published by:
Vydavateľstvo STU
Vazovova 5, Bratislava, Slovakia

ISBN 978-80-227-4342-6

Preface

This volume contains the keynote and student papers selected for presentation and presented at IIT.SRC 2015, the 11th Student Research Conference in Informatics and Information Technologies, held April 23, 2015 at the Faculty of Informatics and Information Technologies of the Slovak University of Technology in Bratislava.

We included in this volume information on all 76 papers presented at the conference, 46 of which are included in their full version, 19 extended abstracts, and information on accompanying events. 11 full papers of master and doctoral students presenting part of their research on ongoing projects were submitted for publication elsewhere and published in this volume in form of one page abstracts. Similarly to previous years these internal publications in most cases are a first step towards later publishing on international established conferences or journals. We suppose that also other papers presented in this volume will evolve into their new (amended/extended) versions and submitted elsewhere.

Research has been one of the main priorities of the university education since its very beginning. It is the case also for our university – the Slovak University of Technology in Bratislava and its faculty – the Faculty of Informatics and Information Technologies. Close connection of research and education leads very naturally to a participation of students in research. This holds not only the students of doctoral study, where research is a substantial part of their study and one of their principal activities. A participation of students in research is "going down" to students of master, even bachelor study.

Universities of technology have a long tradition of students participating in a skilled labour where they have to apply their theoretical knowledge. The best of these results were usually presented at various students' competitions or exhibitions. These were also combined with student research works. Our university has a long tradition in such competition named ŠVOČ (abbreviation of the Student Scientific and Technical Activity). 11 years ago our faculty, FIIT STU, decided to transform former ŠVOČ into the Student Research Conference covering topics of Informatics and Information Technologies (IIT.SRC). Participants are students of all three levels of the study – bachelor (Bc.), master (Ing.) and doctoral (PhD.) study. The conference adopted a form of reviewing as at any other scientific conference, and presenting internally the papers in a form of internal Proceedings, which in most cases means a first step towards later publishing the results on national or international established conferences or journals. However, the aim of reviewing was not in achieving a certain acceptance rate which is currently rather common in our community, but in filtering out papers with low quality or progress and get as best as possible feedback to every student author.

We support students in their (often first) steps in communication their results to the community. We also encourage the students to improve their papers and communicate their work to wider international community. Especially, we support the youngest generation of researchers (bachelor and master students) by travel grants that cover partially travel expenses for the students (which obviously cannot be covered by research grants as these youngest researchers are not formally members of research grant teams).

IIT.SRC 2015 attracted 84 student papers from which 60 were accepted as research papers (19 bachelor, 29 master, 12 doctoral) and 16 as papers to the innovative application and technologies track (6 bachelor, 8 master, 2 doctoral). The number of papers slightly varies each year. This year we have noticed little decrease in bachelor and master categories and an increase in doctoral category comparing to IIT.SRC 2014.

The research track of the IIT.SRC 2015 conference was organized in six sections with papers in two categories – full papers and extended abstracts both presented in course of live discussion in two poster sessions:

- Intelligent Information Processing (12 papers),
- Computer Science and Artificial Intelligence (12 papers),
- Web Science and Engineering (12 papers),
- Software Engineering (11 papers),
- Computer Networks, Computer Systems and Security (8 papers),
- Computer Graphics, Multimedia and Computer Vision (5 papers).

Papers in Innovative applications and technologies track were presented in demo session, where student authors presented their applications online.

The conference was opened by Professor Jan van Leeuwen followed by a keynote titled *The Philosophy of Computation*. Jan van Leeuwen is a professor of computer science (emeritus) at Utrecht University in The Netherlands. His research interests cover all aspects of computer science that involve formal modelling, algorithm design and analysis, and computational complexity theory. Jan van Leeuwen is a member of the Royal Netherlands Society of Sciences and Humanities and the Academia Europaea, and he holds an honorary doctorate from RWTH Aachen University. Last year he received the ACM Distinguished Service Award.

Besides the 76 papers presented at the conference and included in these Proceedings several accompanying events were organized. The RoboCup Exhibition is organised as a part of IIT.SRC from 2005. RoboCup is an attractive project with free participation, designed to support education and research in artificial intelligence, robotics and information technologies. Through several years, our students achieved interesting results, which were presented during the conference. RoboCup exhibition presented both the way the RoboCup simulated league is played and also the progress of current students' research in this field. Seven years ago a new RoboCup league – three-dimensional (3D) robotic simulation was added. The extension of the simulation to the third dimension shows the continuous progress in RoboCup and in our students' skills.

This year we organized for the seventh time as part IIT.SRC a showcase of TP-Cup projects. TP-Cup is a competition of master students' teams aimed at excellence in development information technologies solutions within two semester long team project module. The competition has three stages. 15 teams managed to achieve this stage and presented their projects during the TP-Cup showcase. Extended abstracts of their projects are included in these proceedings.

Accompanying events included for the eight time programming contest. It follows a long tradition at the Slovak University of Technology in Bratislava and our faculty in organizing programming contests, especially the ACM International Collegiate Programming Contest like competitions. This year, we introduced new problem types into the programming contest. Small part of the problems retained the structure of the original contest, which resembles the structure of a traditional ACM-style context. Additionally, for the first time this year, we introduced database querying and error-checking problem types. The goal of this innovation was to motivate students to take part even if they are not very proficient in solving traditional algorithmic problems.

We continued this year with FIITApixel exhibition. FIITApixel brings together both students and staff of the Faculty as well as its potential students and alumni in an effort to create, share and judge pictures. It is organized as an ongoing event, where anyone can contribute pictures. The IIT.SRC FIITApixel exhibition presented the best pictures of this year contest.

For the third time we organized IIT.SRC Junior. It provides a room for presenting inventive high school student projects within the topics of the conference. Four high school students' submissions were selected. All these projects are also presented as extended abstracts for more detailed explanation of proposed ideas and realized prototypes in these proceedings. The first project, authored by Michal Tomáš Buday and Jana Tomanová supports teachers during their work on

student assignment point assessment. In the second project, Dávid Majerčák and Jakub Kozák devised a video-chat which connects people from all around the world. The third project was authored by Dávid Majerčák and Peter Majchrák and involved a prototype of an extension cord, which can be controlled over a cellular network or over the Internet using a website with a user-friendly interface. The fourth project was authored by Martin Pavelka, who created a website about web safety for common users.

This year for the first time we joined our student research conference with International Girls in ICT Day. The Day celebrates girls' interests and strengths, and encourages them to choose a career in information and communication technologies. Our aim was to show high school girls exciting opportunities awaiting them at informatics and information technologies carrier. For this year session, we more than 700 places on 62 different programs were registered. One of program and opening session was directly at the Faculty of Informatics and Information Technologies Slovak University of Technology in Bratislava, who is a key partner of Aj Ty v IT initiative.

IIT.SRC 2015 was for the third time organized in new FIIT building. We all benefited from well-disposed space, which supported live discussions. IIT.SRC 2015 is the result of considerable effort by a number of people. It is our pleasure to express our thanks to:

- the IIT.SRC 2015 Programme Committee who devoted effort to reviewing papers and awards selection,
- the IIT.SRC 2015 Organising Committee and accompanying events coordinators (mentioned in particular reports in these proceedings) for a smooth preparation of the event,
- the students – authors of the papers, for contributing good papers reporting their research and their supervisors for bringing the students to research community.

Special thanks go to:

- Katarína Mršková together with Anton Andrejko and help of several students who did an excellent job in the completion of the proceedings,
- Zuzana Marušincová and the whole organizing committee for effective support of all activities and in making the conference happen.

Finally we highly appreciate the financial support of our sponsors which helped the organizers to provide excellent environment for presentation of the results of student research and valuable awards.

Bratislava, April 2015

Pavel Čičák and Mária Bieliková

Conference Organisation

The 11th Student Research Conference in Informatics and Information Technologies (IIT.SRC), held on April 23, 2015 in Bratislava, was organised by the Slovak University of Technology (and, in particular, its Faculty of Informatics and Information Technologies) in Bratislava.

IIT.SRC General Chair

Pavel Čičák (dean, Faculty of Informatics and Information Technologies,
Slovak University of Technology in Bratislava)

IIT.SRC Programme Chair

Mária Bieliková

IIT.SRC Programme Committee

Michal Barla
Vanda Benešová
Anna Bou Ezzeddine
Peter Drahoš
Elena Gramatová
Ladislav Hudec
Daniela Chudá
Katarína Jelemenská
Peter Kapec
Michal Kompan
Gabriela Kosková
Ivan Kotuliak
Tomáš Kováčik
Alena Kovárová
Tibor Krajčovič
Peter Lacko
Michal Laclavík
Ján Lang
Marián Lekavý
Mária Lucká
Ľudovít Molnár
Pavol Návrat
Peter Pištek
Ivan Polášek
Viera Rozinajová
Petr Šaloun
Jakub Šimko
Marián Šimko
Peter Trúchly
Jozef Tvarožek
Valentino Vranić

IIT.SRC Organising Committee

Alexandra Bieleková
Mária Bieliková
Ivan Kotuliak
Zuzana Marušincová, *Chair*
Katarína Mršková
Ľubica Palatinusová
Branislav Steinmüller
Roman Stovíček

all from FIIT STU in Bratislava, Slovakia

IIT.SRC Accompanying Events Chairs

- *IIT.SRC Junior:* Jakub Šimko
- *TP CUP:* Mária Bieliková
- *RoboCup Showcase:* Ivan Kapustík
- *FIITApixel Exhibition:* Ján Lang

- *Girl's Day:* Petra Kotuliaková
- *Programming Contest:* Márius Šajgalík

IT.SRC General Sponsors

- Capco Slovakia
- Hewlett-Packard Slovakia
- IBM Slovakia
- Softec (supporting the Travel Grants)

IT.SRC Supporting Professional Societies and Foundations

- ACM Slovakia Chapter
- Czechoslovakia Section of IEEE
- Slovak Society for Computer Science
- Informatics Development Foundation at FIIT STU

IT.SRC Medial Partner

- PC Revue

IT.SRC TP CUP Sponsors

- Accenture
- Ditec
- QBSW
- PosAm
- Softec
- Tempest
- Unicorn

IT.SRC FIITAPIXEL Exhibition Sponsor

- Fotolab

Table of Contents

Computer Science and Artificial Intelligence, Volume 1

Web Science and Engineering, Volume 1

Computer Graphics, Multimedia and Computer Vision, Volume 1

Software Engineering, Volume 2

Computer Networks, Computer Systems and Security, Volume 2

Accompanying Events

Innovative Applications and Technologies, Volume 2

IIT.SRC Junior 2015, Volume 2

The Philosophy of Computation

Jan VAN LEEUWEN

Department of Information and Computing Sciences, Utrecht University
Princetonplein 5, 3584 CC Utrecht, The Netherlands
J.vanLeeuwen1@uu.nl

> *"The underlying phenomena is the generative ability of computational systems [. . .]. Knowledge is the posited extensive form of all that can be obtained potentially from this process."*
>
> A. Newell [16]

Abstract. Computation used to be synonymous to calculation. Now computation is what computers do, and it has become the engine of science. Increasingly more powerful machines push the limit of what is known in many fields. Intelligent systems compete with humans and win. Even natural systems like cells or the brain are occasionally modeled in computational terms. Is this use of computation still consistent with Turing's ground-breaking insights from the 1930s that shaped our understanding of computation till the present? Is there a notion of computation that is more fundamental? This challenging question is one of many in the *philosophy of computation*. We consider some of the relevant issues. What does computation actually do for us? Why do we compute? Is it plausible that cognition is computational as a process? Is there some systematic theory that explains it all? We consider possible answers.

1 Introduction

Computation used to be synonymous to calculation. Now it is one the pillars of modern IT and tends to be identified with everything computers do. Ever more powerful machines push the limits of what models can tell us, in any branch of science. Intelligent systems such as *DeepBlue* [15] and *Watson* [7] compete with humans and win. Robots such as *CHAPPiE* [3] make us believe that feelings, emotions, and even consciousness are computational. Can this really be?

Computation is traditionally about calculating functions and solutions to equations, with or without the aid of a calculating machine. The mathematical interest for computation changed character as the result of Hilbert's famous question [13] whether there exists a finite decision procedure for first-order logic. While solving this problem, Turing [22] developed his ground-breaking model of computation that has shaped our thinking of computation since.

One may question the apparent implication that this has defined computation forever. In fact, even Turing himself [24] conceived of various other types of machines that e.g. used real or random numbers, had modifiable programs, or ran forever. He invented, what we now call, "artificial neural nets" and speculated about how machines might think, learn and be intelligent. In 1951, Turing [25] contended that *"machines can be constructed which will simulate the behaviour of the human mind*

IIT.SRC 2015, Bratislava, April 23, 2015, pp. xix – xxx.

very closely", suggesting his belief that mental processes could eventually be understood in terms of, or even *as* computation.

It appears that computation is a much more varied notion than its present technological incarnations suggest. In recent years, the term computation is indeed increasingly being used to describe many other phenomena. Even natural systems like cells or the brain are occasionally viewed as substrates that perform some kind of computation. In his recent book, Valiant [26] even went so far as to explain evolutionary processes like adaptation by computation.

Is Turing's model of computation still the most appropriate for them? Is there a notion of computation that is more fundamental? Is there some overall theory, applicable to whatever sort of computational system, that explains it all? Would it enable us to give satisfactory answers to Abramsky's seminal questions [1]: *why* do we compute, and *what* do we compute?

These challenging question are only some of the many questions in modern *philosophy of computation*. Here we outline some of the issues in the understanding of computation, emphasizing a recent approach by J. Wiedermann and the author [29, 30] and some newer work in [27]. Is it conceivable that eventually a new theory of computation can be devised?

2 Philosophy

Computation has evolved into a highly pluriform notion. This has led to many descriptions of what computation "is". These descriptions invariably aim at some characterisation of *how* computation must work on some underlying model of sufficient generality. We give a brief overview of the issues.

2.1 General

The philosophy of computation began, in principle, with the understanding of calculation and arithmetic in ancient times. It now explores the potential of discrete and scientific computation as we know it today. Other origins are found in information processing and in the development of agent-based systems. Also, the computational modeling of mental and natural processes gave rise to many new perspectives on computation.

The philosophy of computation concerns itself with all questions related to the concrete understanding of computation, in all contexts in which it is recognized. Unlike notions like 'information' and 'communication', computation does not have a long tradition in philosophy. Much of what has been studied arose from questions in logic, mathematics, and the sciences from the late nineteenth century onward.

Some of the major issues include: identifying the objects of study within science and technology, understanding the development of computation and computational modeling in context, understanding computation from e.g. an information- or complexity-oriented viewpoint, the role of representations and theories, and the impact of computation on our thinking and our future. Many questions overlap and have aspects that are not listed here.

2.2 Current Views

Computation is now seen as a much broader notion than captured by Turing's model. Even if there are ways to make e,g, Turing's models fit, this exercise may not give a truthful model [28]. Current approaches either aim to enrich existing models like Turing machines with new mechanisms (such as oracles), devise new models inspired by natural systems (like DNA- and quantum computing), or change the perspective altogether (as in abstract state machines [12]).

Many of the current views hold that computation is some kind of *process* that can be seen as manipulating symbols and transforming information, iteratively or otherwise. The refinements most often added, aim to capture the kind of operational mode by which the transformations are effectuated. Here are some of the typical definitions:

- Computation is the execution of step-by-step procedures for processing information (Valiant [26]).

- The most general definition of computation is as information processing. Computation is the process (or collection of processes) of acquiring information, transforming it, and providing the outcome to the outside world (Akl [2]).

- A computation is a physical process in which physical objects like computers, or slide rules or brains, are used to discover or to demonstrate or to harness properties of abstract objects – like numbers and equations (Deutsch [6]).

- Computation in the broadest sense is anything that happens (as opposed to things being static). If so, then the principles of computation are, in fact, the principles of processes (Frailey [11]).

Even though some scientists prefer to stay with Turing's concept of computation from the 1930s (see [10]), the different viewpoints reflect the broader scope of computation as perceived now. But, where do the definitions point? Have we really gained much so far?

The viewpoints above have in common that they all (try to) describe computation in some absolute, i.e. observer–*independent* way. This may be possible for computation by machines but becomes problematic in other cases e.g. in computational views of natural systems. Shouldn't one focus on the question *what computation does* rather than on *how it does it*? We will outline the viewpoint developed recently by J. Wiedermann and the author [29, 30] that does exactly this.

3 Computation and Knowledge

Can a change of viewpoint help and make a difference? In this section we will argue that it does. Given the limitations of the machine-oriented views, we will look at computation at another level of abstraction. We will argue that computation, above all, is a process of *knowledge generation*.

3.1 A View of Computation

The views of computation have evolved considerably over time. Computation has been looked at as *calculation*, *symbol manipulation*, *information processing* and as a *process*, often in connection to a model of computation that is considered as being realistic now or in the future.

Also the targets of computation have evolved considerably. Computation used to be reserved for evaluating (mathematical) functions and solving models. Nowadays it is seen as the driving force behind any reactive (input-output) behaviour of whatever system or agent, or their components, for which this makes sense. In medicine, the decision schemes for diagnostics or treatments are being used in a similar sense. How can this be read, philosophically?

Following [29], we look at computation as a phenomenon *above* the systems level, at – what Newell [16] has called – the *knowledge level*. The notion of "knowledge" is purposely left open and intuitive. What do we gain by taking this perspective?

The connection between computation and knowledge is not a new one. Philosophers like Lull (1230–1315) and Hobbes (1588–1679) already believed ideas and thought were brought about by computation. Leibniz (1646–1716) contended that all "reasoning" can be reduced to computation, an ideal that is slowly becoming reality. Now we increasingly see processes becoming knowledge-enriched, learning, and *smart*. Computation is increasingly seen as producing knowledge, i.e. information of some kind that is *meaningful* to some substrate or party.

This points to the idea that computation has an epistemical connotation, leading to the following thesis:

- *Computation is a (any) process of knowledge generation, in the context of some suitable knowledge domain* [29].

This viewpoint is attractive because it emphasizes the "goal" rather than the "means" of computation. It does require that one knows what is meant by "knowledge" and how it is taken up, but this is usually understood (and a good test). Some examples of computational systems from different areas are shown in Table 1 (from [29]).

Table 1. Computation as knowledge generation (cf. [29])

Computational system	**Underlying knowledge domain**	**What knowledge is produced**
Contemporary computing systems		
Acceptors	Formal languages	Language membership
Recognizers	Formal languages	Membership function
Translators	Functions, relations	Function value
Scientific computing	Mathematics	Solutions
Theorem provers	Logic	Proofs
Operating systems	Computer's devices and peripheries	Management of computer's own activities
Word processors and graphical editors	Graphical layout, spelling, grammar	Editing skills
Database and information systems	Relations over structured finite domains	Answers to formalized queries
Control systems	Selected domains of human activity	Monitoring, control
Search engines	Relations over unstructured potentially unbounded domains	Answers to queries in a natural language
Artificial cognitive systems	Real world, science	Conjectures, explanations
Natural computing systems		
Living systems, cells	Real world	Life, behavior, intelligence
Brain, mind, social networks	Knowable world	Knowledge of the world
The Universe	Science	Living systems
Non-Turing computing systems		
Compass and ruler	Euclidean geometry	Euclidean constructions
BSS machine [4]	Theory of real numbers	Values of real functions
Oracles [23]	A set $A \subseteq \Sigma^*$	Characteristic function of A
Super-Turing computations	Formal languages in Σ_2	Language membership

The table is mostly self-explaining. Note that the *natural computing systems* in the table satisfy our criterion of computationality, but do not do fit most of the classical definitions of computation. The items in the last part of the table seek the limits of our definition, illustrating that computation may be meaningful as a concept even when there is no immediate physical realization of the computational mechanism at hand.

We do not digress on the further details of the definition here. It would require us to express what it means for a *process* Π with input ω to generate some *knowledge item* κ in the context of a *theory* T and the available knowledge for the knowledge domain at hand. To keep the desired flexibility, we have to allow both *formal* and *informal* theories here, as seen in Table 1. Further details are given in [29, 30].

A major advantage of the given definition is that it is no longer fixed to a specific algorithmic mode of implementation. Abramsky's questions [1] about the nature of computation have a direct answer in this framework as well: we compute in order to generate the knowledge we want or need.

3.2 Observing Computations

The question what constitutes knowledge in a certain context clearly depends on the views or theories of the *observer*. What is knowledge to some, may not be for others. The question is whether observer-dependence is *avoidable* at all when deciding whether some process is computational. We outline a provocative argument from [30] that shows that it may not be.

Consider observers, or agents, that are designed to decide the computationality of any process C that they "see", using the definition of computation above. How might an observer do this? Is there a universal test for computations? Let us specify the question a little further.

Assume, first of all, that observers are computational processes themselves. This is a reasonable assumption if we believe in the testability of computations. A consequence is that it leads to the situation in which one computational process (the observer) has to observe the properties of another, possibly computational process and decide something about it. In the extreme case, an observer may be observing another observer, possible even itself. Before we argue that this is bound to lead to problems, we discuss how an observer could realistically "do its best".

When observing a process C, an observer must verify all requirements of the definition we gave, in order to decide whether C is computational. Thus, the observer must do a systematic check whether the knowledge item κ is indeed derivable from the input ω according to theory T, and whether there is an adequate explanation that the underlying process indeed generates it. In [30] various options are described how an observer might do this, even in the context of "informal" theories. Clearly some observers may be more expert at this than others.

Can there be observers that are competent enough to always make the right decision? This is a difficult question, because we have no absolute notion of decisions being right or wrong when the issues involved are observer-relative like we have. In fact, it might well be that observers inherently disagree many times. The following, informal, result shows that this is indeed unavoidable, under the assumptions we made. The proof is similar to that of Rice's theorem in computability theory.

Theorem 1 (see [30]) *There exists no universal observer whose verdict always agrees with the verdict of every other observer.*

The theorem is almost a "proof" of the observer-relativity of computation. Nevertheless, the claim may be countered by attacking some of the assumptions on which the argument is based. Is the assumption that (universal) observers are computational correct? Can one actually specify and observe processes so a decision of their computationality can always be made? Are observers only useful if they are restricted to the knowledge domains they can handle?

4 A New Model of Computation?

If we accept the view that computation is a knowledge generating process, could it lead to a new kind of theory of computation? It is too early to tell, but we sketch a possible direction in which this may be found. We outline a theoretical model, recently suggested by the author and J. Wiedermann [27].

4.1 Defining Computation

We have distinguished between the "system" level, where the underlying mechanism of a computation unfolds, and the "knowledge" level, where the generated outcomes are observed (as in Newell [16]). Collecting the observables at the respective levels, leads us to consider two separate entities for a computation first: the *action space* of a computation, and the *knowledge space* of a computation.

4.1.1 Action spaces

An action space $\mathbb{A}$ contains the meta-items that capture a given computational mechanism in action. The notion is essential in our understanding of computation and is of course observer-dependent.

Observing the meta-items of a computation in progress suggests that there must be a sense of *proximity* among the meta-items as they occur in sequel. Thus action spaces must have some local structure such that computations can be defined in them in a meaningful way at all. We capture this by postulating the following: *Action spaces have an induced topology.*

The topology of an action space derives from the proximity relation that is obeyed by the action of the mechanism. With this postulate, we can use topological notions and define e.g. continuous mappings over action spaces.

We assume that any action space $\mathbb{A}$ contains a *core set* $\mathbb{A}_0$, consisting of the meta-items that correspond to valid "initialisations" of the underlying mechanism. We make no assumptions on how the mechanism that underlies the computation actually *works*. The mechanism may follow any mode of operation, consist of any number of cooperating components, interact with any environment and more. This gives action spaces the generality we want. We give some examples from [27].

Example 1 *The observable descriptions of a living* cell *form an action space. The meta-items give information about its development, a level of abstraction away from the concrete cell. We may be interested in some special knowledge, e.g. a chemical compound or a property of the cell, which is to be gleaned from the meta-items. (Note that meta-items may be real-valued.) We may also consider the metaspace of a family of cells, as in an experiment. In this case the metaspace is defined by the joint behaviour of the cells over time, with or without taking environmental influences into account.*

Example 2 *The possible "full information descriptions" of a computer executing a (known or unknown) chain of instructions form an action space. Meta-items display the possible instances of registers and memory filled with bits. By "observation" we may read out or interpret any meta-item as knowledge, if indeed it fits the sort of knowledge we are interested in. The meta-items can correspond to any mode of execution (sequential, parallel or distributed).*

4.1.2 Knowledge Spaces

At the knowledge level of a computation, knowledge presumably can be qualified in terms of "knowledge items" (in some domain). We assume that there is some way of delineating the potentially occurring knowledge items in the space (which does not mean that these will all be generable). We assume that every knowledge space $\mathbb{E}$ contains a *core set* $\mathbb{E}_0$ of facts that are initially known, by observation or experience, or just by assumption. The following example is taken again from [27].

Example 3 *The theory of a first-order structure $\mathbb{S}$ forms a knowledge space. The knowledge items are sentences that hold in $\mathbb{S}$. The core set of $\mathbb{S}$ consists of the postulates of $\mathbb{S}$, and the mechanism underlying the space is a combination of first-order inference and the evaluation ("invention") of new sentences. Knowledge in this case follows the standard pattern of a formalized theory.*

One could view the "dispositions" of the *brain* as a knowledge space as well. In this case, the knowledge items would be our possible mind sets, possibly restricted to a certain topic, and the underlying mechanisms would be provided by our thought. The question whether the brain is a "computer" or not (cf. [20, 21]) then reduces to the question how the knowledge space is actually explored.

4.1.3 Computation

We now aim for a definition of computation. The definition will be fully machine- and algorithm-free.

Suppose that we have an action space $\mathbb{A}$ and a knowledge space $\mathbb{E}$. Let $\mathbb{A}_0$ and $\mathbb{E}_0$ denote the core sets of these space, respectively. We first express that (some) action-items x with $x \in \mathbb{A}$ may contain information that maps to knowledge in $\mathbb{E}$. We do this by means of a simple readout function called a *semantic map*.

Definition 1 *A* semantic map *from $\mathbb{A}$ to $\mathbb{E}$ is any partial mapping $\delta : \mathbb{A} \to \mathbb{E}$ with the property that $\delta(\mathbb{A}_0) \subseteq \mathbb{E}_0$.*

We require that $\delta(x)$ is obtained by only a simple "extension" of the observational means that produce x. No substantial extra effort should be involved. The condition that $\delta(\mathbb{A}_0) \subseteq \mathbb{E}_0$ expresses that any knowledge in the items of $\mathbb{A}_0$ should be part of the initial knowledge in $\mathbb{E}_0$.

We now define computations, in this model. Recall that $\mathbb{A}$ is assumed to be a topological space. Thus it makes sense to define *curves* in $\mathbb{A}$. We posit that curves are precisely the sort of trajectories that are traced by computations. Given a curve c, let c^{init} be its starting point and, if it is defined, let c^{end} be its ending point.

Definition 2 *A* computation *is any curve $c \subseteq \mathbb{A}$ with the following properties:*
(i) $\delta(c^{init})$ is defined, and

(ii) if c^{end} is defined, then $\delta(c^{end})$ is defined as well (i.e. $\in \mathbb{E}$).

We require that any computation must start with "some knowledge" but do *not* insist a priori that $\delta(c^{init}) \in \mathbb{E}_0$. At intermediate points, δ need not always be defined. However, if the curve ends, δ must be defined in its ending point.

Definition 3 *A computation c is said to be* enabled *whenever $\delta(c^{init})$ is known.*

The definition of computation by means of curves seems natural. All information about *how* the computation works is hidden, yet one can formulate all usual phenomena like *convergence* (termination) and *composition* with great ease. We refer to [27] for details.

Finally, the term "computation" is often used for a whole *family of computations* that are realized by a same mechanism or some conglomerate of mechanisms. We use the term *bundle* here, indicative of a 'programmed' set of computations.

Definition 4 *A* computation bundle *is any collection of computations $\mathcal{B} = \{c_i\}_{i \in I}$ where I is an index set and for every $i \in I$, c_i is a computation (curve) in $\mathbb{A}$.*

One may argue that all computational systems reviewed in Table 1 can be made to fit the model here, using suitable knowledge- and action spaces. We note that there are various other areas as well in which computations are viewed as 'objects' in suitable spaces. We mention control theory, trace theory (for concurrent systems), and computable topology (in type theory). See [27] for more information.

4.2 Exploring Knowledge Spaces by Computation

Given a knowledge space, how can one discover ("reach") the knowledge items in the space? The problem of knowledge generation is well-studied in philosophy and has given rise to principles like formal inference, informal reasoning, analogy and so on. According to the views in [29, 30], the overriding mechanism for knowledge generation is computation.

In trying to make this more tangible, we consider the definition of computation as given above. What knowledge does a computation generate, in this setting? Can one compute more knowledge items as more are found? Can one characterise the set of all knowledge items that can be generated this way? And, can one "recognize" the knowledge items that can be generated, computationally, in this framework?

4.2.1 Generation

Let $\mathbb{E}$ and let $\mathbb{E}_0$ be as above. Assume that we have a computational mechanism at our disposal for exploring $\mathbb{E}$. Let $\mathbb{A}$ be the underlying action space, $\delta : \mathbb{A} \to \mathbb{E}$ the semantic map that reads out the items, and $\mathcal{B}$ the bundle of computations that we can use.

A crucial question is how "knowledge" is actually extracted from a computation $c \in \mathcal{B}$. If $c^{init} \in \mathbb{A}_0$ and c^{end} is defined, then we may assume that $\delta(c^{end})$ is a "logical consequence" of $\mathbb{E}_0$ and thus knowledge of the sort we are after. However, any knowledge computed "on the way" may be considered as being generated as well. Thus, if c is enabled, the entire set $\delta(c) \subseteq \mathbb{E}$ may be seen generated knowledge. As more and more knowledge is generated, an increasing number of computations from $\mathcal{B}$ get enabled as well. This leads to more knowledge that can be generated, and so on, indefinitely.

Let $K_{\mathcal{B}} \subseteq \mathbb{E}$ be the set of all knowledge items that can be produced and thus become known in this way, using computations from $\mathcal{B}$. In [27] it is shown how to define $K_{\mathcal{B}}$ as a set. More precisely, a set-theoretic operator $G : 2^{\mathbb{E}} \to 2^{\mathbb{E}}$ can be defined such that the following holds.

Theorem 2 (see [27]) *$K_{\mathcal{B}}$ is the least fixed point of G that includes the core set $\mathbb{E}_0$. In particular, $K_{\mathcal{B}}$ is well-defined.*

The proof uses the monotonicity of the knowledge generation process, and relies on the Tarski-Kantorovitch Theorem to prove the existence of the fixed point.

4.2.2 Knowledge Recognition

Finally we consider the *recognition* problem for $\mathbb{E}$. This is the problem of determining, for any given knowledge item $e \in \mathbb{E}$, whether e can be obtained by computation from the core set. Recognition can be of greater concern than generation. For example, recognition processes take place in natural systems such as found on the surfaces of cells and in cognition. Is recognition a computational process?

One may think of recognition as a (new) process that generates all knowledge items that are possible and that "flags" an input item e as soon as e is encountered. It would work precisely for all e with $e \in K_{\mathcal{B}}$. This type of connection between recognition and generation is well-known in classical automata theory [14]. However, is this process computational again, by our definition?

The following definition shows how recognition may be defined as a computational process. For every $d \in D$, let d^+ be a corresponding knowledge item that expresses that d is recognised. Let $D^+ = \{d^+ \mid d \in D\}$.

Definition 5 *A* recognizer *$\mathcal{R}$ for some domain $D \subseteq \mathbb{E}$ consists of the following components:*

- *an action space $\mathbb{B}$ and a knowledge space $\mathbb{F} \supseteq D \cup D^+$,*
- *a semantic mapping $\mu : \mathbb{B} \to \mathbb{F}$,*
- *core sets $\mathbb{B}_0$ and $\mathbb{F}_0$ such that $\{\mu(x) \mid x \in \mathbb{B}_0\} \subseteq \mathbb{F}_0 = D \cup D^+$, and*
- *a bundle of computations $\mathcal{S}$.*

$\mathcal{R}$ is said to recognize *item $d \in D$ if, for some $n \geq 1$, there are computations $s_1, \cdots, s_n \in \mathcal{S}$ with $\delta(s_1^{init}) \in \{d, d^+\}$ such that the composition of $s_1, \cdots, s_n$ leads to knowledge item d^+.*

One can now argue that a recognizer can be construed from the computational process that underlies $\mathbb{E}$. Let $\mathbb{A}$ and $\delta : \mathbb{A} \to \mathbb{E}$ be given for the computational mechanism, and let $\mathbb{B}$ be the bundle we have at our disposal. Assume that $\{\delta(x) \mid x \in \mathbb{A}_0\} = \mathbb{E}_0$.

Theorem 3 (see [27]) *With the given conventions, a recognizer for precisely the items in $K_{\mathcal{B}}$ can be constructed, based on the computational mechanism underlying $\mathbb{E}$.*

The proof shows that one can modify the action space $\mathbb{A}$ and the (view of the) computations, such that a suitably designed semantic map that checks equality between knowledge items does the rest.

The approach using curves enables one to model various properties of computation that seem to be inherent to the notion. The given overview represents 'work in progress' that should eventually learn us how far the model can be extended. We refer to [27] for a more complete description of the ideas.

5 Reflections

We have explored several aspects of the philosophy of computation but focused mainly on the core question, namely understanding the nature of computation itself. We have argued that the rise of the computational paradigm in all sciences calls for a broader concept of computation than is expressed in models like Turing's [22]. Where do we stand?

5.1 When is a Process Computational

In many views of computation, one aims to capture how a process must "compute", in some absolute sense. In order to arrive at a more broadly applicable notion, we have pursued the belief that computations should be understood at – what Newell [16] calls – the knowledge level of the underlying processes. To concretise this, we described the ideas of J. Wiedermann and the author which assert that, in order for a process to be computational, one should be able to view it as a process of knowledge generation. We followed up on this philosophy by designing a topological model of computation.

The details of the model provide a kind of "test" for the computationality of arbitrary processes. The test should be more widely applicable than previous tests, which all use some kind of analogy to information processing by (networks of) computers. Here are the steps needed for testing the broader notion, given some observed or artificial process or set of related processes.

- Specify the action- and knowledge spaces, and their core sets, that play a role. This step involves crucial decisions on the part of the observer on how the processes are to be observed, sensed, measured, and so on.

- In connection to this, a semantic map should be defined that links the observables of the action space, when applicable, to items in the knowledge space.

- Next, the allowable progressions of the underlying mechanism should be described as a bundle of "motions" in action space. Again, this step involves crucial decisions on the part of the observer on what the process is effectuating.

- Finally, there should be a justification that knowledge is generated correctly by the laws of the knowledge space, and that it can be "explained" that this is achievable by the process in action space (regardless of how the process actually does it).

The test has various informal elements. If there is no sense of knowledge being produced, then the test will fail. However, if the test succeeds, the processes at hand can validly be seen as computations, by the philosophy we described. No further assumptions are needed, in particular we do not need any finitary symbolic representations up-front as is often stipulated in the classical case (cf. Fodor [9]). Clearly, some other form of representation may come in when specifying an action space.

5.2 Computation as Knowledge Generation

The understanding of computation as a knowledge generation process brings a variety of advantages that are occasionally lacking in older, mechanistic views. A major advantage is that less conventional cases of computation like processes in cells are covered, in this understanding of computationality. The advantages of viewing computation as knowledge generation can be summarized as follows.

- It gives a criterion for separating computational processes from non-computational ones. We argued that the implicit observer-dependence of this test seems hard to avoid, but also that this seems necessary for applying the notion of computation in modern system contexts.

- It focuses computation on its intrinsic meaning. This gives computation a sound position in philosophy and a potential impact on understanding processes in other areas like cognition and artificial intelligence. For example, in [31] the philosophy is applied in an attempt to understand the nature of *epistemic creativity*. There also may be an impact on epistemology itself, e.g. by the renewed attention for the question of identifying "knowledge generation".

- It does not depend on any operative model of the underlying mechanisms. This also recognizes the computational models which emphasize the role of the knowledge level.

- It resolves some of the notorious "boundary cases" that are difficult with the classical definitions of computation. For example, according to our definition, ecorithms (Valiant [26]) and cognition are computation. For other cases, see [29].

- It allows us to consider computations at a (very) high level of abstraction. This opens the way to new formalizations of computation as a phenomenon, hopefully leading to a new theory of computation that is suited for the applications in modern systems.

- It gives new meaning to the "computation-centric" perspective on computer science. It explains the omnipresence of computer science as a key discipline, as a consequence of the omnipresence of knowledge generation.

- And, finally, it enables one to answer Abramsky's questions [1] about the essential understanding of computation: what do we compute, and why? The answer presents itself at the moment one steps away from the system level to the knowledge level, as explained above.

The philosophy of [29, 30] clearly needs further scrutiny. What are its limits? Boundary cases to investigate are plenty and can be found e.g. in the computational views of cognition [18, 19] and in pancomputationalism [17].

6 Conclusion

The philosophy of computation has become a broad subject, as computational modeling is now recognized (and used) in the study of almost all scientific phenomena. In this overview, we focused on only one question, albeit a crucial one for the entire field namely: what is computation?

We have especially focused on the recent ideas from [27, 29, 30] which relate computation to the generation of *knowledge* in some suitable knowledge domain. What represents knowledge in this context is observer-dependent. The philosophy differs from previous approaches which have tried to find unified models at the systems levels, aiming to explain how computation is performed. In stead, we aim to understand *what* computation does for us, or for that matter, for any agent that values or even relies on the outcome for its future goals. This makes computation to a philosophically sound and relevant notion.

By bringing the worlds of computation and knowledge generation together, the philosophy reflects the fact that the domains have been remarkably converging to each other in the present time. The approach may well lead to a different type of theory than is known for computation in Turing's sense. This may be unavoidable if one wants to capture the broad range of notions that fit under the term computation today. Understanding computation thus continues to be very much on the agenda. Can new, so far undiscovered forms of computation be identified?

Acknowledgement: This paper is based on recent, joint work with Jiří Wiedermann, Institute of Computer Science, Czech Academy of Sciences, Prague, Czech Republic.

References

[1] S. Abramsky, Two puzzles about computation, in: S.B. Cooper, J. van Leeuwen, *Alan Turing – His Work and impact*, Elsevier, 2013, pp. 53–57.

[2] S. G. Akl, What is computation?, *Technical Report* 2013-608, School of Computing, Queens's University, Kingston (Ont.), Canada, 2013.

[3] N. Blomkamp, *CHAPPiE*, Columbia Pictures, 2015.

[4] L. Blum, M. Shub, S. Smale, On a theory of computation and complexity over the real numbers: NP-completeness, recursive functions and universal machines, *Bulletin of the American Mathematical Society* 21 (1): 1–46 (1989).

[5] S. B. Cooper, J. van Leeuwen, *Alan Turing – His Work and Impact*, Elsevier, 2013.

[6] D. Deutsch, What is computation? (How) does nature compute? In: H. Zenil (Editor), *A Computable Universe: Understanding and Exploring Nature as Computation*, World Scientific Publishing Company, 2012, pp. 551–566.

[7] D. Ferrucci *et al.*, Building Watson: An overview of the DeepQA project. *AI Magazine* 31:3 (2010) 59–79.

[8] L. Floridi, *The Philosophy of Information*, Oxford University Press, Oxford, 2011.

[9] J. A. Fodor, The mind-body problem, *Scientific American* 244 (1981) 124–132.

[10] L. Fortnow, The enduring legacy of the Turing Machine, *Comput. J.* 55:7 (2012) 830–831.

[11] D. J. Frailey, Computation is process, *Comput. J.* 55:7 (2012) 817-819.

[12] Y. Gurevich, Foundational analyses of computation, in: S.B. Cooper, A. Dawar, B. Löwe (Eds.), *How the World Computes*, Proc. CiE 2012, Lecture Notes in Computer Science 7318, Springer, 2012, pp. 264–275.

[13] D. Hilbert and W. Ackermann (1928). *Grundzge der theoretischen Logik*,a Springer-Verlag, Berlin, 1928.

[14] J. E. Hopcroft, J.D. Ullman, *Formal languages and their relation to automata*, Addison-Wesley Publishing Company, Reading, MA, 1968.

[15] M. Newborn, *Kasparov versus Deep Blue: computer chess comes of age*, Springer, 1997.

[16] A. Newell, The knowledge level, Presidential address, AAAI'80, 19 August 1980, in: *AI Magazine* 2:2 (1981) 1–20.

[17] G. Piccinini, Computational modelling vs computational explanation: Is everything a Turing machine, and does it matter to the philosophy of mind?, *Australasian Journal of Philosophy* 85:1 (2007) 93—115.

[18] G. Piccinini, S. Bahar, Neural computation and the computational theory of cognition, *Cognitive Science* 37:3 (2013) 453—488.

[19] Z. W. Pylyshyn, *Computation and cognition: Toward a foundation for cognitive science*, MIT Press, Cambridge (MA), 1984.

[20] J. R. Searle, Minds, brains, and programs, *Behavioral and Brain Sciences* 3 (1980) 417–457.

[21] J. R. Searle, Is the brain a digital computer?, *Proceedings and Addresses of the American Philosophical Association* 64:3 (1990) 21–37.

[22] A. M. Turing, On computable numbers, with an application to the *Entscheidungsproblem*, *Proc. London Math. Soc. Series 2*, 42 (1936) 230–265.

[23] A. M. Turing, Systems of logic based on ordinals, *Proc. London Math. Soc. Series 2*, 45 (1939) 161–228.

[24] A. M. Turing, Intelligent machinery, *Report*, National Physical Laboratory, 1948.

[25] A. M. Turing, Intelligent machinery: a heretical theory, in: B. J. Copeland, A lecture and two radio broadcasts on machine intelligence by Alan Turing, in: *Machine Intelligence* Vol 15, Oxford University Press, 1999, pp 445–476.

[26] L. Valiant, *Probably Approximately Correct: Nature's Algorithms for Learning and Prospering in a Complex World*, Basic Books, New York, 2013.

[27] J. van Leeuwen, J. Wiedermann, Knowledge, representation and the dynamics of computation, in: G. Dodig-Crnkovic and R. Giovagnoli (Eds), *Representation and Reality: Humans, Animals and Machines*, Springer-Verlag, 2015 (to appear)

[28] J. Wiedermann, J. van Leeuwen, How we think of computing today, In: *Computability in Europe*, Proc. CiE 2008, LNCS 5028, Springer, 2008, pp. 579–593.

[29] J. Wiedermann, J. van Leeuwen, rethinking computations. in: *6th AISB Symp. on Computing and Philosophy: the Scandal of Computation*, AISB 2013 Convention, Proceedings, AISB, Exeter, 2013.

[30] J. Wiedermann, J. van Leeuwen, Computation as knowledge generation, with application to the observer-relativity problem. In: *7th AISB Symp. on Computing and Philosophy*, Proceedings, AISB 2014 Convention, London, 2014.

[31] J. Wiedermann, J. van Leeuwen, Towards a computational theory of epistemic creativity, in: *Proc. AISB 2015 Symposium on Computational Creativity*, AISB 2015 Convention, Canterbury, 2015 (to appear).

Intelligent Information Processing

Popularity Prediction of Scientific Publications

Adam BACHO*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
adam.bacho@gmail.com

Abstract. With increasing number of scientific publications it is difficult to determine, which of these are of high quality. Citation count is one of the basic ways how to predict the popularity of the articles. However, the main problem arises when we start to talk about newly released articles that have not been cited yet. That is why the aim of this work is to predict the number of citations of the articles in the shortest possible time after publication. We propose a method of citation count prediction that includes some new characteristics, which have yet not been used by the other authors.

1 Introduction

Nowadays, there are dozens of new scientific articles published each day. For researchers it is harder to determine which articles they should use for their further work. Therefore when working with digital libraries authors need to recognize the quality of the articles, because all of the researchers want to write their scientific work on the basis of the best available articles of other authors.

Therefore authors ask themselves: "Which article should I read next?", "Can I trust newly published articles?". In order to answer these questions, we must at first define what being an article of high quality means. One possible definition that also forms the basis for this work, is citation count, i.e., how many times was the given article cited by other articles [4].

The aim of this work is to predict the number of citations of the articles based on their characteristics (title, abstract, authors, journal, etc.) at different time intervals. It is likely, that an article written by the top authors in a specific domain or published by one of the reputable scientific journals, will gain more citations.

In our work we use typical features for our data model, which already have been used by other authors and good results have been achieved. However, we introduced novel features, the most promising of them being *Eigenfactor*[1] that is a characteristic of a journal, in which an article has been published. To the best of our knowledge Eigenfactor has not been used for popularity prediction of scientific publications yet.

---

* Bachelor degree study programme in field: Informatics
Supervisor: Róbert Móro, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

1 http://www.eigenfactor.org/

2 Related work

Nowadays, many authors are engaged in citation count prediction. Either they strive to enhance the current results of other authors or want to introduce new article or journal related features that are able to predict citation count better. One of the key characteristic a prediction model should have is to predict citation count as soon as possible.

According to [8], some authors use features available only after an article has been published and the others try to predict citation count even before publishing an article in any journal or on the internet. The first approach was used in [1, 3, 9] where one of the most often chosen feature was a number of downloads of an article six months after it has been published. However, the second approach is much more interesting for most authors, because no author want to wait several weeks or months. They all want to know a popularity of their article immediately. Author in [6] built his model only with features that are available immediately after or even before an article was published.

The way of predicting citation count differs also by machine learning algorithms that are used. Two basic types authors often use are classification and regression. In some works, e.g., in [10] we can even watch different combinations of these algorithms, when authors used classification algorithm for regression after some small modification.

Classification algorithms were used again for example in [8] and although authors achieved some good results, classification does not get us a concrete number representing citation count, but only classifies an article to its corresponding class. Therefore, when we want to predict a concrete number of citations the regression needs to be used. Mostly used is univariate and multivariate linear regression. This approach has been used in [5, 7] where good citation count predictors were identified while preserving a high measurement accuracy.

If we imagine, what the first thing is that interests us when choosing an article we want to read, the answer is surely its title. This feature is often modified into many different forms and used for predicting citation count. The whole work [7] is devoted to the question if the citation count is influenced by the title length. In this work authors came to the conclusion that the articles with longer titles are prone to be cited more than the articles with shorter titles. However, according to other works using title length as a feature in their model, the title length does not always seem to be a good predictor. The authors in [5] say the title length has negative correlation coefficient in case it is the only predictor. When using this feature in multivariate model it is statistically insignificant. Based on several works and approaches of different authors we can see that title length may be or may not be a good predictor. Therefore, we decided to include this feature in our work and compare our results with other authors.

At the end we want to mention the feature that is, according to many authors, the best predictor of citation count – Journal Impact Factor (JIF). This is not surprising since every author chooses articles primarily from journals that are good, because they assume articles in these journals are good as well. JIF, as one of the many features authors used, was used for prediction in [5]. As in [7], where there were mostly articles from the domain of medicine, even in [5] authors chose articles from medicine journals, specifically from top five journals according to JIF. At the end of the analysis they came with the result showing that only two from all the chosen features were statistically significant, while one of them was JIF. Even in [2] authors say JIF is the best feature in their model.

Based on the results of many authors we chose JIF as one of our features, although our primary goal is to built data model on Eigenfactor. It would be interesting to compare the results of these features for our dataset and as well, if they would be similar to other authors.

3 Method of citation count prediction

Our method of citation count prediction uses linear regression that weighs different features of the articles.

Our goal is to predict citation count for articles 8 years after publication and because of that we parse articles from concrete year for which there are after-8-years citation reports. The very interesting thing we want to achieve is to reduce time interval after which we will be able to predict citation count sufficiently precise. According to this objective we include some citation-related features in our predictive model specifically citations in release year, citations 1, 2 and 5 years after publication. This growth of the citation count over time should have good impact on whole predictive model, especially for linear regression learning phase.

The list of all other article-related features is shown in Table 1. Their hypothesis are also listed and these should increase citation count. However, we do not include just article and citation-related features, but want to try feature that has not been used by other authors yet. This feature is *Eigenfactor* and two its biggest advantages compared with better known *Journal Impact Factor* are:

- *Eigenfactor* is not influenced by journal self-citation and
- the highly cited journals will influence the citation network more than lesser cited journals[2].

However, there is one bigger disadvantage of Eigenfactor and this is that *Eigenfactor* is available only since 2007 what can make it difficult for us to predict citation count 8 years after publication of the article.

To find out, which features are statistically insignificant for our model, we decided p-value to be less than 0.05 in regression results, as it is commonly used by other authors.

4 Evaluation

We evaluated our proposed method by conducting an experiment with the PubMed Dataset. In the next section we at first describe the dataset cleaning process of PubMed Dataset. Subsequently we present the results of evaluation phase, which was done with all features obtained from the dataset. These were used in univariate and multivariate models separately with some modifications.

4.1 PubMed Dataset

At the beginning, the dataset consisted of 907,851 articles from more than 4,734 journals. The dataset is available free on the web[3] and is divided into four logical units, in which all articles are alphabetically ordered by journals that published them.

We cleaned all of these articles while taking following steps:

1. The articles in the dataset are available as XML files, so we had to parse them in order to get the features of the articles.
2. After processing all files in step one, we had 3 files, in which the data, information about citation count and citation frequencies were stored. All these files were formatted as CSV files. The citation count of the articles, which we needed for building our predictive model, we obtained from lists of references at the end of each article with help of PubMed ID. We assume the dataset chosen for this work is complete enough, so that we can consider our predictive model credible.
3. Then we filtered only articles from a specified (2005) year. This year was chosen for 2 important reasons. The first was, that we wanted to predict citation count 8 years after publication whereas our newest citation information was from 2013. The second was, that there were too few articles older than the ones chosen and thus the predictive model would have not been sufficiently precise.

[2] http://admin-apps.webofknowledge.com/JCR/help/h_eigenfact.htm
[3] http://www.ncbi.nlm.nih.gov/pmc/tools/ftp/

4. We also removed all the articles with missing data, that is – all the articles without PubMed ID, year of publication, articles with no authors or no pages and finally articles with no title.
5. At the end of the dataset cleaning, we had *9,582 cleaned articles* with all information necessary for our predictive model. These articles contained only article-related features which were all available in the XML files.

The *Eigenfactor* was not included in our predictive model in the presented evaluation, as it is not available in the dataset by default. However, we plan to obtain it and other journal-related features from the ISI Web of Knowledge, where information about journals is stored by years. After obtaining the features from the web we join them with the articles by ISSN number, which is available in the dataset.

Table 1. Features with their values in the cleaned articles with the corresponding hypotheses.

#	Feature	Median (mean)	Range	Hypothesis
1	special char in a title	-	-	title contains special character
	Length of title:			
2	title (chars)	93 (96)	4-275	more title characters
3	title (words)	13 (13)	1-37	more title words
4	abstract (chars)	1379 (1355)	0-4661	more abstract characters
5	abstract (words)	199 (196)	0-706	more abstract words
	Number of:			
6	references	30 (34)	0-497	more references
7	pages	3 (5)	1-405	more pages
8	authors	4 (5)	1-42	more authors
9	affiliations	0 (0)	0-18	more affiliations
10	keywords	0 (2)	0-22	more keywords
	Citations received after N year(s):			
11	0 (release year)	0 (0)	0-7	more release year citations
12	1	0 (1)	0-19	more after-1-year citations
13	2	0 (1)	0-42	more after-2-years citations
14	5	1 (3)	0-172	more after-5-years citations

4.2 Results

We performed linear regression on a subset of 9,852 cleaned articles published in 2005. We firstly used all the features separately in a univariate model, then we performed multivariate regression with all the features together for predicting citation count 8 years after the article's publication.

For the univariate model, only one feature (special char in a title) had the p-value greater than 0.05 [p = 0.87], thus this features was marked as statistically insignificant predictor. In order to get statistically more significant results we performed 10-fold cross-validation for obtaining R^2 and RSE, which stands for residual standard error. As we can see in Table 2, the most powerful results were achieved with feature "citations received 5 years after publication", where R^2 = 87.8 % and RSE = 3.77, which is much less than it was for the other features. Although these numbers may seem great it is not that good. The reason is that we need 5 years after the publication of the article to obtain this feature and that is really long time when we take into account the fact we are predicting citation count just 8 years after the publication.

Therefore, the multivariate model is much more interesting for us. We did the same process as for the univariate model, but this time we included all of the features for creating only one multivariate model. The results in Table 2 show more features that were statistically insignificant, concretely special char in a title (again), abstract (chars and words), number of affiliations, number of keywords, citations received in release year and 1 year after publication. However, after performing 10-fold cross-validation, adjusted R^2 used in multivariate model was 89.4 % which

means that the model fitted almost perfectly on the test set. RSE, in this case, was counted as the mean of residual square errors of each fold. Its value (RSE = 3.52) is more positive than in the case of univariate model.

Finally, we wanted to predict citation count as soon as possible without waiting 5 years after the publication of the article. Thus, we performed four more multivariate linear regressions with no information about citations, with citations received in release year, 1 year after the publication and 2 years after the publication separately and only with features that were statistically significant in multivariate model of all the features. According to these results shown in Table 3 we can conclude that the information about citation count strongly correlates with power of predictive model and need to be contained in it.

Table 2. The results of univariate and multivariate model.

VARIABLE	UNIVARIATE MODEL					MULTIVARIATE MODEL			
	Reg. coef.	p-value	Cor. coef.	R^2 (%)	RSE	Reg. coef.	p-value	R^2 adj.	RSE
title (chars)	-0.00608	0.04	0.0346	0.0228	10.8	-0.00799	0.0099	89.4%	3.52
title (words)	-0.0462	0.031	0.0324	0.0291	10.8	0.0462	0.0352		
special char in a title	-0.0369	**0.87**	0.0104	-0.0315	10.8	0.0933	**0.2171**		
abstract (chars)	0.00053	0.002	0.1120	0.08	10.8	0.00005	**0.9103**		
abstract (words)	0.00324	0.006	0.1070	0.0577	10.8	-0.00183	**0.5506**		
references	0.06749	< 0.001	0.3740	2.76	10.7	0.00368	0.0083		
pages	0.1822	< 0.001	0.1480	0.717	10.8	0.0315	< 0.001		
authors	0.4823	< 0.001	0.1980	2.06	10.7	0.0731	< 0.001		
affiliations	1.163	< 0.001	0.0838	1.12	10.7	0.0557	**0.1114**		
keywords	-0.1781	< 0.001	-0.0933	0.141	10.8	-0.00511	**0.7107**		
citations in release year	8.381	< 0.001	0.3000	11.1	10.2	0.0649	**0.5366**		
citations after 1 year	5.7576	< 0.001	0.5600	37.9	8.52	-0.0904	**0.2103**		
citations after 2 years	4.0403	< 0.001	0.7020	54.9	7.26	-1.28	< 0.001		
citations after 5 years	1.90472	< 0.001	0.9070	87.8	3.77	2.32	< 0.001		

Table 3. The results of multivariate model after adding citations in different years.

Model after adding citations:	R^2 adj. (%)	RSE
without adding citations	5.51	10.50
in release year	15.1	9.95
until a year after publication	39.7	8.39
until 2 years after publication	56.1	7.16
until 5 years after publication	89.4	3.52

5 Conclusions and future work

In this work we created several predictive models based on the features we chose from the PubMed Dataset that stores mainly medicine-domain articles. We then performed the univariate and multivariate linear regression separately.

We achieved good results from model in which all of the features were used, including citations 5 years after publication [R^2 = 89.4 %, RSE = 3.52]. However our goal was to be able to predict citation count 8 years after publication as soon as possible. Therefore, other predictive models were built using no citation information, citations in release year and so on until 5 year interval been reached. Based on the results of these models we can conclude the more information about how citation count has been evolving over time we have the better results can be achieved. According to these findings we want to include some alternative metrics in our

further work, because they can dramatically reduce waiting time for information about citation growth.

Naturally, our further expectations include greater contribution and improvement of our predictive model after adding journal related features, concretely *Eigenfactor* and *Journal Impact Factor*. These two are available from the ISI Web Of Knowledge.

In conclusion we must take into account the fact that the predictive models and the features obtained from articles vary from one domain to another. Our model is thus with very high probability unusable outside the domain of medicine.

Acknowledgement: This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG 1/0646/15.

References

[1] Brody, T., Harnad, S., Carr, L.: Earlier web usage statistics as predictors of later citation impact. In: *Journal of the American Society for Information Science and Technology*, (2006), pp. 1060–1072.

[2] Callaham, M., Wears, R.L.; Weber, E.: Journal prestige, publication bias, and other characteristics associated with citation of published studies in peer-reviewed journals. In: *Jama*, (2002), pp. 2847–2850.

[3] Castillo, C., Donato, D., Gionis, A.: Estimating number of citations using author reputation. In: *String processing and information retrieval*. Springer Berlin Heidelberg, (2007), pp. 107–117.

[4] Cheek, J., Garnham, B., Quan, J.: What's in a number? Issues in providing evidence of impact and quality of research (ers). In: *Qualitative Health Research*, (2006), pp. 423–435.

[5] Falagas, M.E., et al.: The impact of article length on the number of future citations: a bibliometric analysis of general medicine journals. In: *PloS one*, (2013).

[6] Fu, L.D., Aliferis, C.: Models for predicting and explaining citation count of biomedical articles. In: *AMIA Annual Symposium Proceedings*. American Medical Informatics Association, (2008), p. 222.

[7] Habibzadeh, F., Yadollahie, M.: Are shorter article titles more attractive for citations? Crosssectional study of 22 scientific journals. In: *Croatian medical journal*, (2010), pp. 165–170.

[8] Ibáñez, A., Larrañaga, P., Bielza, C.: Predicting citation count of Bioinformatics papers within four years of publication. In: *Bioinformatics*, (2009), pp. 3303–3309.

[9] Lokker, C., et al.: Prediction of citation counts for clinical articles at two years using data available within three weeks of publication: retrospective cohort study. In: *BMJ*, (2008), pp. 655–657.

[10] Yan, R., et al.: Citation count prediction: Learning to estimate future citations for literature. In: *Proceedings of the 20th ACM international conference on Information and knowledge management*. ACM, (2011), pp. 1247–1252.

Automatic Diacritics Reconstruction in Slovak Texts

Jakub GEDERA*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
xgedera@fiit.stuba.sk

Abstract. In this paper we describe the method of diacritic reconstruction in text written in Slovak language. Slovak language contains words that can have more than one possible diacritic reconstruction. In solving this task, we employ context based approach. Using n-gram language model we are looking the most likely occurring forms of terms in a sentence. The method will be implemented as a web service and will be available in the form of extension for web browser for users as a support tool for everyday work with Slovak text during browsing.

1 Introduction

There is a lot of Slovak texts on the Web written without diacritic. This complicates various tasks related with intelligent information processing such as texts categorization or metadata extraction. For example, text categorization is less effective when working with text without diacritics. Extracted metadata from text without diacritics may be different than extracted metadata from text with diacritics. Reading text with diacritics is more comfortable for users and they miss support tools for diacritics reconstruction on the Web.

The aim of our work is automatic reconstruction diacritic in Slovak texts written without it. We assume that words in the text are written correctly and they are missing diacritic only. Methods of diacritic reconstruction can be broken down as follows [7]:

- Methods independent of the context
- Methods dependent on the context
- Combined methods

Best results are achieved by the combined methods. In this paper, we utilize context to achieve the highest accuracy of diacritics reconstruction. In order to find the correct form of the word with diacritic, we use language model to find most likely occurring sequence of words [5].

Our web service will be available in the form of a web browser extension. Utilizing the extension, we will be able to evaluate the accuracy of diacritics reconstruction as well as user

* Bachelor degree study programme in field: Informatics
Supervisor: Dr. Marián Šimko, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

satisfaction with the service offered. Users of our service allow us to collect data that will be able to analyse further.

2 Related works

The problem of diacritic reconstruction was tackled in [1], where Czech texts are reconstructed. The author used a method independent of context. The created program finds all possibilities of reconstructed word in a vocabulary. In the vocabulary is for every word assigned the number that represents how often the word was occurred in training corpus. The most occurred word is chosen.

In [8], diacritic is reconstructed for mobiles devices. The created application does not use context too. It reconstructs words alone. The problem in mobile devices is their memory. Memory in mobile devices is much smaller than in table computers and language model needs a lot of memory to use.

The two mentioned papers use methods independent of context. Antonín Zrustek compared all three methods for diacritics reconstruction [7]. Methods independent of context yield 96.6 %, accuracy, methods depending on the context yield 94.2 % and combination of these methods yields 97.3 %. Each number represents how many words were reconstructed correctly.

Jakub Vrána in his thesis proposed a method using context [6]. Using n-grams and Viterbi algorithm he is looking the most likely occurring sequence of words. The proposed methods yields 97.4 % accuracy.

Staš, Hládek and Juhár [2] do not deal only with reconstruction of diacritics but also with overall correction of Slovak texts. In a same way as a spellchecker tool, they correct misspelled words. They use the Viterbi algorithm and a hidden Markov model to find the best possible sequence of correct words.

There is a web service for diacritic reconstruction available online, which was created by Ľudovít Štúr Institute of Linguistics Slovak Academy of Sciences [3]. They allow the user to select different ways of diacritic reconstruction. The success of this web service using trigrams is 99.8 %. This web service achieves very good results but it is not integrated for web browser in the form of support tool for users.

None of the mentioned applications achieve satisfactory results except the last mentioned. Results obtained by the last web service are very promising. Other application achieve less than 97.4 % accuracy. This means 3 out of 100 words are reconstructed incorrectly. Incorrectly reconstructed words reduce text quality.

3 Diacritics reconstruction

3.1 Description of the algorithm

In this paper we present a method for diacritic reconstruction based on a statistical language model. Application uses 3 different language models[1]. First language model was trained on journalistic texts, art texts and scientific texts (33.3 % journalistic, 33.3 % art, 33.4 % science). It contains 317,469,718 unique tokens. Second was trained on newspaper articles and it contains 540,812,859 tokens. The last model was trained on texts from linguistic magazines and contains 773,493,137 tokens. This language models are free for use.

Some words can have more than one way of diacritics reconstruction. The problem is how to choose the correct form of the word. Our approach is based on the algorithm of Ludovit Stur Institute of Linguistics Slovak Academy of Sciences [4].

1. Convert all tokens to lowercase

[1] Language models are available here http://korpus.juls.savba.sk/prim(2d)6(2e)1.html

2. Find all possibilities of diacritic reconstruction for every single word
3. For input text take all possible n-tuples and evaluate them
4. Choose the highest evaluated n-tuple
5. Reconstruct the case of the characters

In the first step, for each token in the text we will find its variations of diacritic reconstruction in corpora. All possible variations of word reconstruction we found in word form databases. The token is first converted to lowercase. The algorithm ignores sensitive case. We pass the input text after the tuple of tokens. This will create all possible n-grams of n-tuple tokens. All of these n-grams are evaluated by the language model and the n-gram with highest occurrence in language model is selected. We continue processing the next tuple in the text. Finally, we redesign the case of the characters from the original text.

When computing probability during evaluation in step 3, the language model may sometimes give zero probability for an n-gram. This is due to the fact that the n-gram we are looking for did not occur in training data. Therefore, we are reevaluating zero probability n-grams, and assigning them non-zero probability. This technique is called smoothing [5]. We employ smoothing called Witten-Bell discounting.

3.2 Web browser extension

In order to bring advantages of diacritics reconstruction to a user of the Web, the algorithm will be available as an extension for web browser because its use is quick and easy. After installation, a user is able to reconstruct diacritic in arbitrary selected text.

The architecture of our application is shown in Figure 1.

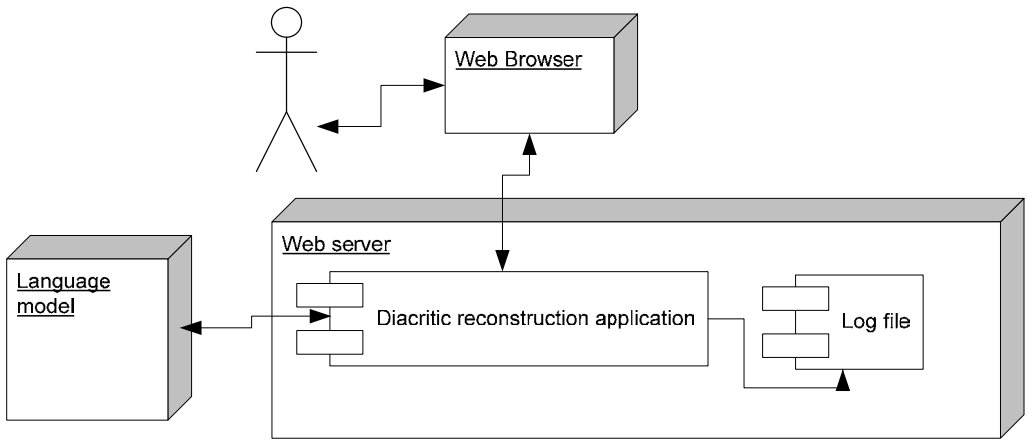


Figure 1. Architectural design of application.

When using our extension, not all words in the text had to be reconstructed correctly. User can correct bad reconstructed word and send us a log with correct correction. Bad reconstructed words are part of n-grams which evaluated probability in a language model was low. User's correction is logged and data are processed. We can use provided log in order to expand our database of words. There is a chance that word was bad reconstructed because the correct form was not in vocabulary. Therefore it could not be evaluated by a language model correctly. The question is how reliable are corrections made by user? In theory, we can include correct or incorrect corrections too. When we will reconstruct text later then for the logged word we will find more possible ways of diacritic

reconstruction in vocabulary and language model will evaluate more possible n-grams but it still will choose the most probably n-gram for reconstruction.

Another aspect of our solution – and an important contribution of our work – is data collection from users of web browser extension. Using data collected from users we are able to know which type of texts they reconstructed the most. Collected data are texts that were reconstructed by user. This text can be analyzed. Then we can extend the language model and increase the accuracy of diacritics reconstruction further, making the diacritics reconstruction continually improving.

4 Evaluation

4.1 General experiment

We have performed two types of experiments. Results how successful application was in different language styles are reported in the Table 1.

Table 1. Success of application used for different language styles.

Language style	Number of words	Number of errors	Successful
Newspaper style	862	12	98.6 %
Speaking style	1616	16	99.0 %
Scientific style	2534	78	96.9 %
Poetic style	2514	173	93.1 %
Colloquial style	4215	58	98.6 %

We have made test on different types of texts. So we tested if used language models are suitable and if our application is suitable for general and wide use. Newspaper style was represented by articles published in newspaper. Scientific style was represented by article about medicine. In scientific style, incorrectly reconstructed words that are very specific to that field of medicine were present. The worst results were achieved in poetic style. This style was represented by Slovak poems Mor ho! or Žltá ľalija. Poetic style contains poetic words that do not use in daily communication. The average successful on the whole test set is 96.7 %. However, if we exclude poetic style percentage increases to 98 %. Tests were made using 3-gram language model.

4.2 Experiment with users

In the second experiment, we chose 5 users and let them to select any text. In this text was reconstructed diacritic. Results are shown in the Table 2.

Table 2. Success of application on user's selected text.

User	Number of words	Number of errors	Successful	Satisfaction
1	145	1	99.3 %	10
2	177	3	98.3 %	9
3	254	2	99.2 %	8
4	244	2	99.2 %	9
5	138	1	99.3 %	9

Column Satisfaction represents how user was satisfied with the application on a scale from one to ten. Ten means he was very satisfied, one means he was not satisfied. The average accuracy of diacritics reconstruction on texts chosen by users was 99.1 %. User 1 chose educational text. Users 2 and 3 chose as a sample a post on the forum. Reading the post with diacritic was more comfortable for them. Users 4 and 5 chose their own emails. Success of diacritic reconstruction on emails was 99.2 %. They were very satisfied with reconstruction. We assume that our application will be the most intensively used as extension for web browser when writing or reading emails and also when reading some posts on forums written without diacritic.

5 Conclusions

In this paper, we have described our method for diacritics reconstruction. We utilize context to choose correct form of words with more than one way of diacritic reconstruction. Our application surely finds use for other problems of natural language processing such as text categorization or metadata extraction. Therefore, it is important that we achieved good results.

Our method will be tested on a greater sample than was presented. In our future work, we also plan to include tests on short texts without context.

We hope that our application in the form of web browser extension will be useful for users when writing emails or reading some text without diacritic on the Web. Text with diacritic looks more professionally and is more comfortable for reading. If the user decides to fix word manually we can this fix use in the next text reconstruction and reconstruct text better.

Acknowledgement. This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG 1/0646/15.

References

[1] Dušek, B.: *Diakritizace českého textu (in Czech)*, B.C. Thesis, Informatic, Charles University, Prague, (2010).

[2] Hládek, D., Staš, J., and Juhár. J.: Unsupervised spelling correction for Slovak. In: *Advances in Electrical and Electronic Engineering*, (2013), vol. 11, no. 5, pp. 392–397 .

[3] Jazykovedný Ústav Ľudovíta Štúra: *Slovak National corpus [online]*, Available: *http://korpus.juls.savba.sk/*.

[4] Jazykovedný Ústav Ľudovíta Štúra: *Slovak National corpus [online]*, Available: *http://korpus.sk/diakritik.html*.

[5] Jurafsky, D., Martin, J.: *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistic and Speech Recognitio.* New Jersey: Prentice Hall, Pearson Education, (1999), pp. 189–210.

[6] Vrána, J.: *Obnovení diakritiky v českém textu (in Czech)*, M.S. Thesis, Informatic, Charles University, Prague, 2002.

[7] Zrustek, A.: *Doplňování diakritiky do českých textu s chybějící diakritikou (in Czech)*, M.S. Thesis, Informatic, Masaryk University, Brno, (2000).

[8] Železník, J.: *Doplňování diakritiky pro mobilní zařízení (in Czech)*, B.C. Thesis, Informatic, Charles University, Prague, (2010).

Recognition of User's Identity or Activity by Analyzing Sensor Data from Smart Glasses

Marek HLAVA*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
xhlava@stuba.sk

Abstract. In this paper, we describe results of preliminary experiments, which support the idea that it is possible to distinguish between users according to their specific movements – head movements and eye blinks – which are tracked by Google Glass sensors – the accelerometer and proximity sensor. Data from the accelerometer has very low ability to distinguish users, moreover, it has greater variation in time. Eye blinks are more stable and therefore more clearly identify the user (51 %). As for different user recognition, better results are obtained from the activity of reading (86 %) than from gaming (69 %). Determination of the activity did not achieve as good results as the user identification.

1 Introduction

There are many different devices with various sensors. They can be used to track an acting individual wearing them. We can obtain their biometric and other characteristics. Biometric ones are often used to verify the user [6], to identify them among a small group of other users [2]. Other characteristics can help identify e.g. the user's activity [5] or even to identify user's emotional state [3]. Such devices can be a mouse and a keyboard [2, 3, 6] or e.g. Google Glass [5]. To look further, these characteristics can become a part of the user model or a domain model. Such models can improve e.g. personalization on the web. Personal computers along with a mouse and a keyboard and later on mobile phones became a common part of our lives. Now it is the age of mobile devices with a host of other sensors. We want to explore Smart Glasses. Such a device creates new possibilities for monitoring actions of a user, in particular whether their explicit or implicit feedback. Inspired by Shoya Ishimaru et al. [5] we decided to closely explore Google Glass' potential to identify a user or their activity. Similarly to as [5] we chose two sensors: the accelerometer and proximity sensor.

2 Related Work

Ishimaru et al. [5] have demonstrated how information about eye blink frequency and head motion patterns derived from Google Glass sensors can be used to distinguish different types of high level

* Bachelor study programme in field: Informatics
Supervisor: Dr. Alena Kovárová, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

activities. Their aim was to show that eye blink frequency data from an unobtrusive, commercial platform which is not a dedicated eye tracker is good enough to be useful and that adding head motion patterns information significantly improves the recognition rates. They evaluted this method on a data set from an experiment containing five activity classes (reading, talking, watching TV, mathematical problem solving, and sawing) of eight participants showing 67 % recognition accuracy for eye blinking only and 82 % when extended with head motion patterns.

Bacher et al. [4] have studied the factors conditioning changes in the frequency of spontaneous eye blinking (SB). They described that eye blinking is related to two personality traits. Adults with higher emotional stability exhibited lower levels of SB than adults who rated themselves as lower in emotional stability. Adults higher in agreeableness responded with reductions in the frequency of blinking across the memory tasks, whereas adults lower in agreeableness exhibited increasing blinking across the tasks. Adults reporting stress the day of the observation had higher rates of blinking than those reporting no stress. They did not observe sex differences in blinking for adults or infants.

Barbato et al. [1] have studied the relationship of Neuroticism, Extraversion and Psychoticism as assessed by Eysenck Personality Inventory (EPI) with spontaneous eye blink rate. A total of sixty-three healthy subjects (40 females, 23 males, mean age 24.2) were studied. Spontaneous blink rate and time of blink suppression were assessed by EOG measurement. Levels of Extraversion and Neuroticism were inversely correlated. They found significant correlation between blink rate measures and Neuroticism. No significant correlation between blink measures and either Extraversion, or Psychoticism were found.

3 Experimenting With User Activity Recognition

3.1 Methodology

We conducted our first experiment with 5 users (User1, User2, ... User5), university students, 2 men and 3 women. The users carried out four different activities in front of personal computer wearing Google Glass. Each activity took exactly 5 minutes and under the same lighting conditions. The activities were:

1. Just sitting (Nothing)
2. Reading set news from a set news portal, same for all users (Reading)
3. Chatting with a friend about recent events, different for every user (Chatting)
4. Playing a simple "rush" game Despicable Me: Minion Rush,[1] same for every user (Gaming)

The head movement via accelerometer (accel.) and eye-blinks via proximity sensor (proxim.) were monitored during 5 minutes for every activity. We averaged accel data for each half minute and subsequently calculated standard deviation of this movements for every average. To detect eye-blinking we used the application [5] to count number of blinks for the same time step – 30 seconds. This way we received two times 10 values for every user's activity. To calculate probability of correct identification, we divided randomly 100 times the data into 5 and 5 values. The first five served as a sample and the rest 5 served as a test data. The comparison between each sample data and test data was calculated by t-test similarly to [2].

[1] Despicable Me: Minion Rush, http://apps.microsoft.com/windows/sk-sk/app/despicable-me-minion-rush/ce53ce3b-63a9-4b89-8c0f-df816f6ca560

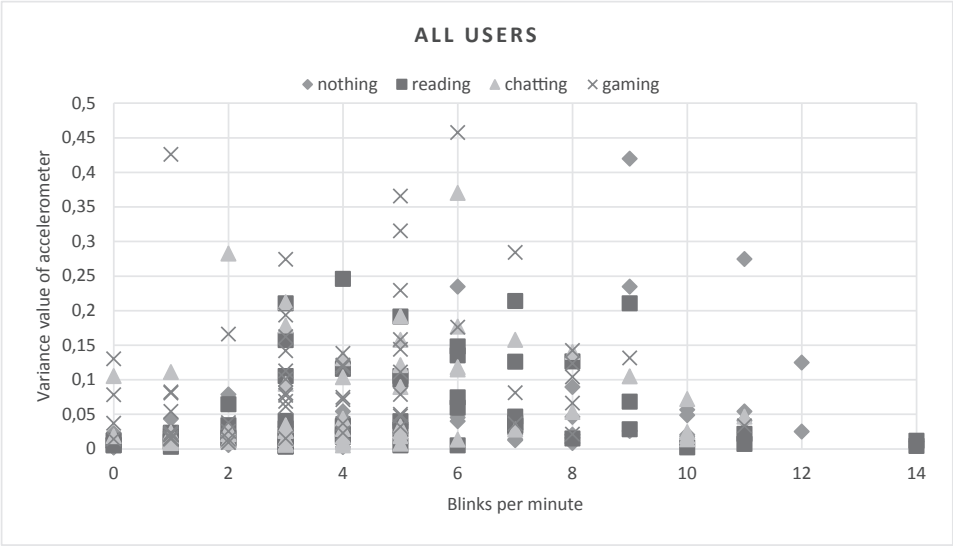


Figure 1. Dependence of the average deviations head movement on number of blinks per minute during different activities depicted for all users.

Table 1. Probability of correct identification of the user.

activity	nothing	reading	chatting	gaming	all
accel.	0.24	0.26	0.64	0.31	0.24
proxim.	0.33	0.62	0.53	0.48	0.51
both	0.44	0.71	0.71	0.56	0.53

3.2 Results

Results of our experiment are in Figure 1 and 2, depicting 12 charts. To make graphs more intelligible, we calculated the values from one minute long time steps instead of 30 seconds. Graphs in Figure 1 display users and graphs in Figure 2 activities, both from the same data set. Already on the graphs themselves can be noticed that different people performing different activities react differently. Simply by visual comparison of "All users" and "All activities" one can see that the probability of correct user identification is higher than correct activity identification. Therefore we skipped results for the activity identification. We also did not compare with Ishimaru et al. [5] since some of their activities engaged the whole body.

3.3 Evaluation and Discussion

Table 1 contains probabilities of correct identification of the user, where our data were 100 times randomly divided in two subsets – sample and test. Since we had 5 users, the probability close to number 0.2 means that it is not possible to distinguish the users. The higher the probability is, the greater the ability to recognize the user. The power of differentiation between users is higher for proxim. (count of blinks), but the best results are obtained by combining results from both sensors. The probability of identifying the correct user without taking into account the activity is 0.53.

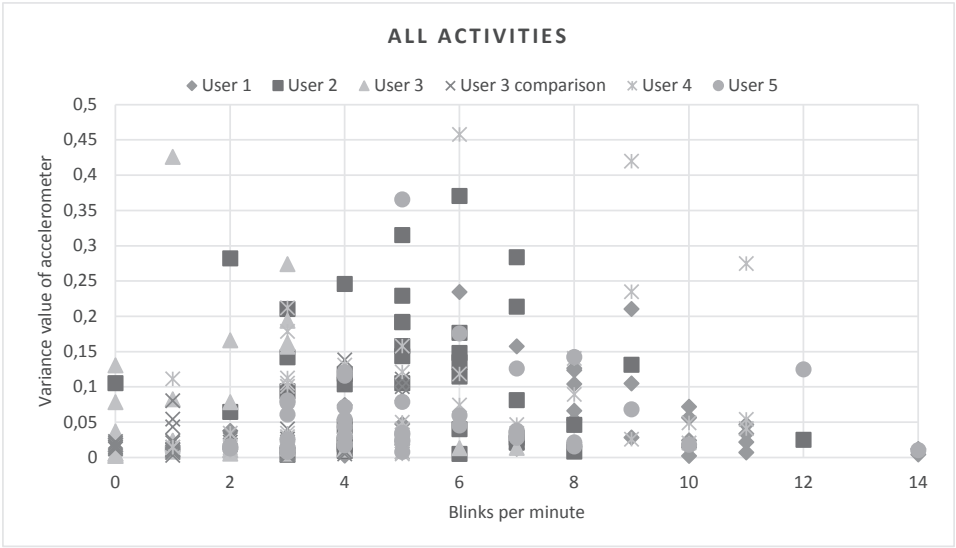


Figure 2. Dependence of the average deviations head movement on number of blinks per minute for all users depicted for all activities.

4 Experimenting With User Identity Recognition

4.1 Methodology

In the second experiment we tried to find out if it is possible to determine if the user wearing the Glass is the owner or a different person. In the light of the results of the first experiment, we conducted our second experiment with a larger group of users, testing only the activity of reading and gaming. There was a total of 14 users (9 men and 5 women). Similarly to the first test, we have collected 10 values from the accelerometer and 10 blink counts per test per user. The samples were then compared to each other using Welch's Two Sample t-test with 95 percent confidence interval. If p-value of the test is lower than 0.05 the samples are from different users.

4.2 Results

Figure 3 shows the dependence of head movements and number of blinks for all users.

4.3 Evaluation and Discussion

Table 2 contains percentage of different user identification for both activities. We compared each user to 13 other users and counted how many of them are recognized as different users. 100 % means 13 out of 13 comparisons were successfully recognized, 0 % means none. The results show that we were able to reject up to 100 % of other users combining data from the accelerometer and blink counts. On the average, better results are obtained from the activity of reading (avg. 86 % from both sensors) than from gaming (avg. 69 % from both sensors). Using blinking only (avg. 63 % from both tests) is better than using accelerometer only (avg. 38 % from both tests).

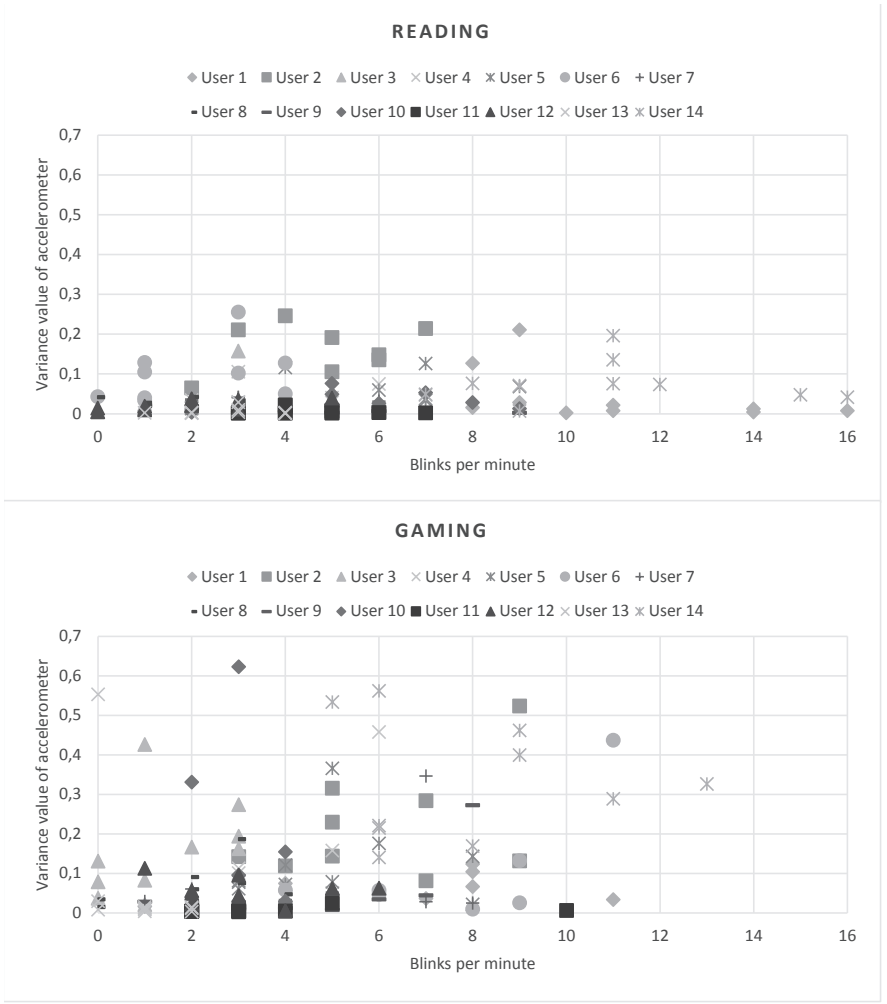


Figure 3. Dependence of the average deviations head movement on number of blinks per minute for all users depicted for both activities.

5 Conclusion and Future Work

The results show that while the head movement and eye blinking frequency may have good potential to identify the user (51 %), they are even better for determining if two different users are wearing the device (78 %). As for different user recognition, better results are obtained from the activity of reading (86 %) than from gaming (69 %). That means that users' reading patterns are more different than their gaming patterns, analyzing that could be an objective of a future work.

We did not compare activity recognition with Ishimaru et al. [5] since some of their activities engaged the whole body. Our tests were situated in front of a computer where there is little to no head movement, which reduce our activity recognition success rate.

In the future work we plan to conduct an experiment with different activities and compare our results with recognition results gained by mouse and keyboard [2].

Acknowledgement: This work was in part supported by the Scientific Grant Agency of the Slovak Republic, grants No. VG 1/0752/14 and VG 1/0646/15.

Table 2. Percentage of different user identification.

reading	**accel.**	**blinks**	**both**	**gaming**	**accel.**	**blinks**	**both**
P01	8 %	92 %	92 %	P01	31 %	31 %	54 %
P02	77 %	54 %	92 %	P02	54 %	54 %	77 %
P03	23 %	69 %	77 %	P03	46 %	85 %	92 %
P04	31 %	54 %	77 %	P04	8 %	46 %	46 %
P05	0 %	62 %	62 %	P05	15 %	46 %	54 %
P06	62 %	69 %	100 %	P06	15 %	77 %	85 %
P07	46 %	69 %	92 %	P07	15 %	39 %	54 %
P08	46 %	77 %	85 %	P08	31 %	77 %	92 %
P09	62 %	62 %	85 %	P09	23 %	39 %	46 %
P10	46 %	62 %	85 %	P10	0 %	46 %	46 %
P11	62 %	54 %	85 %	P11	62 %	39 %	69 %
P12	46 %	69 %	85 %	P12	31 %	54 %	62 %
P13	62 %	69 %	92 %	P13	15 %	92 %	92 %
P14	62 %	92 %	92 %	P14	85 %	77 %	92 %
AVG	**45 %**	**68 %**	**86 %**	**AVG**	**31%**	**57 %**	**69 %**

References

[1] Barbato, G., della Monica, C., Costanzo, A., Padova, V.D.: Dopamine activation in Neuroticism as measured by spontaneous eye blink rate. *Physiology & Behavior*, 2012, vol. 105, no. 2, pp. 332–336.

[2] Chudá, D., Krátky, P.: Usage of Computer Mouse Characteristics for Identification in Web Browsing. In: *Proceedings of the 15th International Conference on Computer Systems and Technologies*. CompSysTech '14, New York, NY, USA, ACM, 2014, pp. 218–225.

[3] Epp, C., Lippold, M., Mandryk, R.: Identifying Emotional States Using Keystroke Dynamics. In: *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems (CHI 2011)*, Vancouver, BC, Canada, 2011, pp. 715–724.

[4] Hubbard, L.A., Lewis, K., Wallace, K., Wynkoop, K., Norton, J., Bacher, L.F.: Factors contributing to changes in the frequency of spontaneous eye blinking in infants and adults. `http://www.oswego.edu/~bacher/Research/EPA%202006%20SB%20poster.pdf`, 2006, [online; accessed 2-Dec-2014].

[5] Ishimaru, S., Kunze, K., Kise, K., Weppner, J., Dengel, A., Lukowicz, P., Bulling, A.: In the Blink of an Eye: Combining Head Motion and Eye Blink Frequency for Activity Recognition with Google Glass. In: *Proceedings of the 5th Augmented Human International Conference*. AH '14, New York, NY, USA, ACM, 2014, pp. 15:1–15:4.

[6] Zheng, N., Paloski, A., Wang, H.: An Efficient User Verification System via Mouse Movements. In: *Proceedings of the 18th ACM Conference on Computer and Communications Security*. CCS '11, New York, NY, USA, ACM, 2011, pp. 139–150.

Extracting Keywords from Movie Subtitles

Matúš KOŠÚT*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
matuskosut@gmail.com

Abstract. In our work we aim at keyword extraction from movie subtitles. Keywords and key phrases although missing the context can be found very helpful in finding, understanding, organising and recommending the content. Generally they are used by search engines to help find the relevant information. Movies and video content are becoming massively available and widespread. The ability to automatically describe and classify videos has a vast domain of application. In our work we aim at movie subtitles as a source of information, which seems to be more efficient compared with video and audio analysis. The main goal of our work is to design a method able to use of specifics of subtitles. First part of method focuses on pre-processing. Pre-processing tries to process timing, information for hearing impaired persons and tags included in subtitles. Second part divides subtitles into conversations according to the speed of speech (words per minute) and the gaps detected between the conversations. Scored conversations are used for keyword extraction. By this we create the set of keywords that can be used by recommending and search engines.

1 Introduction

In this work we deal with keyword extraction from movie subtitles. The importance of having video data annotated and ordered grows exponentially with rising amounts of data created every day. It became almost impossible to assign necessary keywords to all of the movie data, because of a time-consumption. Although a research about video and audio analysis advanced recently, it is still very computationally intensive to process all of the data and the results are not always as expected. It is good to consider alternatives, especially when the subtitles are available. Original subtitles are usually provided with movies on the media (e.g. DVD, Blu Ray, etc.) and also a lot of educational video content available online is available with subtitles provided either by tutors or community. Keywords extracted from the subtitles could be later utilized in movie recommendation systems.

As we know, there still is not a lot of works focused on processing subtitles as a source of information. Small projects usually use keyword extraction algorithms designed to work with pure text, which they obtain from subtitles without using any other metadata. Our aim is to use these specifics of subtitles to improve keyword extraction.

* Bachelor degree study programme in field: Informatics
Supervisor: Dr. Marián Šimko, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 20–24.

2 Related works

One of the related works, which goes further in the subject, is VIRUS (Video Information Retrieval Using Subtitles) [1]. Main goal of this work is to provide users with movie scenes selected on the basis of specific characteristics (e.g. love scenes, violent scenes, funny situations) [1]. They use the simultaneous analysis of video, audio and subtitles to achieve their results. Video is analysed only for simple patterns, motions and colour histograms while audio is processed more deeply in order to search for specific situations of screaming, gunshots and similar noises in addition to subtitles keyword extraction. This concept of combining especially audio and text is believed to result in more precise details of scenes, assuming that "murder" mentioned in subtitles doesn't just mean it is relevant fact unless we are provided with the fact of shooting as a result of audio analysis. They believe this approach to semantic analysis is innovative and maximizes the quality of search results.

Another quite interesting work in this subject is Semantic Video Classification Based on Subtitles [2]. The work although using keyword extraction from pure text of subtitles is noteworthy because it focuses on the classification of their ample description of the algorithm. The authors deal with the design and issue of the service which could define a movie category from a given subtitles. The proposed method [2] uses an algorithm TextRank with its default settings, as suggested in the work [3]. Then they are trying to assign a category, to keywords extracted in previous step, using WordNet domains by mapping the individual category transferred to a specific ranked categories that are ordered by descending and the most successful is selected.

As mentioned before, standard keyword extraction methods such as TF-IDF, RAKE [5] and TextRank [3] with its modifications are used really for the purpose of extraction without consideration of subtitles specifications, timings and closed captioning, which we consider to be a main deficiency of related works.

Therefore, we want to use these specifics of subtitles to split them into conversations, to rate them, to extract the closed captioning and to evaluate their impact on the extraction.

3 Keywords extraction from movie subtitles

Subtitles contain pure text apportioned to individual titles with the timings of when and how long to display. Sometimes, in addition to dialogues, subtitles contain additional information, closed captioning, to help disabled people (deaf or hearing impaired) recognize the sounds in the scene and sounds in the background or to help those who do not understand the spoken language or have troubles recognising it because of the accent used by actors.

3.1 Conversations

The first idea to enrich the keywords extraction with metadata included in subtitles is to recognise individual parts of movie. We call them conversations. We suppose that we could divide dialogues in subtitles into conversations and get the scenes severally, approaching the natural distribution of scenes as it is seen by viewers. In the sense of when there are different characters talking in dialogues or scene is changed we would start a new conversation.

We want to explore possibilities for rating these conversations according to relevance. The higher the relevance is, the higher the ratings of keywords extracted from the conversations become.

These conversations could also possibly help us with joining subtitles created by different authors in case we need to get more information about the scene that would result into more keywords, assuming different authors use different sentences and on condition that it is not a transcript.

To split the subtitles into conversations, we use timings included in subtitles. Using timings we want to detect the gaps between individual subtitles, supposing that if there is a gap bigger

than are the gaps in the surrounding titles, the conversation has changed into next one. We are also experimenting with the speech rate (words per minute) in individual titles and conversations as way to differ the conversations and to rate them supposing there is a relation between the speed and the importance of conversation.

3.2 Hearing impaired subtitles

Subtitles for hearing impaired viewers are a special category of subtitles created to help disabled people understand of what is happening in the scenes. Based on our experience with various types of subtitles we concluded that subtitles for hearing impaired viewers contain descriptions of the most necessary sounds from scenes and backgrounds. We propose to use these metadata to substitute audio analysis. We assume precision increase of the results, because of a better representation of scenes and connecting the sounds with surrounding keywords is a sign of their importance [1].

We also experiment on speech rate (words per minute) and subtitles rate (titles showed per minute) as a helper tools for a stress, action or aggression detection in scenes.

3.3 Algorithm

Our algorithm to keyword extraction from movie subtitles consists of the following steps:

1. Subtitles pre-processing
2. Conversations
3. Filtering stop words
4. Keyword extraction
5. Processing ratings

Subtitles pre-processing. Pre-processing consists of processing information about times of individual titles, counting the speech rate and gaps between them. Irrelevant symbols and subtitles specifics, such as music symbols and colour tags, are filtered out. Text inside hearing impaired tags and html tags is labelled. Then all the text in titles is segmented tokenized and POS tagged.

Conversations identification. We experiment with subtitles splitting into conversation using gaps between individual titles and conversation rate. We suppose that the gaps between the titles of different conversations are wider than those of the same conversation. We also suppose that speech rate of titles differ between conversations according to characters, intensity and tense of conversation. We experiment on ratings using the location, length and speech rate of conversation and filtering out unimportant ones.

Filtering stop words. We are using Rake Text Smart List [5] of stop words to filter out unnecessary and irrelevant words from the text of subtitles.

Keyword extraction. We are using standard keyword extraction techniques in our method to retrieve the ratings of words. These words are usually ordered and the top *n* of them are chosen as keywords, but in our method we also include another step to take account of conversation ratings and closed captioning, subtitles for hearing impaired people, so we could evaluate the assets of our method. We experiment not only with the extraction from whole text but also with the individual conversations used for extraction, which results in more sets of keywords for one subtitles to be processed.

Processing ratings. Ratings of conversations are used with closed captioning to calculate the final ratings of words. We experiment with combination of these ratings and joining the conversations results from previous step. We also suppose the words with general meaning should be underrated because they are not describing the content well and for the same reason we are preferring nouns and verbs over other words.

4 Evaluation

We have already implemented a major part of an environment for experimentation with our keyword extraction method. We have implemented standard methods for keyword extraction from pure text and atop of them we have implemented features we proposed. We are working with conversation splitting and closed captioning.

Our idea is to use standard methods enriched with our features, which are extending these standard methods by influencing the results of keyword extraction, so we can evaluate our results by comparing them with the results of standard methods, which are also used by related works.

For now we have only evaluated standard keyword extraction methods that can be seen in Table 1. We plan to use them with OpenSubtitles dataset[1] to experiment with the settings of conversation splitting and rating of conversations and also the extraction and consideration of closed captioning in order to explore the effect of our features on the final results of keyword extraction.

Table 1. Evaluation of TextRank and TF-IDF [5] extraction from pure text from subtitles.

	TextRank		TF-IDF	
	Average	Best	Average	Best
Precision	17.86	29.41	19.11	32.0
Recall	36.49	60.0	38.71	64.0
F-Measure	23.98	39.47	25.59	42.67

Evaluation was done by comparing the subtitles keyword extraction results with the keywords obtained from movie database[2]. The subtitles for keyword extraction were chosen from an online list of most favourite movies. Evaluation results can be seen in Figure 1.

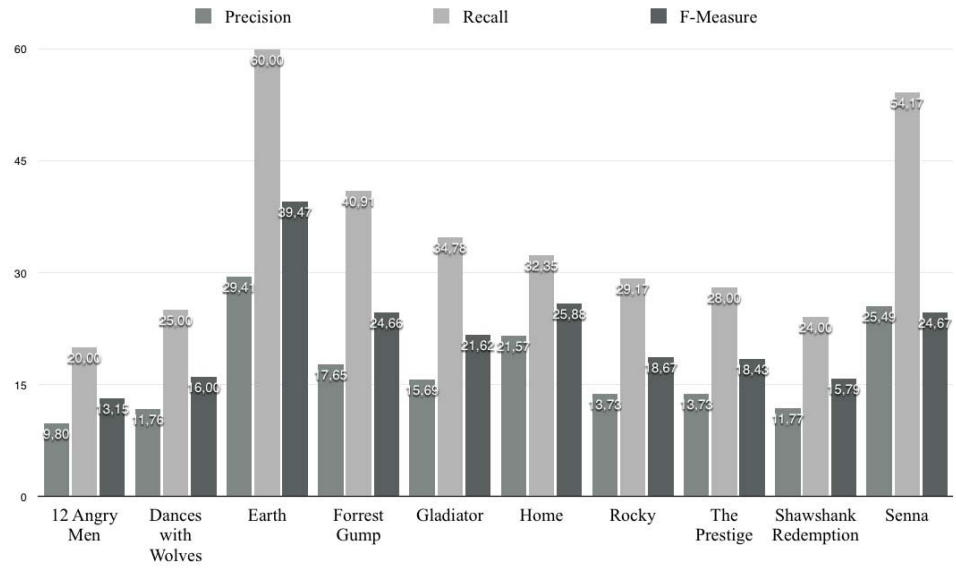


Figure 1. Evaluation overview of TextRank performance with subtitles.

[1] http://opus.lingfil.uu.se/OpenSubtitles2013.php
[2] http://www.moviecus.com/

We did not include RAKE algorithm [5] into our evaluation, because this algorithm extracts only key phrases which although may be pretty interesting to explore, but we did not have the phrases dataset to which we could evaluate our results against and the results of splitting the phrases into keywords were not sufficient.

5 Conclusions

In this paper we tackled the issues of metadata extraction from movie subtitles. Metadata such as keyword are useful for providing intelligent services such as movie recommendation. However, not much work is done in this field, we are aware of only a limited number of works.

The contribution of this paper is our subtitles keyword extraction method which enriches standard methods for keyword extraction. We proposed the detection and rating of conversations to be merged with standard methods and thus influence the rating of extracted keywords to achieve more appropriate results. We also proposed extraction of closed captioning (subtitles for hearing impaired) as a substitution for audio analysis in order to get more detailed view of scenes which can be connected with extracted keywords not only to improve results of keywords extraction itself but also to provide additional information which may be interesting for recommendation systems.

Acknowledgement: This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG 1/0646/15.

References

[1] Langlois, T., T. Chambel, E. Oliveira, P. Carvalho, et al.: VIRUS: video information retrieval using subtitles. In: *Proceedings of the 14th International Academic Mind Trek Conference: Envisioning Future Media Environments, Tampere, Finland*, ACM, (2010), pp. 197–200.

[2] Katsiouli, P., V. Tsetsos and S. Hadjifthymiades: Semantic Video Classification Based on Subtitles and Domain Terminologies. In: *Proc. of the KAMC*, http://ceur-ws.org/Vol-253/paper05.pdf, (2007).

[3] Mihalcea, R. and Tarau, P.: *TextRank: Bringing order into texts. In: Proceedings of EMNLP, Association for Computational Linguistics*, Barcelona, Spain, (2004).

[4] Hasan, K.S. and Ng, V.: Conundrums in unsupervised keyphrase extraction: making sense of the state-of-the-art. In: *Proceedings of the 23rd International Conference on Computational Linguistics*, Beijing, China, (2010), pp. 365–373.

[5] Dagan, I. and K. Church, B.P.: Termight: identifying and translating technical terminology. *In: Proceedings of the Fourth conference on Applied natural language processing*, Stuttgart, Germany, (1994), pp. 34–40.

[6] Kogan, B.: Text mining: Applications and Theory. In: *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, Barcelona, Spain, (2010).

Touchscreen Authentication Using Keystroke Dynamics and Other Touch Based Features

Lukáš MIŠKOVSKÝ*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
miskovskylukas@gmail.com

Extended Abstract

Keystroke dynamics is relatively old research subject. We can distinguish two basic areas in the field of keystroke dynamics and in the biometric systems in general. Identification and authentication. The former represents a recognition of one person, among many others. The latter means being able to tell if a given biometric sample belongs to a given person, provided that the system has a previous data on which it can base the decision.

Smartphones still remain the least secure platform with most of our vital information stored in them. Main objective of our paper is to address this problem and merge known research in keystroke dynamics with other touch based features to strengthen smartphone security during authentication. One of the main advantages of keystroke dynamics is that this process is not intrusive which is important in a day to day smartphone usage.

You can see the summarized comparison of different research teams in the Table 1. Hold time and inter-key time (also known as flying time) are in the mentioned table referenced simply as "time intervals". Success of different teams is measured in FAR (false acceptance rate) and FRR (false rejection rate) or EER (equal error rate).

The proposed system, enhancing authentication with keystroke dynamics, consists of three major phases. First one is the enroll phase, where the user repeatedly inputs his login credentials into our system. After that a template for this user is created using the collected characteristics. Now his login process can be enhanced with keystroke dynamics and while he is logging in, the system evaluates if he is eligible.

Success of a system based on keystroke dynamics highly depends on captured features. Because of this, we decided to use more features in our system and find out their impact. Our goal was to facilitate the use of touch screen display on modern smartphone devices to consider other promising features for biometric authentication. You can read more about established features such as key hold time or flying time in [1–3]. We will describe new or unusual characteristics below.

---

* Bachelor study programme in field: Informatics
Supervisor: Assoc. Professor Daniela Chudá, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 25–26.

Table 1. Comparative table of other research teams using keystroke dynamics for biometric authentication.

Team & year	Keyboard type	Measured features	Method	Accuracy
Tojahn a Ortmeier 2012 [3]	touch screen	time intervals, pressure, touch area	statistical	12.13% FAR 8.75% FRR
G. Kambourakis et al. 2014 [2]	touch screen	time intervals, distance, speed	machine learning	12.5% FAR 39.4% FRR
M. Antal et al. 2014 [1]	touch screen	time intervals, pressure, touch area	machine learning	12.9 – 15.3% EER

– *touch areas* – We are recording two areas in total. One during the first touch and the other one when a user is lifting his finger from the screen. That way we can tell if a user changed an angle of his finger while typing a letter.

– *touch distance* – This represents how far from the center a user touched the key. We normalized the distance to range $\langle 0, 1 \rangle$, 0 being the centre of a key and 1 being the outermost edge of it, to isolate the device characteristics such as screen size and resolution from the biometric nature of this feature.

– *touch sectors* – Imagine a key divided to four sectors. First one is in upper left corner and, going clockwise, the last one is in lower left corner. Appropriate number is then assigned based on touch location.

To capture these features we modified the Biosec Logger application. We had to implement our own virtual keyboard, to get to a more advanced features.

In our experiment, after getting familiar with our keyboard, the users were asked to repeatedly input two different phrases in order to create two separated datasets. First one was a familiar phrase "@gmail.com" that our test subjects should be used to write quite often as a part of an email address. Second one "falosnyklub" was devised to be long enough for a password yet simple enough to write easily.

First analysis, testing just one of our additional characteristics, shows promising results. Our next work will focus on more detailed testing of the obtained datasets.

Acknowledgement: This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG 1/0646/15.

References

[1] Antal, M., Szabó, L.Z., László, I.: Keystroke Dynamics on Android Platform. In: *8th International Conference Interdisciplinarity in Engineering*, Elsevier Ltd., 2014.

[2] Kambourakis, G., Damopoulos, D., Papamartzivanos, D., Pavlidakis, E.: Introducing touchstroke: keystroke-based authentication system for smartphones. *Security and Communication Networks*, 2014.

[3] Trojahn, M., Ortmeier, F.: Biometric Authentication Through a Virtual Keyboard for Smartphones. *International Journal of Computer Science & Information Technology (IJCSIT)*, 2012, vol. 4, no. 5, pp. 1–12.

Evaluating Learnability of Games

Peter DEMČÁK*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
piers.demcak@gmail.com

Abstract. In this paper, we propose a set of methods for the evaluation of gameplay learnability in games, based on mining of designer-prescribed behavioral models in a stream of gameplay events. These events are collected through implicit feedback and include gaze tracking, inputs, and logical events dependent on the context of the game.

1 Introduction

In this paper, we propose a method which aims to become a new useful tool for game development. However, as we will eventually show, our solution can be applied to support the development of any kind of conventional software.

The video game development process can usually be described by the spiral model of software projects. As a part of the development and testing phase, different prototypes of a game are being produced and subsequently tested for various qualities that the developers aim for – general usability of the game interface, performance, or just the plain entertainment value of the game. All of these properties – and others – are essential to create a working game, and must be evaluated in the course of the development in order to evade risks as soon as possible and to provide feedback for improvements in future iterations.

Playtesting [1] is a popular methodology, used to evaluate certain properties of a game. During playtests, the object of evaluation is the experience of the players itself. With a particular gameplay feature in mind, the playtesting players are let to interact with the game as naturally as possible. The head of the experiment has to pay close attention, as to what sort of information they provide the player with outside the game itself, so that the objective of the playtest is not compromised.

With the player experience being an inner phenomenon of the mind and it being difficult to measure overall, various techniques are used during playtests to collect feedback about the game experience from players. These techniques are namely: observation, think-aloud protocol and a questionnaire or an interview at the end of the playtest; often used in combination.

However, all of these techniques have certain shortcoming. Either they are so disruptive for the players, that we can no longer tell, if we are actually measuring the experience real players are going to have with our game, or they are incapable of collecting all of the information about the

* Master degree study programme in field: Information Systems
Supervisor: Dr. Jakub Šimko, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

player experience. These are the aspects of playtesting that our method of Learning Case Mining aims to improve.

The objective of our main method is to provide an additional source of feedback from playtesting, which diminishes the aforementioned shortcomings when integrated into the playtest. Our method is based on implicit feedback, namely gaze tracking, thus, it is less disruptive for the player and uses a different source of information about the player than just observation and explicit feedback.

Gaze tracking is a technology, which enables us to accurately determine what position on the screen the user is currently looking at. It is a relatively new technology, but it is rapidly becoming more widely available. The use of gaze tracking shows its promise in various fields such as marketing, health care [2], education [3], psychology, and also user experience. Usage of gaze tracking in any manner largely depends on the premise, that gaze fixations are synchronized with mental processing of the objects that are being gazed upon.

Our method further concentrates on evaluation of gameplay learnability. Classic learnability is an aspect of usability; the ability of software to be learned by its users. Learnability is especially crucial for games, because in order for games to be entertaining for the player, the player will have to be able to get a quick grasp of the basic game mechanics a dynamics.

2 Related work

Various studies have been performed on the subject of video games in combination with gaze tracking, but none are currently concentrating directly on using the technology to improve gameplay learnability.

Eivazi et al. [4] used gaze tracking to predict the mental states of puzzle solvers using a SVM. Their accuracy reached 53 %, which the authors consider to be improvable with the use of a more advanced machine learning approach. Bartels [5] used gaze tracking in an AB test of an airport simulator game to compare the cognitive workload with different versions of game interface, and to pinpoint critical areas of interest for the players. Kickmeier-Rust et al. [6] used gaze tracking in an educational game to find the correlation between the eye behavior and the improvement of the students' knowledge between the pre-test and the post-test. These are just several examples that show the contributions gaze tracking can bring for games.

In the area of learnability, we can turn to the survey performed by Grossman et al. [7] According to the survey, even in the human-computer interaction community the term of learnability is used loosely by various authors. If all the different definitions are considered, we can separate learnability into two categories:

- Initial
- Extended

In the context of games, we concern ourselves foremost with the initial learnability, as it is far more important for a game to be approachable by players from their very first session with the game, while the complete mastery and proficiency in the game (which are covered by extended learnability) can be let for the most hardcore type of players to discover on their own.

Learnability is usually not measured on its own, but as a part of general usability. For ways of measuring learnability, we can once again turn to Grossman et al. who summarize a collection of various learnability metrics:

- Metrics based on performance
- Metrics based on command usage
- Metrics based on cognitive processes
- Metrics based on explicit user feedback

- Metrics based on usage of documentation
- Metrics based on change of usability
- Metrics based on specific rules

Several of these metrics could be applied to games with only slight modifications. However, games are different from other types of applications in several aspects. While the work performance of the user with a different type of information system can be determined from the user successfully fulfilling a task, in a game, the player might not even know, that the task exists. While in other types of software, users use the software with a certain goal, the goal of playing a game is the entertainment of interacting with the game itself, and so the player needs to learn the elements of the game as well as how to use them.

More insight into what kind of gameplay elements players might need to learn is provided by Hunicke et al. [8], who proposed the MDA framework to better describe game design elements. They split games into mechanics, dynamics and aesthetics.

Mechanics represent the rules of the game, which define how the players can interact with the game and each other. Dynamics represent the game as a system of these rules and the emergent situations that can happen in the borders of these rules. Aesthetics represent the emotional response the game triggers in the player. Example: In a farming game, I feed a dog by pressing the feed button, which is a game mechanic. When I feed the dog, the dog is no longer hungry, it is happier, and guards the farm better from thieves, which is the game dynamic. The emotions of fulfilled challenge from taking good care of the dog and the positive feeling from seeing the dog happy form the game aesthetics.

While the game aesthetics do not need to be taught to the player, as they are the closest the goal of the player to enjoy the game, the player needs to learn the game mechanics and some of the game dynamics to be able to enjoy the game and achieve any game aesthetics. Based on this knowledge, we consider initial gameplay learnability to be connected with the players leaning basic game mechanics and game dynamics.

3 Method

In the course of each iteration, playtesting is performed for new gameplay elements, which are usually represented by new game mechanics and new game dynamics, which we discussed in the previous section. We propose Learning Case Mining as a method to be used for evaluation of the gameplay learnability. Learning Case Mining is to be used during the development and testing phase, simultaneously with common playtesting. It requires a testing device equipped with gaze tracking equipment and software.

Learning Case Mining is based on the principle, that each playtest has a hypothesis in the form of one or more game mechanics or dynamics, which need to be tested for whether the players enjoy them. Therefore, game designers are capable of describing the interaction, in which they plan the player to learn a certain game mechanic/dynamic. This type interaction is called a learning case.

To use our method, the game designer creates a model of the learning case describing the expected player behavior. This learning case is then compared with the data describing the actual behavior of the players during the playtest. Our method finds a behavior that was the most similar to the learning case prescribed in advance, finds the differences and aggregates the results with other players. To implement this method, we created the software tool Geddit.

3.1 Learning cases and logs

Above, we defined learning cases as the models of interaction between the player and the game, which serve to teach the player a certain type of game mechanic or dynamic. In order to compare them with the real run of the playtest however, the game designer needs to be able to both formally

describe learning cases, and to collect the data from the playtest in a fitting form in order to mine learning cases in them.

Let us consider an example, where we are trying to teach a player of a platforming game to bypass the enemies and the pits at the same time. Naively, this can be described as a simple sequence of actions. While playing the game however, players can have multiple choices for solving the situation at hand, all of them correct (decision). Furthermore, a situation can contain multiple problems, which have to be solved simultaneously (parallelism), or the player may need to repeat certain actions (cycle) in order to resolve the problem. Therefore, we formally model learning cases as activity diagrams (see Figure 1), which are capable of carrying this sort of information.

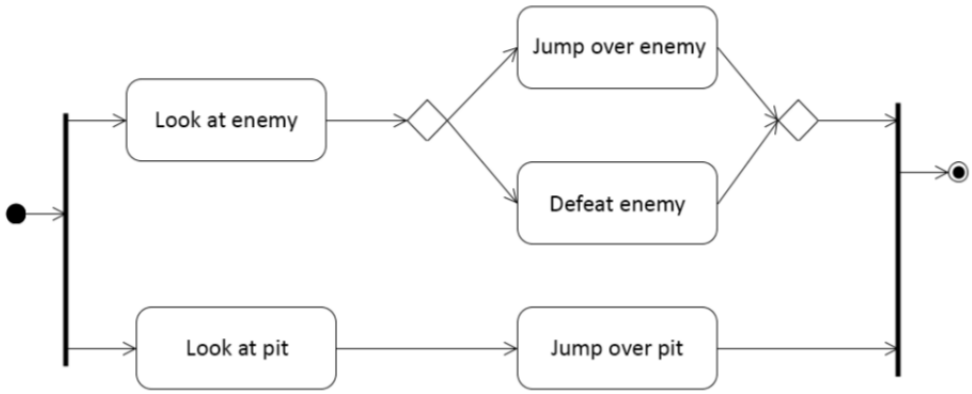


Figure 1. Sample learning case for a platforming game.

Once we have the learning cases at ready, we need to collect the stream of events that happen during the playtest. We have distinguished three types of events that can be observed automatically

- input event
- gaze event
- custom event

Input events represent inputs from a mouse, keyboard or a game controller. Gaze events represent fixations of gaze on game elements on the screen. Custom events can be used to describe events based on the internal logic of the game (example: The player has won)

To make use of the aforementioned events, game designers need provide Learning Case Mining with three sources of data. The gaze tracking data from the gaze tracking tool (for the development of the tool Geddit, Tobii Studio was used), logs about the changes of game elements on the screen and logs of custom game events.

3.2 Matching algorithm

Our matching algorithm has two inputs – the learning case and the stream of events from the playtest. As a preliminary step, the learning case is transformed into a set of all possible passes through its activity diagram. Then, the most similar pass is chosen by comparing it to the playtest stream of events as a sequence.

As the metric for comparing these sequences, we use the number of operations of a modified version of the Damerau-Levenshtein distance [9]. We count each type operation separately because of their different level of impact on the difference between our strings:

- Deletion – highest impact, an expected event is missing
- Transposition – medium impact, an expected event happened, but elsewhere
- Addition – low impact, the playtesting event stream is usually significantly longer compared to the learning case pass. Values too high can still mean a problem, if the player did not know what to do

In addition, the operations of deletion and substitution are considered identical for the same reason the operation of addition is considered of low impact.

After all three metrics are calculated, the optimum pass is chosen depending on the minimum number of operators by the order of their impact. Afterwards, the optimal passes found are mapped back onto the activity diagram, showing the game designers the locations of any learnability issues where the actions are missing, where they are placed in an unexpected order, or with too much different activity in-between.

3.3 Playtest design

When planning a playtest using Learning Case Mining for learnability evaluation, a game developer should thoroughly consider the choice of playtesting participants. Because the initial learnability is the object of evaluation here, all of the participants should come in contact with the tested game mechanics or dynamics for the first time during the playtest. Also, as with all playtests, unless the game is targeted at a certain demographic, the playtesters should represent both casual and hardcore types of players.

Because Learning Case Mining can be automatized, it can also be easily scalable even to a larger number of playtesters. The decision for the number of participants depends largely on the scalability of other techniques used during the playtest, although the empirically proven number of five participants is recommended as the bare minimum.

Before starting a large scale evaluation of learnability, it is also recommended to validate the design of the learning cases themselves, in a smaller-scale playtest.

4 Experiment

The Learning Case Mining tool Geddit was created for the verification of our method. Even during its development, several experiments with the Tobii eye tracking device were conducted in order to develop the tool on a realistic example set of learning cases and playtest action streams.

We have two main hypotheses for our method:

- H1: Using Learning Case Mining, we can tell if game mechanics or dynamics have issues with their initial learnability.
- H2: Using Learning Case Mining, we can determine the reason behind the issue of initial learnability of game mechanics or dynamics.

Our planned experiment requires evaluation of more than one game with a known learnability problem. For each of these games, two learning cases are formulated – one with a known learnability issue and one without.

A set of at minimum ten playtesters is chosen. Before the test starts, each playtester has to provide their distinguishing identifier, level of experience with video games (casual to hardcore) and the confirmation that they never played neither of the games that are used for the experiment. Each playtester then takes part in a playtest of each of the tested games.

The playtests are not strictly limited by time. The players are recorded using gaze tracking software, their gaze is tracked and Learning Case Mining is applied to the data collected from their behavior during the playtest. After the main experiment, the players are interviewed and made to watch their recorded playtest on video replay and provide insight into their behavior during the experiment.

After the experiment, the overall results of the playtests performed are qualitatively compared with expected results and the explicit feedback from players collected from video replay.

5 Conclusions

We have proposed a method, which can improve the state of evaluation of initial learnability of games and can also be integrated in the playtesting methodology. This method is based on comparing the expected behavior of players with the real stream of events collected during the playtest. It uses gaze tracking as an additional source of implicit feedback, and is also undisruptive for the player, and is automatized for even larger scale playtests.

Moreover, this method is general enough that we can claim its usability in every area of software, where checking if users are behaving in an expected way could be beneficial.

Among the possibilities for future work is to redesign the learning case matching algorithm to one based on graphs and using a different set of metrics. Another plan is to make Geddit available through a web interface, so our method is more readily available for use by any interested game designers.

Acknowledgement: This work was partially supported by the Cultural and Educational Grant Agency of the Slovak Republic, grant No. 009STU-4/2014.

References

[1] Schell, J.: *The Art of Game Design: A Book of Lenses*. Morgan Kaufmann Publishers, (2008), pp. 115–126, 399–401.

[2] Faro, A., Giordano, D., Spampinato, C., De Tommaso, D., Ullo, S.: An Interactive Interface for Remote Administration of Clinical Tests Based on Eye Tracking. In: *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*, (2009), pp. 69–72.

[3] Wang, H., Chignell, M., Ishizuka, M.: Empathic tutoring software agents using real-time eye tracking. In: *Proc. 2006 Symp. Eye Track. Res. Appl. – ETRA '06*, (2006), pp. 73.

[4] Eivazi, S., Bednarik, R.: Inferring Problem Solving Strategies Using Eye-Tracking: System Description and Evaluation. In: *Proc. 10th Koli Call. Int. Conf. Comput. Educ. Res.*, (2010), pp. 55–61.

[5] Bartels, M.: Eye Tracking Insights into Cognitive Modeling. In: *Proc. 2006 Symp. Eye Track. Res. Appl.*, vol. 1, no. March, (2006), pp. 27–29.

[6] Kickmeier-Rust, M. D., Hillemann, E., Albert, D.: Tracking the UFO's Paths: Using Eye-Tracking for the Evaluation of Serious Games. In: *Virtual and Mixed Reality – New Trends*, (2011), pp. 315–324.

[7] Grossman, T., Fitzmaurice, G., Attar, R.: A Survey of Software Learnability: Metrics, Methodologies and Guidelines. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (2009), pp. 649–658.

[8] Hunicke, R., Leblanc, M., Zubek, R.: MDA: A Formal Approach to Game Design and Game Research. In: *Proceedings of the Challenges in Games AI Workshop, Nineteenth National Conference of Artificial Intelligence*, (2004).

[9] Damerau, F. J., I. B. M. Corporation, Heights, Y.: A Technique for Computer Detection and Correction of Spelling Errors. In: *Commun. ACM*, vol. 7, no. 3, (1964), pp. 171–176.

Crowdsourcing for Large Scale Texts Annotation

Jozef HARINEK*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
j.harinek@gmail.com

Abstract. There is huge amount of information stored in natural language. In order to be able to process it by computer, we need to pre-process the text into machine understandable form. To achieve this, we can use the process of annotating the texts in various layers. It starts from morphological layer and goes up to contextual layer. In our work we focus on syntactic annotation of large scale texts by employing crowdsourcing principles. We harness a numerous crowd – elementary and high school students who perform syntactic annotations anyway as school assignments. We give them tool to perform these annotations electronically and help to build annotated corpus. First experiment showed promising results that suggest the method is set up well. Proposed method is verified in a software prototype designed especially for this purpose. The aim of our work is to explore the possibilities crowdsourcing has in the field of creating syntactic annotations for Slovak language, identify the quality of annotations created by our method, identify the power of crowd in this field and explore the possible use cases that can benefit from such dataset.

1 Introduction

There is large amount of information stored in natural language on the web and in electronic documents. To be able to process it better we need to transform the text to a machine understandable form or add some additional information about the text that can help to understand its meaning. To this purpose serve several layers of annotation [3]:

- *Morphological layer* – the parser extracts grammatical categories of words.
- *Syntactic layer* – it is focused on identifying sentence elements and relations between them.
- *Morphemic layer* – adds morphemic information to previous layers.
- *Semantic layer* – it is focused on semantic structure of sentences.
- *Contextual layer* – adds to semantic layer information about word context.

In our work we focus on obtaining annotations for syntactic layer, exploring the power of crowd in this field and the possible applications of a dataset annotated by employing our method.

* Master degree study programme in field: Information Systems
Supervisor: Dr. Marián Šimko, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

The state of research in the field of natural language processing of Slovak language in syntactic annotations is showing little progress. There are not many projects aimed on syntactic analysis and the latest research was done in 2010 [2]. There is also a project of Slovak national corpus [5] which is, however, still in progress and not yet finished and released for public use. It was started in 2005 and manually annotated by experts in the domain. The creation process was not yet finished during the last year. Automated language processing always depends on the language itself or at least group of similar languages [1] and in order to process natural language better we can use annotated corpora. Due to the state of the art in the field of syntactic annotations of Slovak language we decided to employ crowdsourcing principles in this area.

Crowdsourcing is recently very popular in field of large datasets processing or completing tasks that require a significant human input in order to solve the task. It uses crowd to complete a task that is usually done by an expert [4]. Crowdsourcing emerged with the Web 2.0 phenomenon that set the trend of masses-created web content. Similar problems to syntactic analysis were solved by employing crowdsourcing techniques. In [7] authors used crowdsourcing methods to acquire annotations for pictures by modifying the popular Concentration game.

There is also a potential numerous crowd to be harnessed in creating annotations. Every student in Slovak republic has to, at least in a part of his studies, come across syntactic annotations training and therefore also has to complete some home and school assignments in this field. Given an appropriate tool we can encourage students in the learning process and also collect annotations generated by them.

These facts encouraged us to employ the techniques of crowdsourcing in the field of syntactic annotation of Slovak language. Promising results in other fields, similar to this suggest that crowdsourcing will be also efficient in this domain. In our work we explore the power of the crowd in this domain, quality and possible further use of annotations created in such way.

The paper is structured as follows. In the next section we describe our proposed method for obtaining syntactic annotations via crowdsourcing, evaluation of experiment is following. In the last section we conclude the results and propose an extensive experiment to prove the concept more elaborately and to improve the method.

2 Crowdsourcing method for syntactic annotations acquisition

We propose a method for syntactic annotations acquisition based on crowdsourcing principles. In our method we harness crowd of elementary and high school students that need to perform syntactic annotations as school assignments anyway. Students from 5^{th} grade of elementary school start to learn basics of syntactic annotations which means that even in the young age they are able to perform a very basic syntactic annotation. As they learn more every year, they are able to perform more complex annotations and in the 8^{th} grade of elementary school the students spend quite significant amount of time studying syntactic annotations.

Our method is designed to employ this crowd of students. We designed a software tool that enables teachers to define an assignment and gives the students environment to draw a syntactic tree of sentences (see Figure 1). The main part of this system is an interactive tool to draw the syntactic tree from given sentence. In this syntactic tree student has to identify particular sentence elements and also identify relationships between these elements. In the next paragraphs we describe the method by means of crowdsourcing dimensions as they are described in [6].

There are several factors that distinguish particular crowdsourcing methods from each other. As a *motivation* factor we use points or grades that students get after completing their assignments in the best way they can. The *quality control* of user generated output is ensured by following the rule of majority. Later in this section you can see more detailed description of our method for correct annotations distinction. *Data aggregation* is covered in the process of annotations collection. We build a dataset of correctly annotated sentences that gradually grows. *Human skill* needed in order to participate in our method is knowledge of syntactic analysis in Slovak language.

We do not specify the level of proficiency, because we can use various detail levels of annotations. In our experiment we focused on secondary school students that are able to annotate whole sentences on a basic level. *Task cardinality* is many-to-many since one sentence is annotated by several students and each student gets many exercises with different sentences.

As mentioned above, motivation is ensured by the need of grade or points for assignment from students. The model is that students get an assignment which they have to fulfil as a part of their course. According to our discussions with teachers and students as well, they need to complete syntactic analysis assignments anyway and doing so by means of such system is for them much more interesting and motivating for them than writing the homework assignment by hand.

Ensuring the quality of output is a challenge for us. We plan to develop a system of combining the results from individual students that will yield a correct solution after certain submissions of the same sentence. Right now we are still setting up the right parameters based on data collected from students. Another way of ensuring the right output is via teachers that can always choose to create a correct solution. We plan to identify solutions that are ambiguous and therefore are likely to come from a problematic task. Such tasks are forwarded to teachers that will create correct solution for that particular sentence.

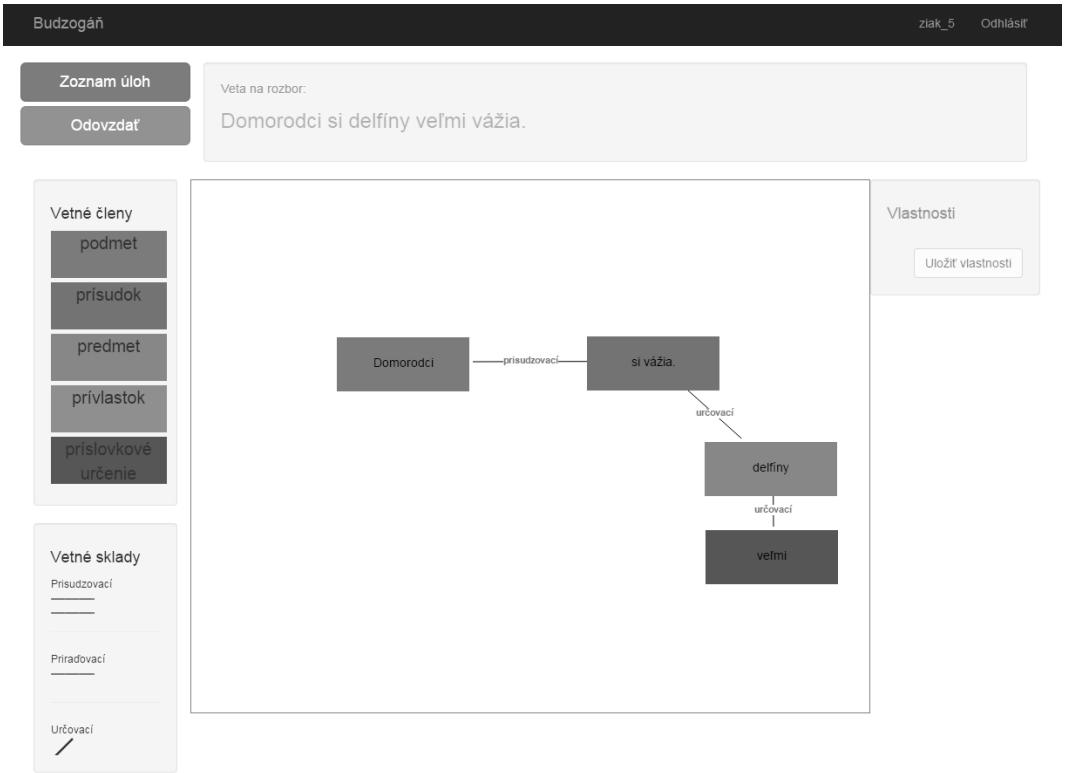


Figure 1. Screenshot of the environment.

In our research before proposing this method we also talked to teachers about such system. They were very keen about such tool and they said that they believe it will help the children to complete their assignments with greater willingness. We also evaluated the market for such tool. Its target group are last five years of primary school and four years at high school. All of these students are capable of creating syntactic annotations of some level of granularity. Of course, the younger students can create only basic annotations but the level of proficiency is increasing every school

year. In school term of 2013–2014 there were 209 133 students at primary schools and 76 711 students at high schools in the appropriate age. This means that the crowd is potentially quite large.

Role of a teacher is also a significant one in our system. On one hand, the teacher is present as one who defines the tasks for students. This means he is cooperating with the requester (as defined in *process order*). On the other hand, teacher can submit a solution himself and add to the crowd wisdom. Such solution can be then considered as a golden standard or a solution with higher weight.

2.1 Method for identifying correct solution

The ideal case is when we can obtain a complete and correct annotation of a sentence. However, also the partial solutions (not all elements or relations in a sentence assigned correctly) are useful and can be used in further text processing. Besides, we can build a complete solution from several partial solutions (one student gets correctly one part another student different part of the sentence).

We propose a method to identify correct solutions automatically from collected data. In this paper we present the first version based on data we collected during the first experiment. It takes into account the majority of votes on each element and if it is over a certain threshold (marked as X in 1) it is marked as a correct solution for that particular element or relation identified by the crowd. If the threshold is not met the element (or relation) is sent to revision to an expert (teacher) or not resolved and left without annotation.

$$element = opt_i; \max\{\frac{|opt_1|}{|evaluators|}, \frac{|opt_2|}{|evaluators|}, \ldots \frac{|opt_n|}{|evaluators|}\} \quad (1)$$

$$\max > X$$

Gradually we build a correct solution for whole sentence from these partial elements and relations annotations.

Our future work is to improve this identifying method to yield even better results. We also plan to add a method for error and unclear task detection based on further data analysis and finding patterns in data. All of this will help the teachers and they will not need to check the student's solutions.

3 Evaluation

To verify the concept we designed a proof-of-concept experiment in which we had a group of 20 students that participated in one 45 minutes session of annotations creation. During this experiment we obtained 186 annotations of sentences. 8 sentences had more than 12 different annotations. For evaluation purposes we also obtained one set of annotations from teacher that served as a golden standard.

We evaluated the results from several points of view. Sentence elements, sentence elements relationships, complete sentence analysis, sentence elements in sentences. The results of the experiment are shown in *Table 1*.

Based on the results we can see that even with such a small group of students with relatively small number of annotations per sentence we can obtain valid syntactic annotations. We can see that this proof-of-concept experiment yields promising results. That encourages us to further work on this project in order to provide a better method validation and to set up the concepts even better.

As a part of the first experiment we also employed a qualitative experiment to verify the concepts as set up in crowdsourcing dimensions and improve the environment for annotations collecting. We gave the students a questionnaire in which we focused on their user experience after experiment and we also posed questions about the concepts. 100 % of the students participating in the experiment said that they like the concept of completing their assignments via such system. Our set up of motivation proved to be correct.

Table 1. Results of first experiment.

	Correct solutions
Sentence elements	92.68 %
Sentence elements relations	90.00 %
Sentences – complete	71.43 %
Sentences – elements	85.71 %

We also collected some useful comments on the environment based on which we could improve it. It was also discussed with teachers to get also their ideas about what functions should such system have. Based on this information we are improving the environment and preparing next version for further experiments.

4 Conclusions

In our work we proposed a method for syntactic annotations acquisition. This method is based on crowdsourcing principles. To prove the concept we created a software prototype to collect syntactic annotations which is also a platform for educational system.

In the first experiment we showed that the idea of collecting syntactic annotations by primary and high school students is reasonable. We were able to obtain correct annotations even with a relatively small number of participants (20).

Our future work is to perform an extensive experiment in which we plan to collect more data and make further decisions on the project. We plan to obtain parameters for automated solution retrieval. We will also focus on identifying the possible usages of a corpus created by our method. To compare the quality of our corpus we can use the Slovak national corpus [5] which we have in possession from JULŠ[1].

The future experiments involve two phases. The first one is collecting data from students. We plan to deploy our solution to several primary schools for 8th and 9th grade students. The first group concludes the studies of syntactic analysis and learn all the basic sentence elements. The second group, 9th graders, revise concepts from whole primary school as a preparation for high school. These two groups are quite large in number and therefore the most likely to participate in our experiment.

The second experiment will consist of data analysis collected in the first phase. We want to identify patterns that suggest faulty solution or difficult/unclear task. To be able to do this we need to set up parameters in correct solution identifying and possibly create a more sophisticated version of the method.

Acknowledgement: This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG 1/0752/14.

References

[1] Cimiano, P.: *Ontology Learning and Population from Text: Algorithms, Evaluation and Applications*. Springer, (2006).

[2] Čižmár, A., Juhár, J., Ondáš, S.: Extracting Sentence Elements for the Natural Language Understanding Based on Slovak National Corpus. In: *Proceedings of the International Conference on Analysis of Verbal and Nonverbal Communication and Enactment, COST'10*, LNCS 6800. Springer, (2010), pp. 171–177.

[1] http://www.juls.savba.sk/

[3] Furdík, K.: *Information Retrieval from Texts in Natural Language, Using Hypertext Structures* [in Slovak]. PhD Thesis, Technical University of Košice, (2003).

[4] Howe, J.: The Rise of Crowdsourcing. *Wired magazine*, (2006), no. 14, pp. 1–4.

[5] Gajdošová, K.: Syntaktická anotácia vybraných textov Slovenského národného korpusu [in Slovak]. In: *Varia. XVI. Zborník príspevkov zo XVI. kolokvia mladých jazykovedcov. (Častá – Papiernička 8.– 10. 11. 2006)*, Bratislava: Slovenská jazykovedná spoločnosť pri SAV – Jazykovedný ústav Ľ. Štúra SAV, (2009), pp. 140–148.

[6] Quinn, A. J., Bederson, B. B.: Human Computation: A Survey and Taxonomy of a Growing Field. In: *Proceedings of the SIGCHI conference on human factors in computing systems*, ACM, (2010), pp. 1403–1412.

[7] Šimko, J., Bieliková, M.: *Semantic Acquisition Games: Harnessing Manpower for Creating Semantics*. Springer, (2014).

User Model to Identify his States

Michal PETRÁŠ*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
xpetrasm@stuba.sk

Abstract. User model and identification his states is a important part of building Intelligent Computer Systems. These computer systems may based on the identification and modeling of the user work more efficiently, users can get better advice, or help to better results. This work focuses on identifying the user experience while using computer in different environments. As work with computers we understand using the keyboard, the mouse and work with the environment itself, in which users perform tasks. With identifying experiences is also possible to map changes in emotional state of user. This paper tries to reveal the influence of emotions on the experiences of users in their work.

1 Introduction

User's model and his identification in various forms are used these days in various kinds of areas. Whether it is web browser and application collecting data about user, or it is desktop application that is capable of recognition and precise identification of user and his identity based on user's model, the recognition is using user's habits of using keyboard or mouse. All of these applications have the common goal – to gain user's characteristics during use of keyboard and mouse. Based on these characteristics, they are able to create model of user that serves for his identification and identification of his various states, e.g. experience or emotions.

The presented topic is subject of discussions in the field of intelligent systems. Experts are having various different opinions about characteristics for each individual and there is conflict when discussing use of keyboard and mouse. The number of characteristics when using keyboard is relatively large. The most important characteristics in this field have been experimentally defined, such as dwell time, flight time and others. Similarly, such approach has been used also for use of mouse, where the most characteristic are parameters such as movement speed and change of direction.

Based on these characteristics, we are able to create model of user that we are then using for his identification and identification of his states. Various experiments proved that sometimes using of proposed characteristics is not enough. In order to improve the overall results, it is necessary to combine these characteristics for each individual user.

Some experts are claiming that using information about use of keyboard and mouse is not effectively describing model of specific user and they are aiming also on actions that is user

* Master degree study programme in field: Information Systems
Supervisor: Assoc. Professor Daniela Chudá, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 39–44.

performing during work in environment. Those actions are then assigned to specific user. In our work, we studied existing approaches and we decided to experiment with combination of all three approaches in a way that would enable us to identify experience of user with work in two different environments. In other words, we used data from work using keyboard, mouse and also from actions performed in environment.

2 Related Work

2.1 GIMP

In work [3] they studied the dynamic detection of novice and more skilled user that were using graphical software GIMP. During the study, they created statistical model that enabled classification of user with 91% accuracy.

Their focus was on low level events such as: mouse move, time needed for action in main menu. Based on collected samples from participants, they created classifier that could be used for classification of individual users.

They worked with this data:

1. Total time
2. X and Y speed of mouse
3. X and Y acceleration of mouse
4. Delay time
5. Number of open submenu
6. Depth of open menu
7. Number of visited items in menu
8. Unique visited items

2.2 User experiences based on interaction of user with application

Work [2] examines the detection of user's experience based on interaction of user with application and its interface.

In mentioned article, definition of skill is following: To know how to do something. Their research was based on such assumption. The work contains two performed experiments that were used for construction of classifier. Experiment was conducted using graphical software implemented in Java. Before experiment, some small changes of source code were performed in order to improve the outcomes.

The experiment had a strictly steps. Every test subject work with these steps: Open Image, Copy image from left bottom corner to the right top corner, pick blue color from the image in the left corner of image, pick font and size, Make text – hello world into the center of image, pick font and size, apply sharpen filter, save and exit.

2.3 Experts identification

Several studies [1, 4, 5] are aimed on research and experiments based on researching of only one specific area, for instance, dynamics of writing using keyboard. However, in mentioned works is not direct focus on experimenting with user's experience.

3 Work with user experiences

Based on analysis of similar works, we proposed method for testing of users where we will investigate user's experience during working on PC. For purposes of experiment, we needed

logged data of various users in various environments. Data recording was performed by logger prototype of our own design that is capable of recording of work with keyboard, mouse and also other actions in environments. Proposed logger is designed in Java programming language and uses libraries that enable it to acquire low-level events in operational system. In such a way we were able to catch all the necessary data about work of user in raw format. Software prototype then saved information into text file in format: action – time – definition.

From data obtained during the test, we exported necessary information into vector of users. Another approach for studying of characteristics was proposed, where characteristics recorded during work with keyboard and mouse were combined and then user's actions performed in a given environment were added into model.

We created vector of user and we used characteristics in Table 1:

Table 1. Our characteristics.

Number of backspace	Number of delete
Number of ctrl,alt, shift	Flight time
Average time – calm cursor	Number of mouse moves
Total time – work with keyboard	Total time – work with mouse
Number of actions	Digrafs
Trigrafs	Total time of work
Velocity of mouse	

We suggest these characteristics by the actions:

Table 2. Actions.

Text editor	Eclipse
Number of total actions	Number of total actions
Number of align actions	Number of undo actions
Number of change color actions	Number of export actions
Number of change font actions	Number of save actions
Number of change heading	Length of work in file
Average time between actions	Number of run program
	Length of work in Windows

Each participant completed the questionnaire related to logged information, prior to test and after it. This questionnaire is supposed for self-evaluation. Users were told to evaluate their experience with work using keyboard, mouse and also overall computer skills.

4 Implementation

4.1 Logger

We implemented logger in Java. We used combinations of open source library and our own code. This is a library to provide global keyboard and mouse listeners for Java. This will allow you to listen for global shortcuts or mouse motion that would otherwise be impossible using pure Java. To accomplish this task, library leverages platform-dependent native code through Java's native interface to create low-level, system-wide hooks and deliver those events to your application.

Following the previous actions, we created software module capable of saving logger data into text file. Structure of data is variable and is defined in source code, so in case of need, we are able to change the structure by simple change of few lines of the code. The main advantage of our logger is its usability possibilities. We used this logger in our text editor as embedded part and we used logger as a separately part with Eclipse.

4.2 Extractor

After raw data collecting, it was necessary to extract the relevant data and we needed to create vector of user. Subsequently, we implemented next part of our logger for attribute extraction. Extractor is capable of attribute extraction from section 3. Proposed software creates large attribute vector for each user.

4.3 Solution scheme

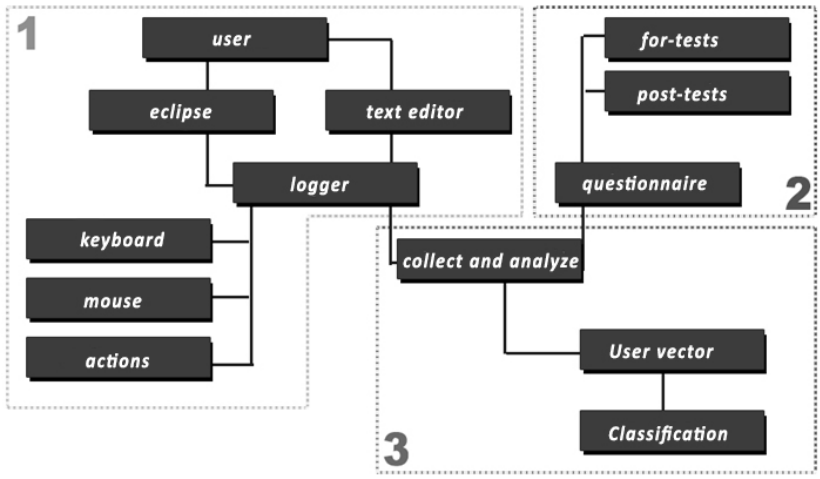


Figure 1. Scheme of our solution.

5 Data Set

Testing has been conducted in two different environment – Eclipse and text editor implemented in Java. For text editor, we had 27 testing participants and their task was to write defined, non-changeable text from field of medicine. Duration of experiment was approximately 40 minutes. For Eclipse, 17 participants were tested with task of creation of project according to instructions. Each participant completed the questionnaire for self-assessment.

6 Experiments

Collected data were used in various experiments. Experiment was divided into two groups with respect to application. User's experience in Eclipse may differ from those in text editor. From recorded data, 16 different experiments were created and various methods were applied, such as SVM and decision trees. For each sub-experiment, the set of various attributes from user's vector were chosen, using the following choices: all attributes of vector, omission of actions, omission of digraph and three graphs, selected attributes of vector similar to related work. In Table 3 are collected results.

Our next experiment we confirmed results of study [1]. Authors in this work collected data from user, which is very interest for our cause. Authors get results: speed of typing rapidly increase if user has positive emotional state, speed of typing rapidly decrease if user has negative emotional state. Keys like backspace, unrelated keys can occur more often. Test providing set of user's sentences has been created and users were told to write proposed sentences and assess each of the sentences by emotion. Using proposed approach, we gained data for work using mouse and keyboard and data are relevant for mapping of emotional state of user.

Table 3. Results of experiments.

Text editor	Eclipse	Accuracy	Precision	Recall	FScore
	SVM				
	All attributes	78.57 %	1.0000	0.6666	0.8000
	Without digraphs and three graphs	64.28 %	0.9444	0.8947	0.9188
	Without actions	64.28 %	0.9444	0.8947	0.9188
	Selected attributes	50.00 %	0.6000	0.6667	0.6310
	Decision trees				
	All attributes	71.43 %	0.7777	0.7777	0.7777
	Without digraphs and three graphs	64.28 %	0.9444	0.8947	0.9188
	Without actions	64.28 %	0.9444	0.8947	0.9188
	Selected attributes	57.28 %	0.7140	0.5555	0.6249
SVM					
All attributes		77.27 %	0.8500	0.8900	0.8700
Without digraphs and three graphs		77.27 %	0.8500	0.8900	0.8700
Without actions		77.27 %	0.8500	0.8900	0.8700
Selected atributes		31.82 %	0.8300	0.2600	0.4000
Decision trees					
All attributes		90,90 %	0.9473	0.9473	0.9400
Without digraphs and three graphs		72.00 %	0.8421	0.8421	0.8421
Without actions		86.00 %	0.8947	0.9444	0.9188
Selected attributes		63.00 %	0.7500	0.8333	0.7890

Problem of emotion states is very large, so we suggest new method of mapping emotion states of user of work with keyboard and mouse. Our method is based on change of speed typing, number of individual keys, unrelated keys or mouse speed.

It was not possible to verify results with 100 % accuracy due to amount of obtained data from tested subjects. Hence as a first we verified it with questionnaires from participants. In other words, for purposes of evaluation we relied on data provided by users before and after tests. These data were considered as reliable and as guiding reference. Of course, we prepared verification for cases when user would self-evaluate himself too bad and thus he would fall into incorrect group. Such inaccurate evaluations did not gain high weight in experiments with the characteristics. In case of gaining multiple testing subjects, we would use various kinds of methods for evaluation of our method, e.g. cross-validation.

7 Conclusion

In our research, we implemented method in form of software prototype that is capable of recording user's activities during work on computer. Gained information is then transformed into vector of user which serves for creating model of user. Tested subjects completed questionnaires of self-evaluation during testing and those were divided according to used environment. All of gained information was used for classification of user according to experience. Promising results were obtained for both environments, for text editor as well as for Eclipse. In case of text editor, 90 %

accuracy in users' experience classification was achieved, whereas for Eclipse 78.57 % was achieved.

8 Future work

Future work will include several steps, such as involving more participants, experiments with emotional states of users and influence of emotion with computer usage.

Acknowledgement: This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG 1/0646/15.

References

[1] Nahin, A. F.M.N.H. et al. Identifying emotion by keystroke dynamics and text pattern analysis. In Behaviour & Information Technology [online]. 2014. Vol. 33, no. 9, pp. 987–996. [accessed 2015-02-15].
Online: http:// www.tandfonline.com/doi/abs/10.1080/0144929X.2014.90734

[2] Ghazarian, A., Noorhosseini, S.M. Automatic detection of users' skill levels using high-frequency user interface events. In User Modeling and User-Adapted Interaction [online]. 2010. Vol. 20, no. 2, pp. 109–146. [accessed 2015-02-15]. Online:
http://link.springer.com/10.1007/s11257-010-9073-5.

[3] Hurst, A. Dynamic Detection of Novice vs. Skilled Use Without a Task Model. In [online]. 2007. [accessed 2015-02-17].
Online: http://www.cs.cmu.edu/~akhurst/publications/chi574-hurst.pdf.

[4] Epp, C. Identifying emotional states through keystroke dynamics. In [online]. 2010. [accessed 2015-02-15]. Online: https://graphics.im.ntu.edu.tw/~itspeter/194-ceppthesis.pdf

[5] Killourhy, K. A Scientific Understanding of Keystroke Dynamics. In [online]. 2012. no. January.[cit.2015-02-15]. Online: http://oai.dtic.mil/oai/oai?verb= getRecord &metadataPrefix= html&identifier=ADA557327

Visual Analytics Tool for Energy Consumption Data Set

Andrej ŠTAJER*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
andrejstajer.sk@gmail.com

Abstract. In our work we focus on data analysis. One of the way to make these analysis is Visual analytics. Visual analytics uses knowledge of visualization techniques and automated data analysis. We analyzed state of the art and discovered challenges, which could be examined. We made analysis of tools supporting Visual Analytics. Based on our energy domain dataset we examined different visualization techniques with well explaining ability.

1 Introduction

Visual analytics is not just about view and visualization of data, as the name may suggest. It is a process where knowledge is a goal. In order to success in this task we may use visualizations or data mining [1]. By interconnection of these two different approaches we can speed up understanding of dataset. Effectiveness depends on degree of interaction. Automated analysis needs high degree of interaction to be effective, yet visualization with lot of interaction is sometimes not effective. However, if we merged these two methods we would get maximum effectiveness in certain level of interaction [2].

Visual analytics is a young domain which is based on multiple areas. It brings new winds into waters of sheets of analytical ships. It uses knowledge from the statistical analysis, the cognitive science, the data management and many more. It tries to simplify complex data to discover new relationships which data suggests. However Visual analytics is not just multiple domains interconnected [3].

There are many tools claiming they support Visual analytics. However they mostly use simple line or bar charts visualizations. In addition they are designated for experienced analyst who knows which visualization he should use. Customization of view is also an issue. The user is not allowed to adjust parameters to gain the better overview. Another functionality needed for an advanced visual analytics is predictive analysis. If tools provide this kind of analysis, it does not allow to choose prediction model. Thus it is proprietary [4].

Field of Visual analytics has multiple pioneers and its popularity is growing in recent years. Many works are based on the pillars defined in [5]. Authors separated simple visualization of data and applied statistical knowledge to make his data more understandable.

---

* Master degree study programme in field: Information Systems
Supervisor: Assoc. Prof. Viera Rozinajová, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 45–49.

Applying knowledge of multiple fields allows to look at data from different perspectives and angles. In [6] authors define challenges of this field as:

- Scalability
- Uncertainty
- Interaction
- Evaluation

Authors of [8] take evaluation as more complex than others. Key aspects for evaluation are user satisfaction, effectiveness and efficiency. Visual analytics has three levels for evaluation approach component, system and environment. There are isolated evaluations for components. For system level we need to track history of data analyzing. For environment level technology adoption is considered.

2 Goals of work

In our work we set these main goals:

- Complex data visualization with orientation on pattern discovery
- Interaction with big datasets
- Time series forecast
- Deeper and faster information retrieval

3 Dataset description

The domain of our project is the energy consumption. The chosen dataset contains 21 000 households and consumption places from around 1 300 localities in Slovakia, where smart-meters were installed. These devices are able to measure consumption online in 15 minute intervals. This makes 96 measurements per day and 2 880 per month. Therefore, it is possible to predict consumption, even for very short periods. The dataset is made from measurements from January to December 2014. We extended this dataset with historical data of weather. This weather data consists of main weather characteristics like temperature, humidity or rainfall.

We added another dimension by creating a dataset of social and cultural events. This contains roughly 2 000 events of year 2014 in Slovakia. We extended this set by Slovak holidays.

4 Proposed solution

Based on the analysis and Visual analytics mantra “Analyze first, show the important, zoom, filter and analyze further, details on demand” [1] we designed a process *(Figure 1*) to accomplish our goals. The main process is divided into three sub-processes. First of them is *an Integration*. The raw data input is transformed into the integration model with main entities. The other two sub-processes are alternating in a cycle after the completion of the first sub-process. *An Aggregation & DrillDown* is process of changing a granularity of the model on demanded level. The last but not least is *a Visualization* process. Here we use the output of *the Aggregation & DrillDown* and apply advanced visualization techniques and graphs.

We are not trying to replace traditional methods of data analysis. We are trying to help the user to understand faster and help him with first steps or where should he look at the beginning. By [7] there are four strategies of data analysis. Our goal is to give a clue to the user. This step is supposed to reduce time needed to understand connections between entities. Without this clue the user would be forced to apply bottom-up or top-down strategy which are time consuming.

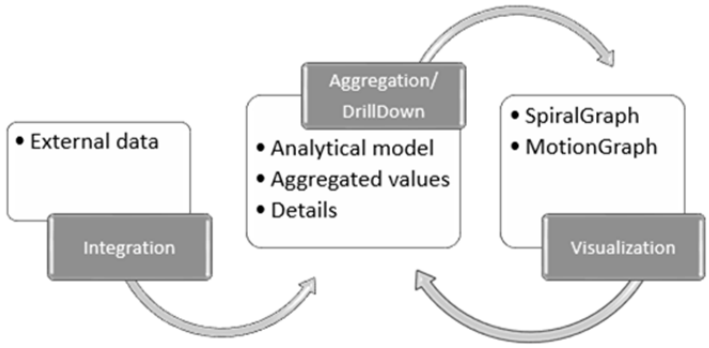


Figure 1. Designed Visual analytics process.

4.1 Prototype

We have designed and implemented a prototype system for Visual analytics. Prototype has client-server architecture and it is based on three components, shown in *Figure 2*. Every component has a role in the designed process. The integration component contains integration model specified for energy consumption domain. The role of the analysis component is automated data mining and granularity changing. The client part of application is visualizing analyzed model in interactive view. Advanced visualizations are used i.e. SpiralGraph [8].

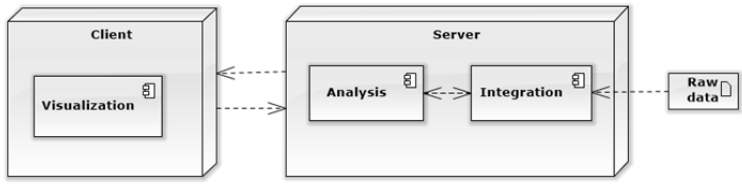


Figure 2. Architecture design of protype.

4.2 Modules

Integration module contains four main entities, which are indirectly connected via time and specific area:

- Consumption Place – location where smart-meter is installed with specified ZIP code
- Consumption – representation of actual energy consumption in exact timestamp. It is directly connected with Consumption Place
- Event – represents special events of different social and cultural influence, i.e. holidays, concerts, annual markets, etc.
- Weather – represents meteorological information about location in specific time

Analytical module is place of statistical calculations and knowledge discovery methods. It will maintain model of dataset with changing granularity. It is designed to execute two main tasks:

- Prediction – time series prediction based on Box-Jenkins methodology for creating ARIMA models. The user is able to choose elements for the prediction model
- Outlier detection – outliers are interesting from analyst point of view. They might indicate special event occurrence

However to improve the user experience it is important to maintain response time of interactions. This module will precompute values to react faster on the analyst's requests.

Visualization module is also graphical interface of our prototype. It allows to the user to work with dataset interactively change granularity or focus on some parts of dataset. It shows various representations of dataset including SpiralGraph (Figure 3). It shows time series in shape of the spiral. This allows the user to discover reoccurring patterns easier. Color saturation describe amount of energy consumption. Saturation levels are normalized to reveal differences of smaller values. The user is allowed to parametrize view by number of elements in one circle and offset of first value of spiral. For paper purpose colors of view were inverted and set to grayscale. Original visualization uses green color for spiral line and black as background. Benefit of this is higher contrast between values. Furthermore beginning of each month is set red for easier orientation and chosen date is highlighted by yellow border.

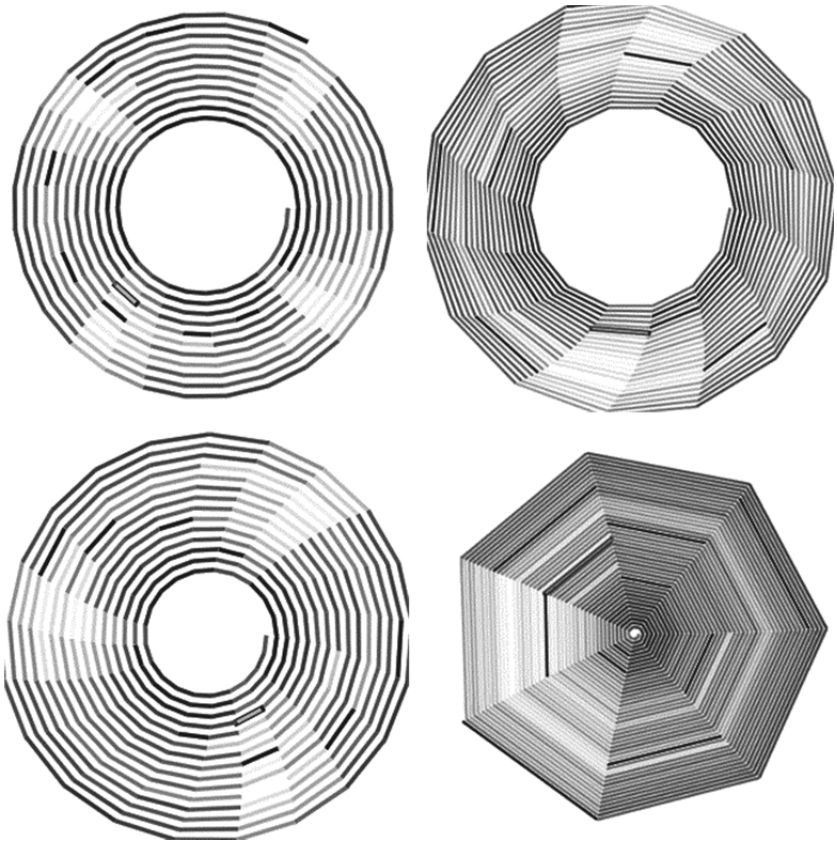


Figure 3. SpiralGraph implementation parametrized for display of 28, 14, 21, 7 days.

5 Experiment

A default view or initial analysis is given to the user as multiple visualizations. The user can interact with this view. He chooses what is interesting and what is worth of further analyzing. If the user needs he is able to change parameters of view.

The user was able to recognize energy consumption pattern in the weekend a weekdays. There were special cases correlating with special dates i.e. Easter. These days were shown in calendar view. Thanks to interconnection same day is highlighted at all visualizations.

6 Conclusion

We use visualizations to retrieve hidden relations and patterns. Advanced visualizations reveal more of data potential, but data are processed before they are shown. However, our work is very dependent on chosen dataset. Most of potential of Visual analytics is in multi-dimensional datasets.

Implementation of the prototype is sufficient to show its improvement of analysis. It supports multiple analyzing strategies and provides a good overview of big dataset.

Our goal for the future is to implement suggestions of points of interest. This will significantly reduce time of the analysis.

Acknowledgement: This contribution was created with the support of the Research and Development Operational Programme for the project International centre of excellence for research of intelligent and secure information-communication technologies and systems, ITMS 26240120039, co-funded by the ERDF. The paper was partially supported by the website www.cokdeje.sk, portal provided their information of events in year 2014.

References

[1] Keim, D. et al.: Mastering the Information Age Solving Problems with Visual Analytics. *In. Eurographics Association*. Goslar, Germany, 2010.

[2] Keim, D. et al.: Visual Analytics: How Much Visualization and How Much Analytics?. *In. ACM SIGKDD Explorations*, ACM, 2009.

[3] Thomas, J., Cook, K.: Illuminating the Path: Research and Development Agenda for Visual Analytics. *In. IEEE-Press*, IEEE, 2005.

[4] Zhang, L. et al.: Visual Analytics for the Big Data Era – A Comparative Review of State-of-the-Art Commercial Systems. *In. IEEE Symposium on Visual Analytics Science and Technology 2012(VAST 2012)*, IEEE, Seattle, USA, 2012.

[5] Thomas, J., Cook, K.: Illuminating the Path: Research and Development Agenda for Visual Analytics. *IEEE-Press*, IEEE, 2005.

[6] Keim, D.; Zhang L.: Solving problems with visual analytics: challenges and applications. *In Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies (i-KNOW '11)*, ACM, New York, USA, 2011.

[7] Youn-ah Kang; Gorg, C.; Stasko, J.: Evaluating visual analytics systems for investigative analysis: Deriving design principles from a case study. *In. VAST 2009*. IEEE Symposium, pp.139–146, 2009.

[8] Rind, A.; Lammarsch, T.; Aigner, W.; Alsallakh, B. & Miksch, S.: TimeBench: A Data Model and Software Library for Visual Analytics of Time-Oriented Data, *In. IEEE Transactions on Visualization and Computer Graphics, Special Issue "VIS 2013*, IEEE, 2013.

The Impact of Citation and Co-Citation Analysis on the Relevancy of Keywords in Digital Libraries

Máté VANGEL*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
mate.vangel@gmail.com

Extended Abstract

Keyword extraction methods are widely used for different purposes like automatic text summarization or user and domain modelling. In digital libraries there is much information related to each article, which may help to improve the results of the extraction methods. This information can be in the form of metadata, such as the name of the author, the year of publication, the category of the article, tags, the abstract of the article, bibliography and citations. In our work we analyse references and citations of research articles, and we use them during the keyword extraction process.

Since citations highlight different aspects of the articles from other researchers' point of view, it is possible to extract the most important information using only citations [1]. The authors in [2] extracted keywords from citation sentences; they evaluated that using citations has a positive effect on the quality of the extracted keywords. The optimal size of citations' contexts, i.e., the size of the text surrounding the reference was evaluated in [3] finding out that it should be at least 100 words before and 100 words after the citation string (reference). The authors considered words nearer to the citation string more important than words far from it. They also evaluated that if at least a citation context with only one sentence is analysed during keyword extraction, the precision of the method is higher compared to the variant with no citation contexts.

In our proposed method we prepare three sources of text for each research article. The first source of text is the full text of the given research article. The second source consists of all citation contexts citing the given article, while in the third source we have co-cited articles' text (two articles are co-cited when they are cited together by at least one other article). We parse citation contexts as shown in [1]. After the sources of texts are prepared, we use them separately to extract keywords using statistical method TF-IDF from each source individually. We normalise weights in each source and after that we compute the final weight of the words as a single value using a simple linear combination. For the graphical overview of the proposed method see Figure 1.

We evaluate the proposed method in a domain of digital libraries in a web-based bookmarking system *Annota*[1] using the explicit feedback from users. Our hypothesis is that using

* Master degree study programme in field: Information Systems
Supervisor: Róbert Móro, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

1 http://annota.fiit.stuba.sk

co-citations produces higher precision of the keyword extraction method compared to the method using only citations, especially when the article has only a small quantity of citations. In order to evaluate the hypothesis the keywords are extracted in two different ways. The baseline extraction method does not use co-citations (words are extracted only from the first and second source of text), while the second does. The participants will assess the extracted words from the provided list as relevant, less relevant or irrelevant. For us to accept the hypothesis, we should observe higher precision when all the sources (including co-citations) are used.

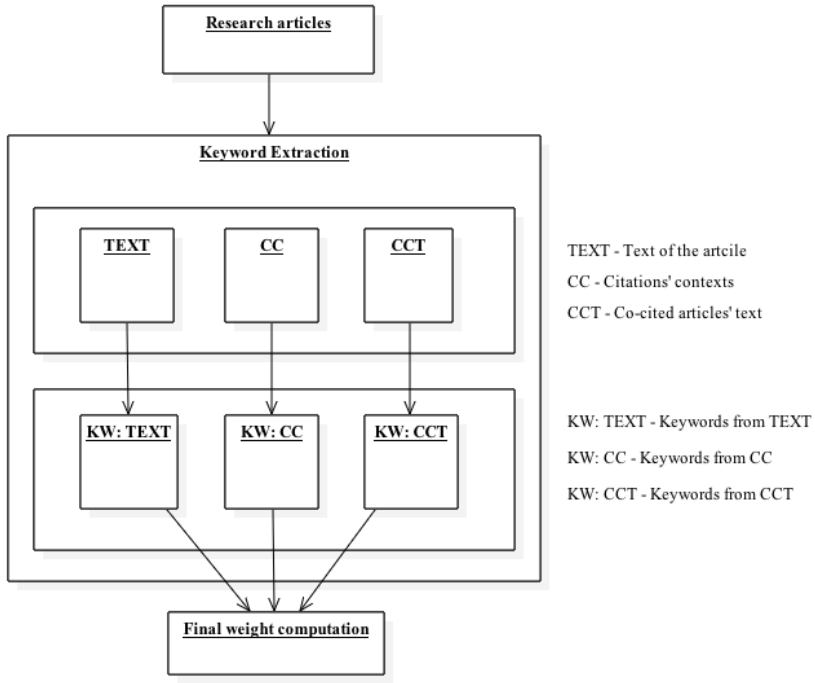


Figure 1. Graphical overview of the proposed method.

The main contribution of our proposed method is in using co-citations in the process of extracting the keywords and determining their relevancy; to the best of our knowledge, they have not been used for this purpose yet.

Acknowledgement: This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG 1/0646/15.

References

[1] Elkiss, A., Shen, S., Fader, A., States, D., Radev, D.: Blind Men and Elephants: What Do Citation Summaries Tell Us About a Research Article?. *Journal of the American Society for Information Science and Technology*, 2008, vol. 59, no. 2003, pp. 51–62.

[2] Qazvinian, V., Radev, D., Özgür, A.: Citation summarization through keyphrase extraction. In: *COLING '10: Proc. of the 23rd International Conference on Computational Linguistics*, (2010), pp. 895–903.

[3] Ritchie, A., Robertson, S., Teufel, S.: Comparing citation contexts for information retrieval. In: *CIKM '08: Proc. of the 17th ACM Conference on Information and Knowledge Management*, ACM Press, (2008), pp. 213–222.

Modelling User Interests in Latent Feature Vector Space based on Document Categorisation

Márius ŠAJGALÍK*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
marius.sajgalik@stuba.sk

Abstract[1]

User modelling includes modelling various different characteristics like user goals, interests, or knowledge. However, evaluation of each of these characteristics can be difficult, since every user is unique and objective evaluation of each modelled feature often requires huge amount of training data. That requirement cannot be easily satisfied in public research environment, where personal information is too confidential to be publicly accessible. In a common research environment, we are confronted with training the model on only a small sample of data. It mostly requires humans to evaluate the model manually, which is often very subjective and time-consuming.

We liken the problem of personalised keyword extraction to document categorisation. We consider users as being different category labels and web pages visited by user as documents. By extracting personalised (discriminative) keywords for each document and by mapping them into a multidimensional continuous latent feature vector space, we can easily infer user interests by aggregating document keywords contained within the web browsing history of the respective user.

We examine a novel approach to quantitatively evaluate user interests by formulating an objective function on quality of the model. Since every user tends to be unique in their interests, we focus on user interests as a discriminative factor between different users. We formulate our objective function to propagate more discriminative (possibly unique) interests, which are supposed to be more informative than the generic ones. Another view on this formulation is that given a particular domain of interest, we are interested only in such user interests that can be used effectively to provide the user with personalised content or services.

By developing an automated evaluation method, we can assess the quality of user models on much larger scale, since we are no more dependent on any manual assessment. The quantitative evaluation method may have a big impact on user modelling research. By automating evaluation, the development of new methods and models can progress at much higher speed. That means that researchers can make more iterations to improve their methods and user models, which may result in bigger and faster research improvements in the field of user modelling.

Acknowledgement: This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG 1/0752/14.

---

* Doctoral degree study programme in field: Software Engineering
Supervisor: Professor Mária Bieliková, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

[1] Full paper available in printed proceedings, pages 52-59.

Issues of Computer Mouse Data for Related Applications

Peter KRÁTKY*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova2, 842 16 Bratislava, Slovakia
peter.kratky@stuba.sk

Extended Abstract

Mouse data representing cursor movement or button clicks are often employed in applications such as user interface evaluation methods, implicit feedback gathering methods and recently even mouse-based biometric systems. Especially, the last mentioned application requires high quality data in order to get desired performance. This paper provides an overview of issues and recommendations for a researcher who wants to study and work with computer mouse data.

Mouse events provided by the physical device are mouse *movement*, mouse *button up/down* and mouse *wheel scrolling*. Most basically, each event is described by time in milliseconds and coordinates (X and Y) of the cursor on the screen. The first issue that should be considered by a researcher is log size. Each event is valuable piece of data and should not be missed. Especially, movement data are very frequent (approximately each 8 milliseconds). Thus, we analyzed how big storage space is necessary. Based on four experiments, we summarized average number of events per minute in an e-learning system, tourist portal, e-shop and a game with values 342, 496, 673 and 2770, respectively. Consequently, we assume that every minute a user interacts with a web page requires around 20 kilobytes of disc space.

Another issue of raw data is movement coordinates precision. Positions occur very often, but the coordinates are rounded to integers. One solution could be decreasing number of events per second. However, fine-grained information is lost in this case. So, we propose smothering of the recorded path and aligning positions to this new path. Another benefit from this approach is that uneven spaces between points could be covered with new points what solves some problems in processing stage. To smooth the path we calculate an approximation curve fitting the measured points. Various interpolation methods, such as natural cubic splines or Bezier curves, use measured points as control points which are crossed by the calculated curve. A bit different method is B spline which does not cross the control points, what is desired feature as those are assumed to be imprecise. Figure 1a shows illustration of an original path measured as well as curve fitted. In the Figure 1b, corresponding angle (tangent) changes are compared within original and smoothed curve.

Raw data need to be translated into something meaningful. There are a lot of features that could characterize user's behavior, such as movement velocity, curvature, etc. Especially mouse movement data hold very small amount of information rather than sequence of such points do.

---

* Doctoral degree study programme in field: Software Engineering
Supervisor: Assoc. Professor Daniela Chudá, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 60–61.

A window of N movement events could be used to calculate values of features. Most research works use rather forming events into meaningful actions in order to represent intent of a user. Usually, *strokes* or *curves* are actions composed of movement event sequence separated by a pause or click [5, 3]. In [1] there are four actions recommended to study – point and click, movement only, drag and drop, silence (no activity). A very complex hierarchy of actions is described in [2].

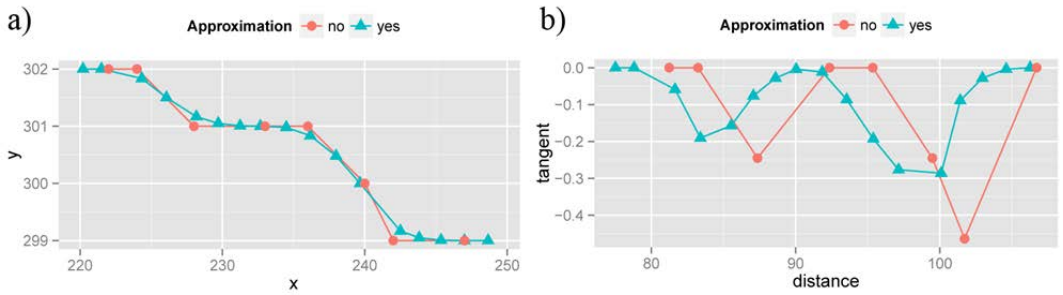


Figure 1. Example of smoothing curve of a) movement coordinates (x, y) and b) angles between points.

Aggregated data often include weird values that should be omitted. One should be careful when discarding values that not easily distinguished what is real and what not. The values of mouse features follow asymmetric distributions, thus proper outlier detection method should be used. For example, Tukey's method used in boxplots is suitable, which discards values out of the range 3 IQRs (inter quartile range) below the first quartile and 3 IQRs above the third quartile [4].

When conducting this kind of research which includes monitoring users, it is quite obvious that the users might feel uncomfortable. We conducted a short survey in which two questions about the privacy issues were asked: 1) *Imagine that you anonymously browse a web page (e.g. e-shop). It records all cursor movement for further analysis. Would you worry about your privacy?* 2) *Would you leave such a web page if you knew it records your cursor movement?* There were 78 participants of different nations, different age groups, both males and females in the study. Completed questionnaires show that 63 participants would worry about their privacy and 58 would leave the page if they knew their cursor is being monitored. We could definitely conclude that this is a sensitive topic for common web users who could feel kind of distress of something unknown that has a potential to be misused.

Acknowledgement: This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG 1/0646/15.

References

[1] Ahmed, A.A.E., Traore, I.: A New Biometric Technology Based on Mouse Dynamics. In: *IEEE Transactions on Dependable and Secure Computing*, (2007), pp. 165–179.

[2] Feher, C., Elovici, Y., Moskovitch, R., Rokach, L., Schclar, A.: User identity verification via mouse dynamics. In: *Information Sciences*. (2012), vol. 201, pp. 19–36.

[3] Schulz, D.A.:Mouse curve biometrics. In: *2006 Biometrics Symposium: Special Session on Research at The Biometric Consortium Conference, (2006)*, pp. 1–6.

[4] Seo, S.: A Review and Comparison of Methods for Detecting Outliers in Univariate Data Sets. M.S. University of Pittsburgh, (2006).

[5] Zheng, N., Paloski, A., Wang, H.: An efficient user verification system via mouse movements. In: Proceedings of the 18th ACM conference on Computer and communications security – CCS '11, (2011), pp. 139–150.

Computer Science and Artificial Intelligence

Motif Finding in DNA Sequences

Róbert CUPRIK*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
robertcuprik@hotmail.sk

Abstract. In this paper we present possible improvement of Multiobjective Artificial Bee Colony (MOABC) algorithm with Opposition Based Lévy Flight and changes to NSGA-II's Crowding-Distance Computation. MOABC is Swarm Intelligence algorithm used to solve Motif Discovery Problem (MDP), specially discovery of novel Transcription Factor Binding Sites in DNA sequences. We describe the changes made to MOABC and its implementation. We aim to improve the quality of found motifs.

1 Introduction

In recent years, algorithms based on Swarm Intelligence (SI) and bio-inspired computation in general have been applied to various real-world problems. One of such algorithms, developed at our Institute, has recently been applied to the problem of identifying and tracking multiple news stories [7]. These algorithms are inspired by metaphors of biological systems and, being a kind of Metaheuristic Algorithms, they are often used to solve challenging optimization problems.

Motif Discovery Problem (MDP) is a NP-Hard problem of finding motifs [4], which are short substring patterns that are repeatedly present at various positions in given sequence or set of sequences. Motifs in DNA sequences are sequences of nucleotides that hold some biologically significant information. However, these patterns are subjected to various mutations [2], so often a solution is not a sequence of nucleotides but it is represented as a Position Frequency Matrix (PFM). This matrix gives us information about motif and how often are nucleotides present at given position. MDP can be seen as optimization problem, where quality of found motif is being optimized. It is one of the problems connected to processing of DNA sequences. For these problems, too, algorithms inspired by bees behavior are increasingly applied, e.g. [1] at our Institute.

To address this problem Multiobjective Artificial Bee Colony (MOABC) [5] algorithm was proposed. It is SI algorithm inspired by foraging behaviour of bee colony. In MOABC MDP was defined as a problem with several objectives:

- Length of Motif,
- Support – in how many sequences is motif present,

* Bachelor study programme in field: Informatics
Supervisor: Professor Pavol Návrat, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 65–70.

– Similarity – how similar is motif among given sequences.

All of the found motifs are subject of given constrains:

– Motif Length Constrain – there is a minimum and maximum motif length,

– Minimum Support value,

– Minimum Complexity of Motif [4] – this concept is considered to be avoiding motifs such as "AAAAA".

In our research we are not trying to improve evaluation of found motifs but to improve a global solution finding process by using Opposition Based Lévy Flight [9] in MOABC and by changes to NSGA-II's [8] Crowding-Distance Computation [3] that is used when sorting colony. Opposition Based Lévy Flight ABC (OBLFABC) [9] was proposed as novel algorithm tested on 14 continuous optimization problems. We aim to improve quality of motifs found in MOABC using these changes.

2 Multiobjective Artificial Bee Colony (MOABC)

MOABC is characterized by two main parts of the algorithm. First is foraging behaviour of bees hive itself and second is evaluation of food sources (motifs).

2.1 Foraging behaviour of bees in MOABC

Artificial bee colony mimics foraging behaviour of a honeybee colony. This colony consists of three groups of bees. First, employed bees which evaluate located food source and subsequently fly back to hive and dance to reflect food source fitness. Second, onlooker bees that watch employed bees dancing decide on which food source to visit. Third, scout bees that explore new food sources randomly.

In MOABC, food sources are represented by position vector of bees in problem space. This position vector representing one solution gives us information on motif length and motif positions in given DNA sequences. First dimension of this vector is assumed to be motif length and other S dimensions are motif positions in S DNA sequences. Foraging behaviour of bees in MOABC can be described in following steps [5]:

1. Initialization of colony,
2. Employed bees phase,
3. Onlooker bees phase,
4. Scout bee phase,
5. colony is sorted using NSGA-IIs non dominated sort and crowding distance,
6. steps 2–5 are repeated until time limit is reached.

In MOABC, colony is represented as an array of bees. This array is sorted using NSGA-IIs non dominated sort and crowding distance [8]. Roles in colony are only set by bee's position in this array. Therefore, first N_W bees are considered to be employed bees, following N_O solutions are onlooker bees. Remaining – and the worst – N_S bees are considered scout bees and there is randomly generated food source for them in every iteration.

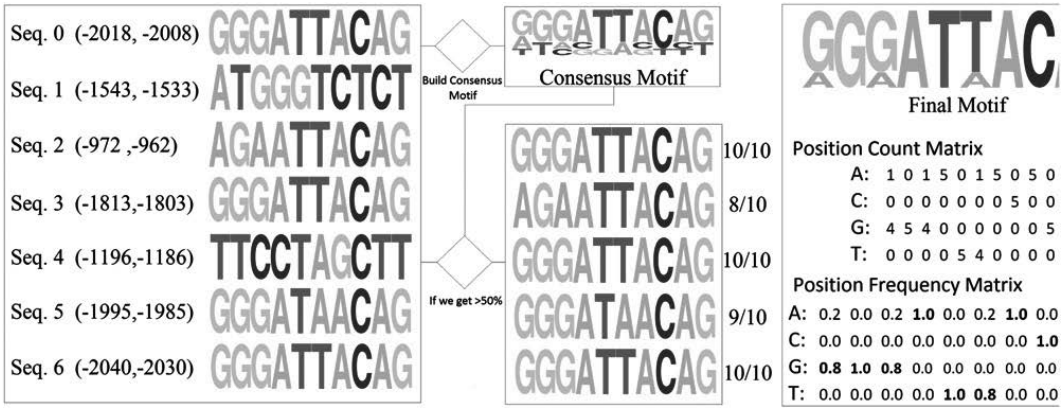


Figure 1. Building of PCM and PFM for given bee [4].

In our implementation, initialization of colony consists of colony array generation. Bees are generated in a way that every bee satisfies constrains given by MOABC. Generated array of bees is then sorted using NSGA-IIs domination concept and crowding distance with changes proposed by [3] explained in Section 3.

In employed bees phase first N_W bees try to improve their problem solution by moving in space using equation [9]:

$$v_{ij} = x_{ij} + \phi_{ij}(x_{ij} - x_{kj}) + \psi_{ij}(x_{bestj} - x_{ij}), \tag{1}$$

where $j \in \{1..D\}$, D is number of dimensions in position vector, v_i is a new possible position of ith bee ϕ_{ij} is random number between [-1, 1], x_k is randomly choosen bee and ψ_{ij} is uniform random number between [0, C]. First bee in colony array is considered x_{best}. This is modification to original ABC where bee movement is computed without x_{best} component [11]. New position of bee is then evaluated and if proven better, previous position is forgotten and replaced by new one.

In onlooker bees phase, bees in positions from N_W to $N_W + N_O$ randomly choose bees from employed bees group and using equation 1 try to find new solution. In our implementation, probability of bee x_i being chosen depends only on position in colony array. Bees with lower position are more likely to be chosen. If the generated position is better than previous one, onlooker bee forgets the old and moves to the new one.

In scout bees phase new positions are being generated and evaluated.

Next step consists of sorting new colony. Then, according to proposed changes, we apply Lévv Flight Search Strategy [9] and with some probability we use Opposition based optimization method [9] in some iterations.

2.2 Bee evaluation

In [4] found solutions are evaluated by using three objective functions. First objective is motif length and it is given by bee position vector. Second, we build a consensus motif. Consensus motif consists of the most frequent nucleotides at a given position in motif from evaluated bee. Then we can calculate support value. Only motifs that have 50 % or more nucleotides matching their counterparts of consensus motif are considered and counted to support value. In Figure 1 in middle column we can see consensus motif and motifs that satisfied support constraint. From these motifs, that satisfy this 50 % threshold we can compute position count matrix (PCM) and PFM. PCM is computed as number of each nucleotide at given position. PFM is based on PCM as a frequency of

these nucleotides on a given position. From PFM we can compute third objective – similarity, with following expression:

$$Similarity = \frac{\sum_{i=1}^{l} max_b f(b,i)}{l}, \tag{2}$$

where $f(b,i)$ is nucleotide b frequency on i-th position from PFM and l is motif length. Complexity of motif is then computed according to expression:

$$Complexity = \log_N \frac{l!}{\prod(n_i)!}, \tag{3}$$

where n_i is number of nucleotides of type $i \in \{A, C, G, T\}$ in final motif sequence.

3 NSGA-IIs Non Dominated Sort and Crowding Distance

Original MOABC [5] incorporates two mechanisms from NSGA-II [8]. Non dominated sort and crowding distance. In non dominated sort bees are assigned to pareto fronts based on number of other bees that dominate them. One bee dominates another if every objective function value of first bee is equal or better than value of the other and there is at least one value that is better. Therefore in first pareto front are bees that are not dominated, in second there are bees that are dominated only by one other bee, in third by two bees etc.

In these fronts we have to somehow sort the bees and this is what crowding distance is used for. In normal NSGA-II, crowding distance is computed as euclidian distance between every bee based on m objectives in m-dimensional space. For each objective bees are sorted using their fitness. Then for a given bee, the difference of fitness between previous and next bee in sorted group is added to ones crowding distance. Bees on borders (bees with the smallest or the highest fitness) are always selected as they have infinite distance assigned from definition. This crowding distance was, however, improved in [3]. Whilst in normal NSGA-II bees with same fitness value in m-dimensional space have different crowding distance values, in our implementation using [3], crowding distance is computed for unique fitness values only and then assigned to corresponding bees.

This modification was proposed in [3] to address the issue of instability when two bees share the same fitness. That is more likely to happen when fitness values can be discrete such as length of the motif or motif support in our domain.

4 Comparing Two Solutions

Many times during foraging process bees have to decide whether they should remember the new solution and forget the old one or to keep the current solution. In our implementation new solution is remembered if it dominates the old one. On the other hand, if none of the solutions dominates the other, all objective function values are normalised and summed together. Solution with higher sum of values is remembered.

5 LéVy Flight Search Strategy

After the original foraging process of MOABC we propose to use Lévy Flight Search Strategy. This strategy for ABC was described in [9] and consists of taking the best solution in every iteration and trying to improve it. This improvement consists of moving the bee using the step length drawn from lévy distribution in multiple iterations. The bee moves only when the position of next step is better than previous. This can be applied only on one bee or it may be applied on whole first pareto front. This process should enhance exploitation of MOABC.

Lévy distribution is used, as it was hypothesized, to be an optimal search strategy adopted by many foragers [11]. In our implementation we use Mantega's algorithm [6] to compute the drawn numbers. Parameters for distribution and number of iterations are to be set experimentally.

6 Opposition Based Optimization Method (OBOM)

In the end of the iteration process, there is a chance of applying OBOM [9]. First, we compute opposition bees. Opposition bee x_{op} to x is defined as:

$$x_{op_i} = a_i + b_i - x_i, i = 1, ..., D. \tag{4}$$

where $x_i \in [a_i, b_i]$ and D is number of dimensions in bee position vector. a_i and b_i are borders of a given dimension in position vector. They are computed for each iteration, where a_i is minimum and b_i is maximum value of ith component of position vector in swarm, thus preserving swarm convergence. Second, we combine original and opposite bees. Third, we sort the combined set and keep only the first bees to fit colony size. This process should enhance the exploration of MOABC.

7 Evaluation

Alghoritm will be evaluated using two different perspectives. First will be algorithm performance viewed from more generic perspective where obtained fitness values will be compared. Second evaluation will be based on results gained using Assessment of Computational Motif Discovery Tools[1] where several other motif finding tools such as MEME, MotifSampler etc. were evaluated. In this evaluation three different datasets are used (real sequences with known motifs [10], sequences randomly generated using Markov chain with inserted motifs and randomly chosen sequences with motifs inserted). This evaluation will provide us with statistical information on our found motifs such as number of motifs in both known sites and predicted sites by our algorithm (true positive) etc. Furthermore we will compare results obtained with and without proposed changes.

8 Conclusions

In this paper, we propose changes to MOABC [5] using Lévy Flight Search Strategy, Opposition based optimization method [9] and improved Crowding distance computation [3]. Our goal was to improve searching process of MOABC, namely quality of motifs found, from both generic fitness perspective as well as their biological significance. We are yet to obtain any test results so conclusion in this matter cannot be made. However, if proven helpful, these changes of MOABC can be applied on other multiobjective problems solved by Artificial Bee Colony.

Acknowledgement: This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG 1/0752/14.

References

[1] BouEzzeddine A., Kasala S., Návrat P.: Applying the Firefly Approach to the DNA Fragments Assembly Problem. *Annales Univ. Sci. Budapest., Sect. Comp. (2014)*, 2014, vol. 42, pp. 69–81.

[1] http://bio.cs.washington.edu/assessment/

[2] Das, M., Dai, H.K.: A survey of DNA motif finding algorithms. *BMC Bioinformatics*, 2007, vol. 8, no. Suppl 7, pp. S21+.

[3] Fortin, F.A., Parizeau, M.: Revisiting the NSGA-II Crowding-distance Computation. In: *Proceedings of the 15th Annual Conference on Genetic and Evolutionary Computation*. GECCO '13, New York, NY, USA, ACM, 2013, pp. 623–630.

[4] González-Álvarez, D.L., Vega-Rodríguez, M.A., Gómez-Pulido, J.A., Sánchez-Pérez, J.M.: Comparing multiobjective swarm intelligence metaheuristics for {DNA} motif discovery. *Engineering Applications of Artificial Intelligence*, 2013, vol. 26, no. 1, pp. 314–326.

[5] González-Álvarez, D., Vega-Rodríguez, M., Gómez-Pulido, J., Sánchez-Pérez, J.: Finding Motifs in DNA Sequences Applying a Multiobjective Artificial Bee Colony (MOABC) Algorithm. In Pizzuti, C., Ritchie, M., Giacobini, M., eds.: *Evolutionary Computation, Machine Learning and Data Mining in Bioinformatics*. Volume 6623 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2011, pp. 89–100.

[6] Mantegna, R.N.: Fast, accurate algorithm for numerical simulation of Lévy stable stochastic processes. *Phys. Rev. E*, 1994, vol. 49, no. 5, pp. 4677–4683.

[7] Sabo S., Kovarova A., Návrat P.: Multiple developing news stories identified and tracked by social insects and visualized using the new galactic streams and concurrent streams metaphors. *International Journal of Hybrid Intelligent Systems*, 2015, vol. 12, pp. 27–39.

[8] Seshadri, A.: NSGA-II: A multi-objective optimization algorithm. *MAT-Lab Central, Implementierung*, 2009.

[9] Sharma, H., Bansal, J., Arya, K.: Opposition based lévy flight artificial bee colony. *Memetic Computing*, 2013, vol. 5, no. 3, pp. 213–227.

[10] Wingender, E., Dietze, P., Karas, H., Knüppel, R.: TRANSFAC: A Database on Transcription Factors and Their DNA Binding Sites. *Nucleic Acids Research*, 1996, vol. 24, no. 1, pp. 238–241.

[11] Yang, X.S., Karamanoglu, M.: *1 – Swarm Intelligence and Bio-Inspired Computation: An Overview*. Elsevier, Oxford, 2013.

Usage of HyperNeat to Train Evasive Behavior in Ms. Pac-Man Video Game

Michal FARKAŠ*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
michalfarkas1@gmail.com

Abstract. Ms. Pac-man is famous video game published in 1982 by Midway, Namco's American distributor of original Pac-man. In comparison to its predecessor, Ms. Pac-man features non-deterministic ghost behavior and more mazes. Recently, this video game caught attention of wider artificial intelligence community. Between years 2007 and 2012 there were two large competitions held on various conferences. HyperNeat was introduced by Kenneth O. Stanley and others in 2009, it is neuroevolutionary method, which means it uses evolutionary algorithm to train neural networks. In this paper we have used HyperNeat to create artificial neural networks intended to evade ghosts in first maze, without any power-ups. Finally we devised several methods of selecting right action from output and compared them.

1 Introduction

HyperNeat is extension of prior neuroevolutionary method called Neat. Neat is capable of evolving topology as well as weights of connections. This means that human researcher does not have to create topology of network by hand. Neat solved problems of its predecessors by several mechanics, namely innovation numbers, speciation and complexification [8]. First mentioned is critical in Neat ability to mate neural networks, some other approaches relied solely on mutation to evade this problem. Speciation and complexification are vital to proper evolution, they secure diverse genofond and prioritization of simpler networks.

1.1 CPPN

Difference between HyperNeat and Neat lies in their encoding method. While Neat uses direct encoding, HyperNeat utilizes indirect encoding method based on CPPN. Aforementioned acronym stands for Compositional Pattern Producing networks, in their essence CPPN are separate neural networks capable of drawing pictures. These pictures can contain features such as symmetry, repetition, repetition with variation [6].

* Bachelor study programme in field: Informatics
Supervisor: Professor Jiří Pospíchal, Institute of Applied Informatics, Faculty of Informatics and Information Technologies STU in Bratislava

CPPN in HyperNeat work by taking coordinates of two points, which corresponds to neurons and returning value. Resulting value is processed and assigned as a weight of connection between two points fed into CPPN. In HyperNeat it is CPPN, which is evolved by Neat. This means that CPPN acts as a genotype, resulting network is phenotype.

1.2 Substrate

For this mechanic to work, every neuron needs to have defined its position. Likewise, every possible connection needs to be defined as well. This construct is called substrate. Dimensions of substrates can vary, most substrates used are 2D and 3D, depending on problem in question. Topology of substrate should reflect geometry of problem. This way HyperNeat can use information about topology [7].

By topology and dimensions we can categorize few basic substrate types. Most basic substrates are 2D and 3D grid. However substrates do not have to be aligned to a grid, topology can form any shapes. Sandwich substrate is restricted 3D structure, consisting of two layers, connected in one direction. One layer contains input nodes and second contains output nodes. This way input into CPPN can be minimized to 4 values.

1.3 Ms. Pac-Man

Ms. Pac-Man is famous video game released in 1982. Much like in original Pac-Man, player controls character, which is being chased by four ghosts. At the same time player should strive to collect all dots, by collecting all dots, level ends. Player can use four power-ups, which give player ability to eat ghosts, these also cause ghosts to run from player.

Unlike its predecessor, Ms. Pac-man features more than one maze and non-deterministic ghost behavior. These changes have varying impact both on human player and on possible game-playing agent. While humans can rarely exploit deterministic moves of original ghosts, some approaches can find best case scenario, rather than analyzing game state for best move or using some sort of strategy. Influence of new maps had greater impact on human player, however impact on agents depends on used approach. One such map is shown in Figure 1.

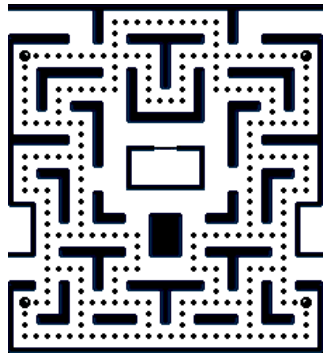


Figure 1. Maze.

In the past there were two notable competitions in Ms. Pac-man, namely Ms. Pac-Man AI Competition and Ms. Pac-Man vs. Ghosts Competition. While first named used screen-scraping approach, second competition used its own Java implementation. In Ms. Pac-Man vs. Ghosts competitors were able to submit controllers for ghost behaviour as well.

2 Related Work

HyperNeat and various versions of Pac-man were used in vast number of papers and studies, we will cover only few of those.

In 2007 work of Gallagher and Ledwich [1], neuroevolution of weights of fixed topology neural network was used. Input representation was handled by three matrices. Those were limited in size, centered around pac-man, essentially making pac-man oblivious to far end of maze. To combat this, four inputs were added, each representing information about dots beyond limited matrix. In most experiments only one deterministic ghost was used, however more complex variations were tried as well. Also all power-ups and bonuses were removed. Results of this study shows that agents were able to learn novice level ability.

Schrum in his work [5] used Modular Multiobjective NEAT to create agent with multiple modes of function. It was argued that Ms. Pac-man require multimodal behavior to evade ghosts at one time and trying to catch them if they are edible at another time. Worth mentioning is also fact that neural networks used were not fed with raw data. Rather there are 30 both direction oriented and non-direction oriented sensors, indicating information such as distances to ghosts, whether ghost approaches, whether ghost is edible, distance to nearest pill, power pill and many other types. Also direction evaluation is done by choosing highest valued direction. In this paper same implementation of game was used as we are using.

HyperNeat was used in this paper [3] to produce agent that is able to play two different Atari games with similar game screen topology. Screen scraping approach with additional visual processing was used to fill input layer. Substrate used is of sandwich type with very similar output processing as in this paper. For neural network to operate, objects must be identified on input layer, this was done by assigning each class of objects some integer value, however authors speculate that this approach is effective only with small set of classes.

3 Approach

3.1 Game implementation

Game implementation used in this paper was java clone created for Ms. Pac-Man vs. Ghost League [4]. This implementation allows user to implement and use his own controllers, for both pac-man and ghosts. Game has several modes of operation. In competition, game behaves asynchronously with time limit for controllers. This means that each tick, game uses move provided by controllers to update game. However, controllers have only 40 ms each to generate new move, if they fail, previous or random move is used. This limits complexity of controllers submitted.

Since we expected that our evolved networks will became rather complex, we could not used mode that was used in competition. Luckily for us, two other modes work differently. They always wait for controllers to determine next move. This way, complex networks are not handicapped and can compete in evolution. This was not used in competition, it was used merely for display and experiments.

To save some computational time, we changed number of pac-man lives to one and changed initial ghost lair times. Also during evolution no visuals are shown and game does not wait for player to be ready.

Our goal is not to finish level, but to evade as long as possible. This is why, we chose number of player-ghost move cycles as a fitness score rather than actual game score.

3.2 Substrate

Anytime we work with HyperNeat we need to think about substrate, which is integral part of this neuroevolutionary method. In [2, 3] sandwich substrate or its derivatives were successfully used for

board games or video games. We could argue that sandwich substrate is natural solution for this type of problems.

However pac-man map topology is different from game board. If we discrete game screen by matching grid, we can clearly see that in all four mazes in Ms. Pac-man, traversable cells are actually a minority of all cells. This allow us to add only those cells to substrate, which are traversable. This lessens number of connections and number of object classes on other hand we need different substrate for different mazes.

3.3 Representation of information

In [3] few classes of objects were successfully represented in neural network by integers, we opted for same approach and decided to designate pac-man as one and ghosts as negative one.

In used implementation, mazes are defined in separate files as a list of nodes, where each node has attributes such as neighboring nodes, pills, powerpills and other. However, these mazes are four times as big as they need to be. Reason behind this is to create illusion of smooth movement. Since we want our substrate to be minimal, we chose to create reduced maze from which we make substrate.

This creates difference between original maze, which is still used by game and our substrate. Even if we transform coordinates between substrate and maze we can not avoid rounding errors. These happen when pac-man arrives at intersection. Due to rounding, in substrate pac-man is already at intersection, however in game, pac-man has yet to arrive at said intersection.

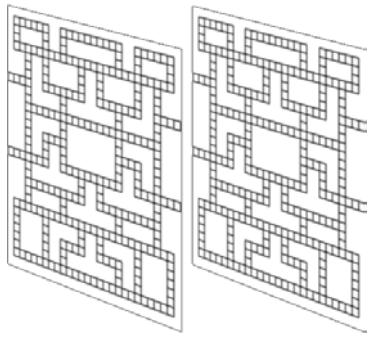


Figure 2. Schema.

Schema of resulting network will resemble that in Figure 2. However any neuron from left layer can be connected to any number of neurons on right layer.

3.4 Output processing

To determine next move from neural network, we can design neural network to return value for each possible move or we can return values of each node and process them with hard coded methods. Since second option is more natural to sandwich substrate, we chose this approach. This way nodes of output layer are painted by their desirability. We designed several hard coded methods to chose next move, on basis of values achieved by nodes neighboring pac-man.

These are rather similar to each other. First one is to simply pick node with greatest value. Due to its implementation, if more than two or more nodes have same value, priority is on move farther clock-wise, so from least prioritized to most is up, right, down, left. Advantage of this method, is its simplicity, this means that neural network must rely on itself and not on method used.

Second method is extension of first method. It improves on cases, when more than one node has same value. This method prioritizes previous move, it results in a much more smooth movement.

Using this method even weak individuals can achieve good score, however this is result of method itself and not neural network.

Third method is even more sophisticated than previous one. It improves on case when previous move is not possible but all remaining moves have same value. When this case happen method chooses randomly.

We can order our methods by their generality. We suspect that more general methods are better in the long run on the other hand, more sophisticated methods are better in short runs.

4 Results

We tested each direction picking method five times, averaged values were put into graphs below. Score calculation was averaged over three attempts. During these experiments, initial pac-man position stayed always the same. In our experiment we used example ghost controller called Legacy, in this controller every ghost uses different pathfinding method and one moves randomly. Population size was set to 500 individuals and maximum generation count was set to 250.

Rarely, an exceptional evolution happened. Evolutions, we call exceptional were performing far better than majority we observed. Sadly almost none of these happened during main experiments. This is why we added fourth graph, which contains individual evolutions that got our attention. Most of these did not for various reason run for 250 generations.

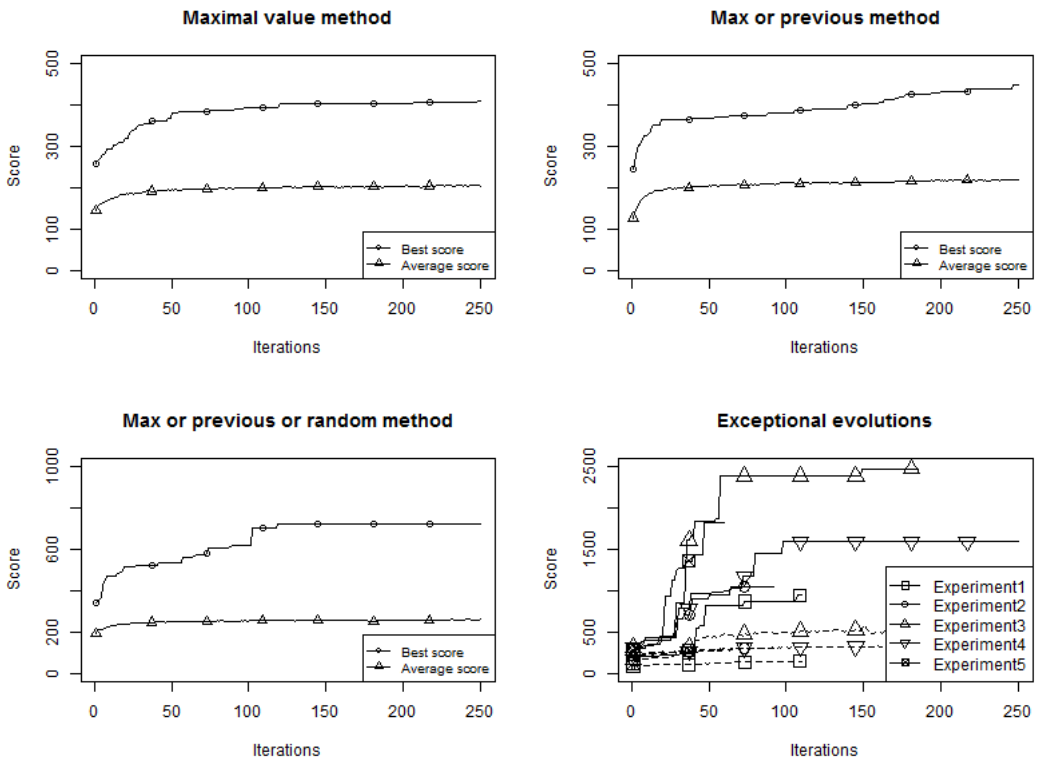


Figure 3. Results.

Our results are shown in Figure 3. In all cases we can see rather sizable difference between average and best score, this is expected. However, average score stagnates or rises very slowly with all three

methods after first 50 generations. It seems that maximal value or previous move method performs better. After reviewing individual logs we conclude this happened due to semi-exceptional evolution, which can happen with every method used and thus it does not prove that this method is superior.

Using maximal value or previous move or random move method, we seemingly achieved two times as good results. However, this seems as a global increase and curve has same shape. Thus we can conclude this method results are not caused by better neural network performance but thanks to method itself.

Aside from main experiments we carried out few more. Some of these ended prematurely due to device failure or were inadmissible to main experiments due to small code, methods and settings changes. We picked some exceptional evolutions from them. Experiment 2 and 3 used another, easier to evade ghost controller. Experiment 1 used changing initial pac-man position to encourage more general approach for evading. Due to low average scores and harder learning this method was not used. All experiments except first one used maximal value or previous move method.

5 Conclusion and Future Work

Our results were too inconsistent to safely draw any conclusions about best methods used. Inconsistency and observation of exceptional evolutions could be caused by too small population. However all three methods and even unexceptional evolutions did manage to significantly improve their abilities. It is possible that evolution was too short for maximal value method to outperform more specialized methods.

Our future work will try to extend this method to pill collection as well as more fine tuning of methods and settings of algorithm. With more time to each evolution and bigger population size, efficiency of this approach could be dramatically increased

References

[1] Gallagher, M., Ledwich, M.: Evolving Pac-Man Players: Can We Learn from Raw Input? In: *Computational Intelligence and Games, 2007. CIG 2007. IEEE Symposium on*, 2007, pp. 282–287.

[2] Gauci, J., Stanley, K.: Indirect Encoding of Neural Networks for Scalable Go. In Schaefer, R., Cotta, C., Ko lodziej, J., Rudolph, G., eds.: *Parallel Problem Solving from Nature, PPSN XI*. Volume 6238 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2010, pp. 354–363.

[3] Hausknecht, M., Khandelwal, P., Miikkulainen, R., Stone, P.: HyperNEAT-GGP: A hyperNEAT-based Atari General Game Player. In: *Proceedings of the 14th Annual Conference on Genetic and Evolutionary Computation*. GECCO '12, New York, NY, USA, ACM, 2012, pp. 217–224.

[4] Rohlfshagen, P., Lucas, S.: Ms Pac-Man versus Ghost Team CEC 2011 competition. In: *Evolutionary Computation (CEC), 2011 IEEE Congress on*, 2011, pp. 70–77.

[5] Schrum, J., Miikkulainen, R.: Evolving Multimodal Behavior With Modular Neural Networks in Ms. Pac-Man. In: *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO 2014)*, Vancouver, BC, Canada, 2014, pp. 325–332, Best Paper: Digital Entertainment and Arts.

[6] Stanley, K.O.: Compositional Pattern Producing Networks: A Novel Abstraction of Development. *Genetic Programming and Evolvable Machines*, 2007, vol. 8, no. 2, pp. 131–162.

[7] Stanley, K.O., D'Ambrosio, D.B., Gauci, J.: A Hypercube-based Encoding for Evolving Large-scale Neural Networks. *Artif. Life*, 2009, vol. 15, no. 2, pp. 185–212.

[8] Stanley, K.O., Miikkulainen, R.: Evolving Neural Networks Through Augmenting Topologies. *Evol. Comput.*, 2002, vol. 10, no. 2, pp. 99–127.

Parallel Bucket Sort Algorithm for Ordering DNA Short Reads Data Sequences

Tomáš FARKAŠ*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
farkas0tomas@gmail.com

Abstract. Proposed paper deals with optimizing and enhancing performance of parallel Bucket Sort algorithm. First we will discuss its suitability for this kind of data. Then, we propose concrete configuration properties, whereby some of them are valid for the latest Intel architecture and the other in general. For the Bucket Sort algorithm we find the optimal bucket size and suggest an idea for improving its performance by about 40 %.

1 Introduction

In the modern computer era, parallelism tends to be the most dominant way of increasing computing capabilities. However implementing basic algorithms in parallel is much more complicated, while developers need to consider many options for to avoid possible mistakes. Not doing so, while processing huge amount of data such as DNA read sequences, running times may rise and the program - even giving correct results, becomes useless.

Data sorting is a crucial part of most DNA oriented applications such as DNA sequence assembling, genome indexing and BWT transformation-based compression [10]. Many papers have been published that cover this field of study but they in general do not suggest and compare specific strategies and configuration properties. It is correct, that some of them are platform-specific. In the following text we would like to move our attention to those properties that are valid in general.

2 Data Suitability Analysis

NGS DNA data in the genome assembly process is formed as k-mers, typically 31-mers consisting of four types of nucleotidic bases [5]. In fact, 4 different values can be coded in 2 bits, so the entire information can be stored as 62 (64)-bit long integer and generic sorting algorithms can be used. However, Bucket Sort, for being $O(n)$, requires near-uniform data distribution. To see the

* Bachelor study programme in field: Informatics
Supervisor: Assoc. Professor Mária Lucká, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 77–82.

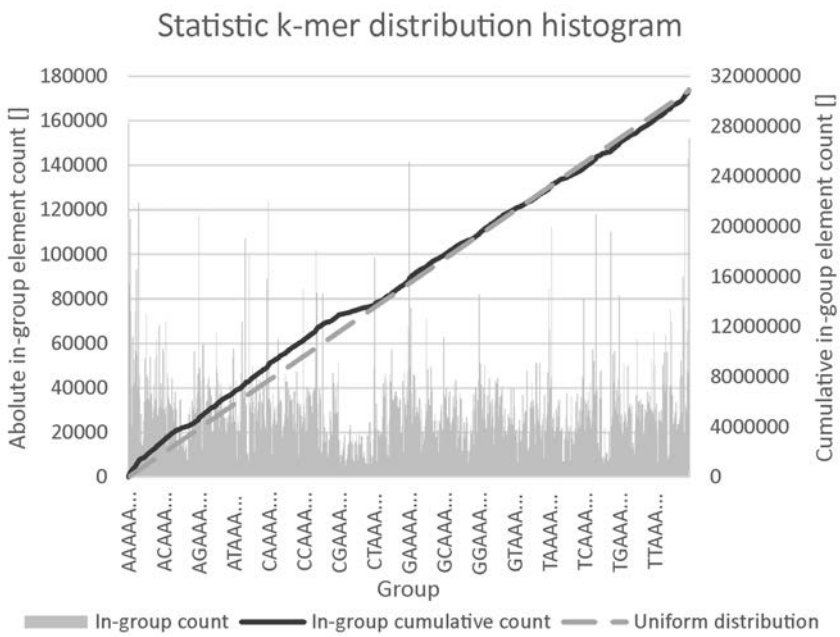


Figure 1. DNA data statistic distribution histogram.

distribution of real DNA sequenced data, we counted a histogram of random part of SRX040485[1] dataset, split into 31M 64-bit samples in the way the assemblers do. The results, as shown on Figure1, have proved near-uniform global distribution. However there are some variances of 1.95 % (604769 elements) in average. Still, after being indexed to 1024-element long sequences, what is hardly achieved, as we will discuss later, remains the average difference of real distribution to the uniform linear one of 54.4 elements. The difference is too high for Insertion Sort-like algorithms that perform well on nearly sorted data, but highly suitable for Bucket Sort.

3 Bucket Sort

Bucket Sort [2,6] is an efficient non-comparative sorting algorithm based on reducing of logarithmic factor present in the complexity of comparative sorting algorithms. The main idea behind is to set specific bounds across the possible data values span, to divide the elements into groups – buckets and finally, to sort each bucket separately. This can be safely done in parallel. Unless any comparative data sorting algorithm, Bucket Sort runs in $O(n)$ time, when using the constant bucket size.

- Best time complexity: $O(n)$, when condition above is satisfied.
- Typical time complexity: $O(n \times log(\frac{n}{b}))$, for *b* as fixed number of buckets.
- Worst time complexity[2] $O(n \times log(n))$

[1] Dataset SRX040485 – Drosophila melanogaster. Available online: http://sra.dnanexus.com/experiments/SRX040485/sample

[2] This is in controversy to [6]: evaluating wost time complexity to $O(n^2)$. However worst complexity can be no worse than complexity of the final-step comparative sorting algorithm. In the case of Quick Sort or Merge Sort it means $O(n \times log(n))$.

– Space complexity: $2n + k$.
Extra memory of $O(n)$ size is required for creating buckets. After this step, original memory can be freed or used as auxiliary memory for Merge Sort algorithm.

The algorithm seems simple to implement, but there are few problems to resolve, that will be described in the following 2 sections:

3.1 Data splitting policies

Data splitting is the most vital part of Bucket Sort algorithm. In this step we need to resolve questions about buckets's memory allocating and the concurrent data access. The establishing of array sizes for buckets can be performed in two ways:

– Once finding out statistic quantities for special kind of data, then allocating memory according to this with a little reserve. This is especially interesting, as the gcc compiler is able to allocate array on-the-run, which consumes short, but significant amount of time.

– Performing an extra pass to find out the exact size of each bucket. In this situation, cumulative quantities can be used as indices to n-element long array, so there will be no need of copying back the buckets for creating a complete sorted array.

As we have found, the extra pass needs only one half of the time needed for any of the next passes and no extra copying of data, so we recommend this. However, the run times may vary according to the used platform. The first approach can be also useful for clustered MPs MPI implementation, where (1): Data is stored externally and never in memory on the whole, so another pass would be time consuming (2): on-the-run memory allocating takes no extra time, as is much faster then transferring received data, (3): data collecting is needed anyway.

Multi thread processing requires concurrent data access. Basic critical section, even if using multiple mutexes for multiple array parts blocks the other threads and prevents good parallelization. Mutex locking and unlocking is also time consuming and should not be used after processing each element. So there will be need for data buffering – creating smaller buckets separately for each thread and copying multiple elements to the destination simultaneously. The run time decreases with the growing buffer size. While storing data, cache misses do not affect the running time, so the buffer size can be increased safely.

3.2 Determining of the bucket's size

During the sorting process, comparative algorithm performs $log(n)$ passes through $O(n)$ element long array. This can be clearly seen on iterative bottom-up Merge Sort implementation. Reducing the size of the buckets to $1/2$ results in reducing the number of passes required for correct sorting of data by one, but advantage of the decreased array length itself is canceled by doubling the number of buckets. So there might be a theory, that the more buckets is used, the less time for data sorting is needed. This fact follows from the inversely proportional number of passes to the logarithm of the number of buckets.

The graph below shows constant time required for 1st pass completion, linearly decreasing time of sorting the buckets, what is correct due to logarithmic axis x scale. However the time required for the 2nd pass tends to grow rapidly to the end of measured interval and when reaching the 10000 border, its growth cancels the benefit of decreasing logarithmic factor of complexity and overall time starts to rise. Now, let's discuss the cause of the growth: Increasing size of the index-table may lower cache-hit-ratio and Intel Performance Counter Monitor[3] tool proves this theory. Whereas splitting

[3] Official tuning tool published by Intel inc. Available online:
https://software.intel.com/enus/articles/intel-performance-counter-monitor

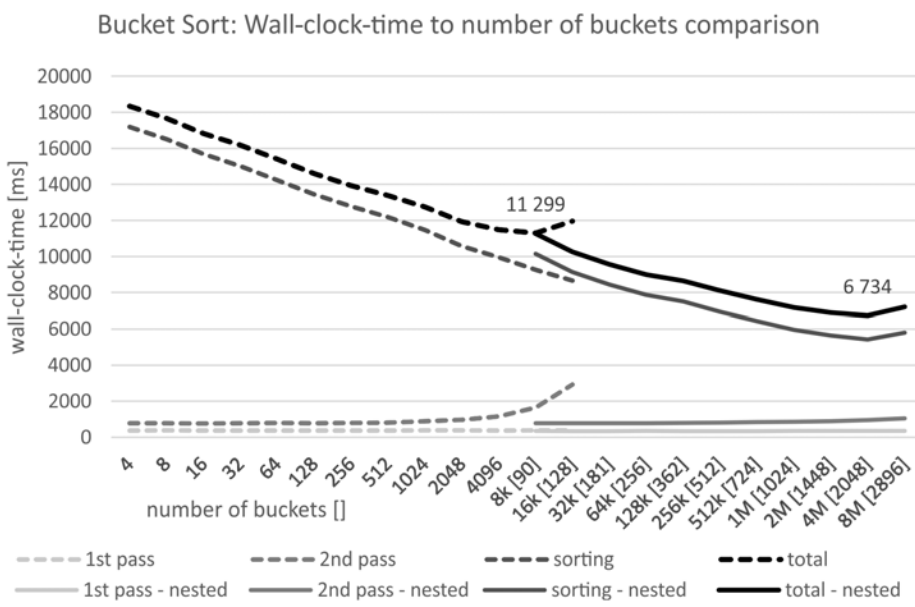


Figure 2. Bucket Sort: Wall-clock-time to number of buckets comparison. 5 runs average data.

the array into 4 buckets produces 3500 kB/s L3 cache misses and 4800 kB/s L2 misses, processing 2nd pass with 16384 element-long table results in 13 MB/s L3 and 105 MB/s L2 misses. This also reduces the number of completed instructions from 10 G/s to 6800 M/s. These values are platform specific.

3.3 Nested Bucket Sort

As shown in the previous section, the number of levels and corresponding size of buckets cannot be set to be optimal. However nothing prevents us to use Bucket Sort algorithm also for sorting buckets themselves. This version can be called Nested Bucket Sort and it not only allows to use much more buckets effectively, but it also reduces the total memory overhead coupled with using huge amount of buckets in the non-nested version. Basic idea of this algorithm, to be a mixture of Bucket, Radix [4] and American Flag Sort [7] is shown on Figure 3. The first two steps may look similar to MSD Radix Sort, however the idea is slightly different. Proposed algorithm does not rejoin the splitted partitions before the next step, even if they are stored in the same array. The original Radix Sort has 15% poorer performance due to long (64-bit) word and higher number of passes needed, as we have discovered. The enclosed diagram compares measured wall-clock-times in meaningful range of the nested version to the overlapping part of the original – Bucket Sort graph. This does not propose the 4M border to be ideal for any size of data, though, but using buckets containing a few hundreds of elements seems to be the right choice if using the nested version.

4 Conclusion

As we have shown, properly chosen parameters used in known algorithms can significantly change running wall-clock-times. For practical reasons, we chose dataset containing 10^8 8-byte integers for tuning the algorithms. This dataset is much smaller than those when sorting real data. Due to this fact, however the matter of seconds changes to the matter of minutes and hours and so every misconfiguration can be critical. Therefore we propose:

```
bucketSort (data, aux){
        //counts buckets starting position indices
        histogram <- bitWiseHistogram (data);

        //copies data content to~corresponding indices
                        //in~an~auxilliary data space
        aux <- bitWisePartition (data, histogram);

        //sorts aux array parts locally
        foreach (partition: aux)
                comparativeSort (partition);
}

bucketSortParallel (data, aux){
        //Map-Reduce-like frequency count
        histogram <- parallelBitWiseHistogram (data);

        //buffered data copying
        aux <- parallelBitWisePartition (data, histogram);

        //runs multiple instances of Bucket Sort, reusing old data array
        #parallel
        foreach (partition: aux)
                bucketSort (partition, data);
}
```

Figure 3. Pseudocode showing Nested Bucket Sort principle.

- Not to overcome the border of 10000 when setting the number of buckets for not-nested Bucket Sort algorithm version.
- To use nested Bucket Sort algorithm version described above if sorting data requiring more than 8k buckets.

Acknowledgement: This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG 1/0752/14.

References

[1] Butenhof, D. R.: *Programming with POSIX Threads*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1997.

[2] Cormen, T. H. et. al.: *Introduction to Algorithms*. 2nd edn. McGraw-Hill Higher Education, 2001.

[3] Hagit, A., Welch, J.: *Distributed Computing: Fundamentals, Simulations and Advanced Topics (2nd edition)*. John Wiley Interscience, 2004.

[4] Hollerith, H.: US. pat. Nr. 395781, 395782, 395783.

[5] Jackson, B. G., Aluru, S.: Parallel Construction of Bidirected String Graphs for Genome Assembly. In: *Proceedings of the 2008 37th International Conference on Parallel Processing*. ICPP 08, Washington, DC, USA, IEEE Computer Society, 2008, pp. 346–353.

[6] Knuth, D. E.: *The Art of Computer Programming, Volume 3: (2Nd Ed.) Sorting and Searching*. Addison Wesley Longman Publishing Co., Inc., Redwood City, CA, USA, 1998.

[7] Mcllroy, M. P. et. al.: Engineering Radix Sort. *Computing Systems*, 1993, vol. 6, no. 1, pp. 5–27.

[8] Quinn, M. J.: *Parallel Programming in C with MPI and OpenMP*. McGraw-Hill Education Group, 2003.

[9] Radenski, A.: Shared Memory, Message Passing, and Hybrid Merge Sorts for Standalone and Clustered SMPs. In: *PDPTA'11, the 2011 International Conference on Parallel and Distributed Processing Techniques and Applications*, 2011, pp. 367–374.

[10] Seward, J.: On the performance of BWT sorting algorithms. In: *Data Compression Conference, 2000. Proceedings. DCC 2000*, 2000, pp. 173–182.

[11] Čerňanský, M.: Paralelné programovanie. http://www2.fiit.stuba.sk/ cernans/pp/detailes.html, Lectures 2014.

Determining the Parts of Speech in Slovak Language

Dalibor MÉSZÁROS*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
xmeszarosd@stuba.sk

Abstract. In last few years, machine learning has made big progress and majority of approaches moved from manually created rule-based systems to data-driven approaches. That is also the case for part-of-speech tagging. In this paper, we focus on part-of-speech tagging in Slovak language. We try to adapt the current state-of-the-art approaches in English language, which are based on conditional random fields and achieve 97.3 % accuracy. We evaluate our approach on annotated dataset obtained from Slovak Academy of Sciences.

1 Introduction

In the past, part-of-speech (POS) tagging in Slovak language seemed much easier than in English, since many tags can be determined easily by word suffix. But as English is dominant language not just among the researchers of natural language processing, much work has been done. Since then, machine learning has made a big leap of progress and nowadays, there are state-of-the-art machine learning methods for part-of-speech tagging, which gradually become a standard [2].

Currently, all freely accessible solutions for Slovak language are based mostly on using dictionary lookup for part-of-speech tagging. Such solutions are of course usable, and they can even provide sufficiently good results in some cases, but it is not universally usable for many other languages because of high inaccuracy. For English, this kind of solution is absolutely unusable, because English language suffers from words, which can share the same part-of-speech tags, if the solution does not consider the context, like for word "being", which may be noun or a verb. Slovak language is not limited as much, because words have most of time multiple meanings, but only rarely share part-of-speech tags. But the problem raises because of rich morphology, which causes used dictionary to be much greater in content, but less sturdy. Moreover, dictionary based methods fail to tag words not present in the dictionary.

As for many languages, the challenging problem for a researcher is to consider and define all the rules, which are used for determining the part of speech tags based on a word and its context. Whole fragile dictionary and pile of unique rules forced us to look after a solution, which could use machine learning for determining the relationships and using context. For this task we chose

* Bachelor degree study programme in field: Informatics
Supervisor: Márius Šajgalík, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 83–88.

conditional random fields, which proved to be reliable for English language, because they do not need dictionary itself. They also proved to be an efficient solution for other languages as well.

Conditional random fields (CRF) belong to statistical and graph models, concretely under category of discriminative models. Those can be used to categorize words into parts of speech based on their unique features, context and probability based on weight, which represents the chance that a word belongs to certain part of speech. Because model is not trying to represent individual words, but rather sequences, this solution may look as superior to a dictionary solution in many ways.

2 Related work

There are few works for part-of-speech tagging problem in Slovak language. Noteworthy are reports from institutes and individuals who are focusing on certain aspect of Slovak language and working on projects [1], which could help others to progress in natural language processing for Slovak language further. However, regarding part-of-speech tagging, there are only few mentions and no results that could be used for comparison.

Morphological analyzer[1] is a tool of Ludovit Stur Institute of Linguistics, which in summary uses multiple existing dictionary databases for Slovak language, to gather any information about the searched word. The dictionaries ranges from regular dictionaries to historical, phrasal and even dictionary of surnames containing meanings of the word and grammatical categorization.

Slovak POS Tagger[2] is a projects of Faculty of Informatics and Information Technologies, which can determine part-of-speech tags and grammatical categories in Slovak language. On the outside, this solution is capable of the same thing we are trying to achieve, because this service can return results with POS tags, but also with the rest of possible tags, which had lower probability chance. But on the inside, the service is limited by the used dictionary. This way, even if it chooses correct answer with high accuracy, it cannot predict POS tags for unknown words, such as slang, technical terms, foreign words or even partially misspelled words. In this case the tool marks the words as indeterminable.

Slovak Lemmatizator[3] is another service and project of our faculty. The goal of this project and web service is lemmatization of Slovak text. That means finding out the basic form of the word, from which is the inputted word derived. This tool supports two types of lemmatization: fast and complete. In case of fast lemmatization output is inputted text, with each word substituted for its lemma. In case that the tool cannot choose lemma for a particular word, the word is unchanged. For complete lemmatization output is XML text, where is each word tagged with correct lemma and other possible lemmas. To sum up, this tool cannot determine parts of speech. It can tokenize sentences and using the context, it chooses the best possible lemma for a word.

Stanford Part-Of-Speech Tagger is represents solutions, which are known as universally usable for many languages. It is implemented in JAVA and is able to train for any language and is even distributed with trained models for few languages: English, Arabic, Chinese, French, and German. From the English language point of view, was Stanford POS Tagger able to achieve 96.86 % accuracy [5], and there are even reports about achieving standard level of 97.24 % accuracy [6].

It has been shown on languages such as Czech, Finnish, Estonian and Romanian [8] that CRF is capable of determining not even parts of speech, but even grammatical properties with high accuracy using more complex approach in combination with simple features, which we are also using in this work. Basic features are still powerful, when used at small collection of tag set and may have effect at tagging certain grammatical categories.

[1] *http://slovnik.juls.savba.sk/*
[2] *http://morpholyzer.fiit.stuba.sk:8080/PosTagger/*
[3] *http://text.fiit.stuba.sk/lemmatizer/*

3 Part-of-speech tagging

The process of part-of-speech tagging can be divided into several steps. It begins with raw text without any annotations and outputs sequence of tagged words. These are the required steps:

1. Tokenization
2. Determining word features
3. CRF model computation
4. Post-processing of CRF output

3.1 Tokenization

This is the first step, which assumes, that the input text consists of sentences, because the POS tagger is more efficient, if it can fully use the context in determining parts of speech.

First step is therefore tokenization, which breaks the text into separated sentences with words. This step is done by using the Stanford Tokenizer. This way will be words separated as one word per line and sentences will be separated by empty line. This is done, because it makes rest of text processing much more trivial and supports ability of CRF over context.

3.2 Determining word features

Second step, is to determine features for each word. Those are also predetermined for words in training, validation and testing set. Because features need to be generated from an unknown words, so they can help in categorization with CRF, they were kept simple.

The most commonly used features are:

- Prefixes and suffixes

  Which are a substring from the beginning and end of the word respectively. For our experiments we used pairings of varying ranges, from which we chose to present ranges from 1 character up to 3, because further they have minor or no effect at all on the tagging accuracy.

  Prefixes cannot unveil the part of speech of words as accurately as suffixes, but we can gain additional information based on the first letter. For example a word starting with capital letter may be most likely a name of person, with higher chance or weight on being a noun. This is also further supported by transformation into lowercase word.

  Suffixes are more reliable, because we have more cases, in which we can categorize words into parts of speech based on the ending of the word. Like in the case of "ť", there is a high chance that the word is a verb, or in the case of "ý" there is a chance, that the word is adjective.

- Words in lowercase

  Because the CRF model is considering words and features as mere strings, there is a reason to transform words into purely lowercase words and generate additional features from it. This way characters would not be mistaken, if there is comparison between uppercase and lowercase of the same character. This way we lose the information, where was uppercase character and therefore it is required to save the information in a feature. We use simple number to represent the state in which the word was:

  - 0 – The word was assembled from lowercase characters
  - 1 – The first letter in the word is capital
  - 2 – The word had somewhere else an uppercase character, this is the rarer possibility
  - 3 – The word was assembled from uppercase characters

- Vector representation of words

Currently we are working at integrating vector representation of words, which could be used to find words, which have closest relationship with the tagged word. This way we could include a powerful set of features, which may greatly improve tagging accuracy.

3.3 CRF model computation

We decided to use an existing implementation of CRF [7]. The inference process follows a set pattern, where the CRF looks and focuses at each word separately along with their features. Then it looks at the surroundings of the word. That means that it takes into consideration neighboring words and also their features, when determining relationships and predicting parts of speech.

The output of the CRF model computation is text consisting of predicted labels for the words. Such output requires to be mapped on the original words, so it may be processed further.

3.4 Post-processing of CRF output

Last step is to interpret and process the raw output text into required readable format. As was mentioned, the output of CRF model are labels only, which may be mapped at the input text using notation of Slovak national corpus, for example text would look like this: "… v:E starovekom:A Egypte:S…" or mapped on our tokenized text, whichever would be easier to process further for user.

Also the output could be formatted in more user friendly way, such as using full name of part of speech: "…v Preposition, starovekom Adjective, Egypte Substantive…"

4 Evaluation

4.1 Text corpus

Training, validation and testing sets were assembled from text corpus of Slovak Wikipedia and Necyklopédia, which were accessed by Slovak Academy of Sciences [4]. These text consists of text from web pages of Wikipedia and Necyklopédia hierarchically separated by XML tags into documents, sentences, blocks. Each word along with tags (parts of speech, grammatical category, lemma), are separated by tabulator space as special character. The text contains 204 943 documents, containing 37 548 997 words in 3 360 218 sentences.

4.2 Training and testing dataset

These texts required some processing, before they could be used. Main task was to divide the text into sets: training, validation and testing in ratio 60:20:20 respectively. We chose this ratio, because 60 % of training set represents the base, which is used to create the model and is required to be greater than testing sets. Validation set will serve for choosing the feature template, which has reached the highest accuracy and choosing the best possible approach. Testing set is used to evaluate representative accuracy achieved in this work with chosen template.

Texts are divided into following sections: documents, their sentences and words with belonging tags. Division by the ratio was made based on the number of documents, so the context would preserve. This way we are not keeping the exact ratio and there may be small difference between amounts of words for each set, what we took into consideration. Documents are also distributed randomly between the sets. This way we can find different documents between each set and we can always generate different sets or generate old ones by using stored seed.

From this dataset we extracted only words with their parts of speech as labels for CRF. Lemmas became unusable for our project, because there is no way to guarantee correct prediction of lemma, before we even start determining parts of speech. Also this kind of solution would go against philosophy of our solution, because it is a solution based on dictionary.

We decided to keep grammatical categories along with parts of speech, because this way the POS tagger could create more categories for words with different set of rules. Those would be ignored in the final processing of CRF output.

4.3 Feature templates

When generating training, validation, testing sets or tagging an unknown set, the most important factor is, how good are the features used for determining the parts of speech. Also there raises a question, how the determination of features for unknown words affects the accuracy or even complexity of the whole solution.

For testing purposes we created variations or templates for used features we decided to use. This way we created combinations based on the use of lowercase solution and length of prefix and suffix features.

In the case of lowercase solution, we divided the possible feature templates into two groups. The one which used the original unchanged word, even for prefix and suffix. And then the second, which used the word with lowercase characters only alongside with feature, which described, if the word contained an uppercase character, how many and where were they positioned. We questioned, how this would affect the accuracy of our solution in the end, because machine processing of CRF treats the words as mere strings.

Then we expanded the templates by prefix and suffix with length 1, 2, 3 as described and used by Google N-grams [2]. This way we could test how much would the prefixes and suffixes length affect the accuracy individually, because they still can be used all at once for a single template.

Table 1. Accuracy of specific Feature Template used on the same data.

Case indicator	Prefix & suffix length	Accuracy
Yes	1	90.58 %
Yes	2	92.91 %
Yes	3	92.99 %
No	1	89.92 %
No	2	91.98 %
No	3	91.98 %

On these six simple templates used on a testing set we can see, how strong is each combination based on the accuracy of predicting words.

As we can see, lowercase word pushed the accuracy higher than in the case of simple word. We can see, that words are indeed considered as mere strings, and there is a need to extract as much information from the original words, as possible, such as the information about the uppercase letters. The almost 1% accuracy increase may be explained by the words at the beginning of the sentence, where these words were treated differently because of capital letter.

Also as we can see, the lowest accuracy is achieved by prefixes and suffixes length being 1. We can assume that as we use wider prefixes and suffixes the accuracy will rise, but also will stall at some point as we can see in the case of prefixes and suffixes length of 2 and 3.

We would like to note, that in final version of project we are aiming to use all three length variations in a single template and this way we wanted to show which length has greater effect at tagging accuracy independently.

5 Conclusion

In this paper, we evaluated the current state-of-the-art method commonly used for English language and adapted it for Slovak language. We have shown that with correctly trained CRF model with simple features, such as used in Google N-grams [3], we can achieve almost 93 % accuracy for Slovak language. This accuracy is showing, that a non-dictionary method is usable also for Slovak language and also that the accuracy is high enough considering the simplicity of feature template, which may be improved in the future by tweaking the features or bringing more complex solution alongside the CRF.

Presented POS tagger has potential to be expanded with new methods, which would help in understanding new features. This work mainly focused at verification of fact, that CRF model may be used effectively for Slovak language, while using simple features.

Acknowledgement. This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG 1/0646/15.

References

[1] Laclavík, M., Giglan, M., Krajči, S., Furdík, K., Hluchý, L.: Dostupné zdroje a výzvy pre počítačové spracovanie informačných zdrojov v slovenskom jazyku. In: *1st Workshop on Intelligent Knowledge Oriented Technologies*, (2006), pp. 66–70.

[2] Lafferty, J.D.: Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data. In: *Proceedings of the Eighteenth International Conference on Machine Learning*, San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2001. pp. 282–289.

[3] Lin, Y., Michel, J., Aiden, E.L., Orwant, J., Brockman, W., Petrov, S.: Syntactic Annotations for the Google Books Ngram Corpus. In: *Proceedings of the ACL 2012 System Demonstrations*. Stroudsburg, PA, USA : Association for Computational Linguistics, 2012. pp. 169–174.

[4] Slovak national corpus (Slovenský národný korpus) – wiki-2014-02. Bratislava: Jazykovedný ústav Ľ. Štúra SAV 2013. Available at: *http://korpus.juls.savba.sk/wiki.html*

[5] Toutanova, K., Manning, C.D.: Enriching the Knowledge Sources Used in a Maximum Entropy Part-of-speech Tagger. In: *Proceedings of the 2000 Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora: Held in Conjunction with the 38th Annual Meeting of the Association for Computational Linguistics – Volume 13*. Stroudsburg, PA, USA : Association for Computational Linguistics, 2000. pp. 63–70.

[6] Toutanova, K., Klein, D., Manning, C.D., Singer, Y.: Feature-rich Part-of-speech Tagging with a Cyclic Dependency Network. In: *Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology - Volume 1*. Stroudsburg, PA, USA : Association for Computational Linguistics, 2003. pp. 173–180.

[7] Okazaki, N.: CRFsuite: a fast implementation of Conditional Random Fields (CRFs) [Online]. 2007. Available at: *http://www.chokkan.org/software/crfsuite/*

[8] Silfverberg, M., Ruokolainen, T., Lindén, K., Kurimo, M.: Part-of-Speech Tagging using Conditional Random Fields: Exploiting Sub-Label Dependencies for Improved Accuracy. In: *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. Baltimore, Maryland : Association for Computational Linguistics, 2014. pp. 259–264.

Utilizing Vector Models for Automatic Text Lemmatization

Ladislav GALLAY*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
ladislav.gallay@lentil.sk

Extended Abstract

Unifying various word forms is the first step in understanding the natural language. Lemmatization is the process of transforming the word into its root form – lemma. Understanding of each individual word is followed by understanding the whole sentence and the context of given document. Current approaches in this area are limited to knowing either full grammar rules or building the translation matrix from the word to its basic form [2].

While the former approach is impossible to fully support the fusional languages such as Slavic ones, the latter is hard to maintain and needs to be manually updated every time new word is introduced into the language. It requires a lot of human input and is error prone.

We have realized that these linguistic rules must be are already captured in the natural text. Vector models are capable of extracting these linguistic regularities from the text into mathematical vectors. The operation such as *vector('king') – vector('man') + vector('woman')* results in vector that is close to *vector('queen').* Recently the word2vec has been introduced [1]. It is capable of training the model using continuous bag-of-words and skip-gram architectures even more effective. The input is some meaningful text in natural language and the output is list of words and their latent vectors.

Based on the results of word2vec tool we have discovered that not only semantic but also morphologic relations are stored in vectors. This allows us to make a query *'dogs' to 'dog' is the same as 'cats' to* which results in the word *cat*. These regularities are shown in figure 1.

We propose new algorithm that utilizes vector model of words. The above example shows how our algorithm works. We take two reference words, for example *'vodníkom'* and *'vodník'* and try to find similar shift from the input word *'rybníkom'*. As the result we are given several words around the word *'rybník'* which is the root form. These words are ordered by the distance starting at the closest one.

We do not want user to put these pairs for each input. Instead a small list of those pairs will be provided to our algorithm. We expect that building such a list is not that expensive as building the whole translation matrix for every word. Our algorithm will then choose the best pairs to run with the input word. We propose several methods of choosing the pairs. The longest suffix seems

* Bachelor degree study programme in field: Informatics
Supervisor: Dr. Marián Šimko, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 89–90.

to be the most reasonable, because the words in same grammar category usually have the same ending. However we will test also random choosing or choosing the semantic related words first. Doing several iterations we end up with many words probably around the basic form.

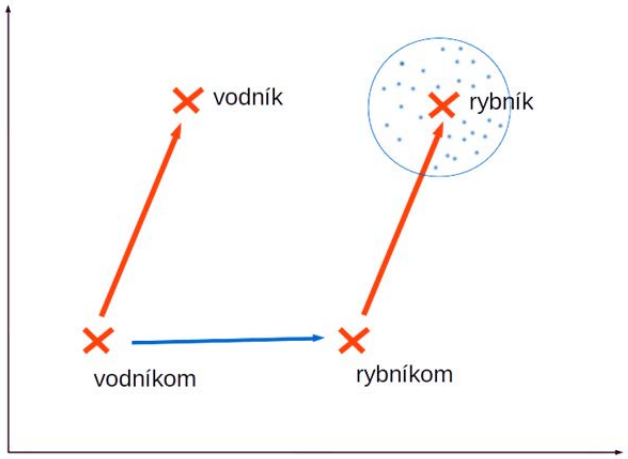


Figure 1. Relationship between vodníkom and vodník is the same as between rybníkom and rybník.

Then we need to filter the correct word. The words usually differ only in suffix and the prefix should remain the same. Although this is not necessary true for all languages, it can be applied in many cases for Slovak language. We calculate the relative prefix length multiplied by the distance for each output word. Having multiple iterations with the input word and various reference pairs the same output word can occur more than once so the results are summed. This number should represent output accuracy which we call *weight*. Following formula illustrates the calculation for given input word A, reference pair R and output word B:

$$weight(A,B) = \sum_{R=various\ pairs} \frac{prefix(A,B)}{\big(\,length(A) + length(B)\big)\,/\,2} * distance(A,R,B)$$

We have tested the algorithm on corpus from SME.sk and Wikipedia. The results were not impressive and are heavily dependent on how well the corpus is trained. Recently we were given the access to Slovak national corpus which outputs much better results. Our next work will include final testing and comparison of different methods of evaluation that we proposed.

Acknowledgement. This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG 1/0752/14.

References

[1] Mikolov T. et. al.: Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781, (2013).

[2] Horák, Gianitsová, Garabík, Šimková, M. Tokenizácia, lematizácia a morfologická anotácia Slovenského národného korpusu (2005).

Refinement of Methods for DNA Sequencing

Daniel KĹČ*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
xklcd@stuba.sk

Abstract. Widely used method for assembling short reads of DNA are NP-hard and struggle to assemble genome quickly, while addressing repeating sequences within genome and keeping results error-free. Described new method has polynomial complexity and offers solution for assembling repeating sequences while keeping results error-free. To further lower time requirements of devised algorithm, it was designed to be parallelized using MPI technology. Testing this new approach on synthetic data on single machine showed that all errors were eliminated and all contigs were assembled correctly. Results show high coverage by contigs, generally above 96 % at 50-fold coverage of synthetic genome.

1 Introduction

DNA sequencing is a field of study, which concerns itself with assembling short reads of DNA into original genome (or its parts). Most methods used today, are originally devised from graph theory [3]. All of these methods continually struggle to deliver better quality, lower error rate of results, and to reduce the cost and time required for a genome to be assembled. Results of these methods are currently used in medicine and biology.

Methods commonly used in DNA assemblers are proven to be NP-hard in [2]. Due to this fact, final goal of DNA sequencing, sequencing of human genome, is hard to achieve. To ease this problem, assemblers support various forms of parallelization. Due to RAM memory requirements and nature of methods used, many solutions struggle to use parallelization technologies in an effective way. As a result, parallelization takes place mostly on computer clusters, in a model where reads are distributed to nodes and then nodes communicate among each other in search of possible connections, which use big amounts of network resources for computation [5].

Furthermore, commonly used methods experience difficulties while processing similar regions in genome. Similar regions in genome "entangle" produced graph during assembly introducing structures like cycles, which are hard to resolve in a deterministic way.

To address the complexity problems, Medvedev et al. [2] suggest the use of methods devised from maximal flow problem, a well-known optimization problem. For this category of problems solution exists with quadratic complexity. We know of no widely adapted method, which effectively solves the second issue (e.g. graph cycles).

---

* Master degree study programme in field: Computer Engineering
Supervisor: Dr. Peter Lacko, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

During the process of reading, DNA errors are introduced to the data set. Every assembler offers tools, to deal with errors in reads at low error rate; however, with error rate above 5 % in the third generation of sequencing technologies [4], it becomes increasingly difficult to distinguish reads with errors.

In this paper we will describe new method intended to address both, complexity and similar regions issues, while keeping results error-free. Method is designed to be easily parallelizable on computer cluster. Methods used within this new composite method are inspired by several existing solutions and mathematical theorems.

2 Method outline

Newly developed method focuses mainly on lowering computation complexity, while keeping results error free, and offering a way to resolve internal repeats in genome in deterministic way. In addition, the internal architecture of the solution allows for effective implementation on computer cluster using MPI technology.

The method to achieve these goals consists of 5 autonomous steps. Each step uses output of the previous step as input, and produces input for the next stage. This design pattern is known as pipes and filters. Steps are as follows:

1. Initialization
2. Data distribution
3. Data processing
4. Results distribution
5. Graph processing

2.1 Initialization

In this phase, the data is read from the memory. Reads are used to initialize graph, which will be processed in the step 5. This graph is following maximum flow model. After graph initialization all reads are connected to source and sink. In maximal flow graph terminology, source is the starting point and sink is the ending point of the graph search.

To distribute the work among the working nodes we use keys. The keys are short sequences of DNA. Number of keys distributed to each working node is based on predefined key length. Set of keys calculated and later distributed among nodes represents all combinations of DNA sequence language (a, c, g, t) of given length.

Every working node is initialized with keys set. Based on key set each working node creates state automaton. This state automaton will in step 2 distinguish reads which are relevant to given node. Relevant read contains key assigned to given node.

The complexity of this step for main node is relatively small, with memory complexity $O(n + m)$, where n is number of reads and m is number of keys. Computational complexity is also $O(n + m)$, since we have added each read to the graph and created each key. Memory complexity on cluster node is $O(m)$. Computational complexity is $O(m)$, since each key has to be added to created state automaton on each node. The network usage complexity of this step is $O(m)$.

2.2 Data distribution

In this step data is distributed to the cluster nodes for processing. Each read is sent from the main node by broadcast on network. Broadcast is necessary since we don’t know which read is relevant to the given node at this time (and each read is relevant to multiple nodes based on key and read length).

When the node receives batch of reads, it will use the state automaton prepared in step 1 on each read. The state automaton will divide received reads on relevant and irrelevant for the given

node. Relevant reads are those reads, which contain keys assigned to the given node. Irrelevant reads are then discarded.

Memory complexity on the cluster node is technically O(n), where n is the number of reads received. However, since the network is the slowest component, reads are sent in batches and state automaton is very effective method to perform this kind of computation. Real memory complexity will be O(k + j), where k is number of relevant reads and j is the size of one batch. Computation complexity of step 2 is O(n) since each node has to process each read once. Network complexity is O(n) since each read has to pass through the network once (via broadcast).

2.3 Data processing

The method based on evidence presented here was originally inspired by the method presented in [1]. During the data processing for each relevant read with key as prefix extension sets are found and their multiplicity is computed. Notice that all the reads required for this are already present on the node, making further communication with other nodes unnecessary until results are sent to the main node. This makes this method ideal for implementation on computer cluster. Step is divided into 4 phases:

1. Division to S and D boxes – this divides relevant reads to reads with key as a prefix and all other
2. Computation of C boxes – this divides reads to sets which extend concrete read with key as a prefix
3. Reduction to E boxes – this divides read sets from 2 to subsets, which are excluding
4. E box filter – this filters sets from 3 based on achieved evidence for given extension.

During the division to S and D boxes relevant fragments from the state automaton are divided to S boxes and D boxes. S box is logical set of all relevant fragments which has concrete key as a prefix. D box is logical set of all relevant fragments, which contain concrete key on any index other then 0(which is prefix and goes to S box). S and D boxes are created for each key assigned to the given cluster node. This further reduces complexity of search for extensions for given relevant read. Memory complexity of this phase is O(m*(o + p)), where m is the number of keys, o is the number of S box members and p is the number of D box members. Computational complexity is O(j*h), where j is the number of relevant reads and h is the number of key positions within read (multiple relevant keys can be on multiple positions which is calculated in step 2) and j >> h.

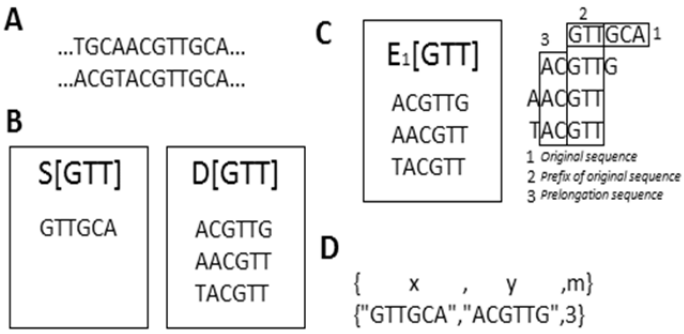


Figure 1. A – example of original genome. B – example of content of S and D boxes based on A. C – example of E box content based on B. D – example of information sent to the main node (in example istead of ID of sequence is sequence itself).

In next step C boxes are created. C box is logical set of reads from D box extending single read from S box (reads with key as prefix). Notice that reads in C box might offer different extensions

for given read from S box. Memory complexity is O(m*o*p) where m << o << p and computation complexity is O(m*o*p).To further reduce C box sets, E boxes are calculated from C boxes. E box is a logical set of reads which extends reads from S box and supports the shortest common extension. The shortest common extension is a read which extends the E-box base read (with key as a prefix) by the shortest sequence in set and is supported by all other reads in the E-box set. Notice that beyond the shortest extension length, members of the E box do not have to support each other further since it is unnecessary (this is due to the fact, that the shortest extension read is being analogically member of some S box on some node in cluster, therefore its extensions are also picked by this process later). Memory complexity remains the same as in the last step (since number of references remains unchanged). Computational complexity is O(c*r*log(r)), where c is the number of C boxes and r is the number of members in the C box (since some members of C box will be considered several times before it is assigned to the concrete E box).

The last step of processing is to filter the created E boxes. Filtering is achieved by considering the size of evidence for the given shortest common extension (there might be several for each read from S). If the evidence reaches predefined value, we consider it as a real extension. Otherwise we consider it a consequence of error in reads (or connections) and discard it. This eliminates almost all reads with errors, if minimal evidence parameter is set correctly. Memory complexity remains unchanged. Computational complexity is O(e), where e is the number of E boxes.

If a read is not extending any other read anywhere in the data set, it might be due to low coverage region of original genome in data set. However, it might be also due to an error in read causing wrong postfix which fits to no other read. This fragment would bypass E box filter, since at the time the read is on the node we do not know if the given read extends some other read. This problem can be easily solved in the step 5, therefore it is overlooked in the step 3.

2.4 Results distribution

During the results distribution, the E boxes are converted on set of three pieces of information <ID x, ID y, m>, where ID x is ID of read being extended, ID y is ID of extending read and m is the size of evidence for extension (multiplicity). These represent information gained from the step 3. We send all of these gained sets from all nodes to the main node to form the final graph of extensions in the form of maximal flow graph.

On the main node modification to the graph prepared in the step 1 is made. Upon reception of set <ID x, ID y, m>, read x is disconnected from source and y is added to it as predecessor with multiplicity m. When all sets are received, the final extension graph is finished.

The memory complexity on given node is O(e), where e is number of E boxes. Computational complexity is equally O(e). The memory complexity on the main node remains unchanged since no additional memory is needed. Computational complexity on the main node is O(x), where x is the number of found extensions, $x \geq n$. Network complexity is O(x).

2.5 Graph processing

During the graph processing contigs are created. In essence it will simply pass through the graph and add all extensions together forming contigs. In reality, however, few more things are needed.

In the first phase of graph processing, all the nodes still connected to the source and sink are found. Before the actual graph processing begins, clipping algorithm is applied eliminating impact of remaining reads with errors on results. This clipping algorithm will search the nodes connected to the source and sink and if the length of path starting at these nodes is equal or shorter then specified parameter (path ends by crossroad), it is clipped at crossroad since there is not enough evidence that the extension is correct at the start/end (notice that we might just have reduced future crossroad to simple path). If no crossroad is reached, path is marked and in the end of processing predefined number of bases will be clipped off to prevent the impact of errors in the reads on the results. Note that the results would be the same if we simply clipped all resulting contigs at the end. However, the graph would be much more complicated. The memory complexity remains

the same. Computational complexity is O(l*(so + si)), where l is the maximal length of clipped path, so is the number of nodes connected to the source and si is the number of nodes connected to the sink.

In the second phase, all simple paths are collapsed. Memory complexity remains the same, computational complexity is O(n), since every node in the graph is visited exactly one time.

In the third phase of graph processing we create raw contigs, or paths in the graph, that create contigs. This is different from simple paths since some crossroad can be deterministically passed while at the others contigs end. There are two types of crossroad:

- Y crossroad, where there are multiple incoming simple paths and one leaving, or there is one incoming path and multiple leaving. These can be solved if the multiplicity of incoming paths is close to equal to the multiplicity of the leaving paths. Otherwise incoming contig ends here and new one starts at crossroad (or multiple starts).
- X crossroad, where there are multiple incoming and multiple leaving paths. In this case we end contigs and start new ones (for each leaving path one).

In the phase three, the memory complexity remains the same. Computational complexity of phase is $O(n_p)$, where n_p is the number of simple paths, since every simple path is visited exactly once. Notice that Y and X crossroad can occur only when repeating sequence within a genome is present, so it is typical that contig is composed of only one simple path.

In the phase 4, the graph representation of contigs is converted into an actual base representation, and ends are clipped according to information from phase 1 (notice that contig end resulting from the crossroad is not clipped). The information about contig neighbours from graphs (result of crossroad) is preserved. The memory complexity is O(c), where c is the number of resulting contigs. The computational complexity is $O(n_p)$ since every path is visited exactly once during the process. Notice that if there is internal repeating sequence in a genome, it will be represented as a contig with higher multiplicity, which will have neighbour contigs noted. This allows us to resolve repeating sequences deterministically.

3 Testing

Method was so far tested on a single computer on synthetic genomes with various sizes up to 2^{18}. Between iterations of tests, the genome size was doubled to follow the rise in time and memory resources used at given size of input data. Synthetic genome reads represented 50-fold coverage at random indexes of genome, with 10% of read errors at the random indexes of read. Results are average of several runs on given genome size and are shown in the table below.

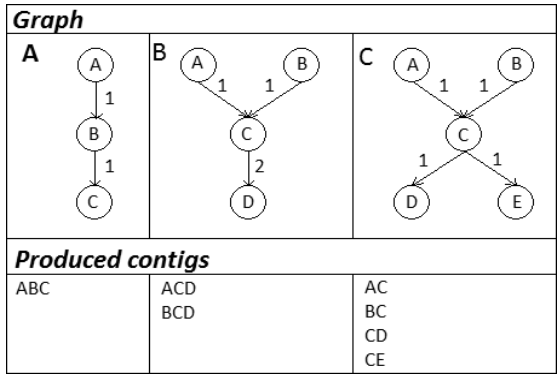


Figure 2. Example of produced contigs in three different situations.

Table 1. Results of described method on various sizes of synthetic genome.

Size of genome (bases)	Coverage by contigs (%)	Time required (s)	Memory required (MB)	Number of contigs (bases)	Average contig size (bases)
8192	98	3	16.2	11	741
16384	97	6.3	29.5	24	661
32768	98	14.6	54.6	48	671
65536	97.5	33.3	105.5	102	630
131072	97	87.3	208.8	208	619
262144	97	224	432.5	408	632

4 Conclusions

We have presented a new method for DNA sequencing. The method is based on maximal flow optimization problem, extension graph and searching for evidence of extension. In this method the read is extended only if enough evidence was presented to the algorithm. Then all the extensions are reformed into an extension graph, which is devised from the maximum flow graph.

Unlike conventionally used methods, we decimate all repeating reads and regions and then try to interpret them in terms of multiplicity and neighbouring contigs. This approach allow us to determine with relative precision number of times given repeat is present in genome. We believe this approach leads to more complete error-free information about real composition of genome.

In comparison with methods which inspired us complexity of our method is quadratic. No errors in assembled contigs were recorded when parameters were properly set making this method slightly more reliable (but to achieve this, ends of some contigs are clipped making coverage slightly lower). Finally our method require lower coverage by reads for high final coverage.

Our method has quadratic computational complexity. It introduces solution for internal repeating sequences in the genome by noting contig multiplicity and its neighbours in the extension graph, which can be used to resolve repeating sequences. By introducing extension evidence system and clipping base on heuristic, we have eliminated all errors from the results. Finally the method is devised in a way so that it can be efficiently implemented in a computer cluster with the MPI technology.

Acknowledgement: This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG 1/0752/14.

References

[1] J. Dohm, C. Lottaz et al. SHARCGS, a fast and highly accurate short-read assembly algorithm for de novo genomic sequencing. CSH PRESS, Genome Research Issue 11, (2007), Vol. 17, pp. 1697–1706.

[2] P. Medvedev et al. Computability of Models for Sequence Assembly. Springer, 7th International Conference, WABI 2007, (2007), pp. 289–301.

[3] P. Mihai. Genome assembly reborn: recent computational challenges. Briefings In Bioinformatics, Issue 10, (2009), Vol. 4, pp. 354–366.

[4] E.E.Schadt, S. Turner, A. Kasarskis. A window into third-generation sequencing. Oxford Journals, Human Molecular Genetics Issue 2, (2010), Vol. 19, pp. 227–240

[5] J. Simpson et al. ABySS: a parallel assembler for short read sequence data. CSH PRESS, Genome research Issue 6, (2009), vol. 19, pp. 1117–1123.

Visual Codebook Generation by Self-Organizing Map

Juraj RABČAN*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
rabcanj@gmail.com

Abstract. Visual codebook generation is important task of Bag of Visual Words. This generation can be done by clustering descriptors of local features. This is challenging task, because we do not know how many visual words is in image dataset. State of the art for clustering descriptors of local features is k-means algorithm. The drawback of codebook generation by k-means algorithm is that cluster centers are located in high density areas. In this paper we investigate Kohonen's algorithm and we compare it with k-means. Kohonen's algorithm is similar to k-means in his goals, but has an algorithmically different approach. Comparison is done by evaluating of searching images task on UK-bench dataset.

1 Introduction

Systems for searching and indexing text documents are nowadays considered as mature and so effective, that they can operate with millions files at once. Common representation for documents in this field is Bag Of Words. Text documents are represented by a set of words and phrases occurring in the document. This set can be stored as feature vectors, which represents documents by their words count (histogram of word occurrences in document) [3]. Since the occurrence of a given word tends to be sparse across different documents, we can create query based on keywords, which produces relevant content in real time.

In computer vision, similar representation of images is named Bag Of Visual Words. This representation found inspiration in classical Bag Of Words. Words form text documents are replaced by visual words. Visual words are analogy to words from text documents. Visual words represent small parts of an image which carry some kind of information. This representation is popular in field of image retrieval systems and in classification tasks.

Bag of Visual Words rely on a visual codebook. Visual codebook is vocabulary where all visual words from documents are stored. Visual codebook is generated by clustering descriptors of local features. Cluster centers are then considered to be visual words. Codebook generation is one of the

* Master study programme in field: Information Systems
Supervisor: Dr. Gabriela Kosková, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

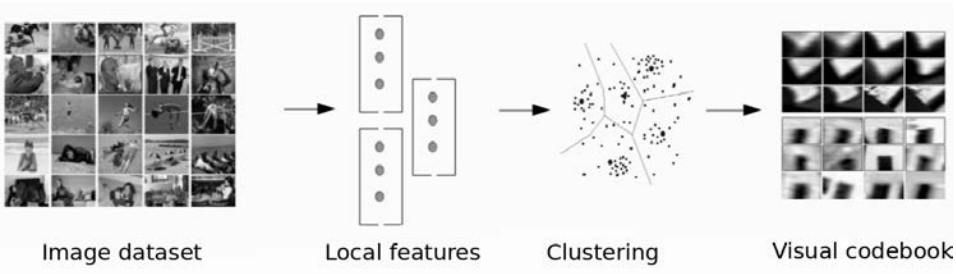


Figure 1. Algorithm of visual codebook generation.

most important part of the Bag of Visual Words, because good creation of codebook can increase within-class and between-class discriminative power of system based on Bag of Visual Words.

2 Codebook of Visual Words

Visual codebook represents vocabulary for all visual words which occur in images. Visual words can be seen as analogy to words from text documents and these words can be derived by clustering of local features descriptors. For searching these features and their descriptors we have used SIFT (Scale-invariant-feature-transform) algorithm. This algorithm is popular for this task, because it is invariant to scale and rotation and it is partially invariant to change in illumination and 3D camera viewpoint too.

When all features are detected we can create visual words by clustering their descriptors. Approximately 90000 samples of SIFT features selected from whole dataset have been used as an input for Self-organizing map and k-menas algorithm. Our experiment shows that number of clusters is important for Self-organizing map and for k-means too. It is not easy to say how many clusters we need, because we do not know how many visual words occur in the dataset. Clustering must be done several times with different number of clusters to discovery of a suitable amount of clusters. The algorithm of visual codebook generation is illustrated in Figure 1.

Visual codebook is important part of Bag of Visual words representation. This representation can be used for searching images [2] and it is briefly described in the next section.

3 Bag of Visual Words

This approach found inspiration in natural language processing domain. In Bag of Visual Words image is represented as histogram of visual words from the image. This representation is popular in image retrieval systems and in object recognition tasks. Bag of Visual Words can be briefly described as follows:

1. Create SIFT descriptors for images from the whole dataset.

2. For experiments presented in this paper PCA has been used. Dimensionality of SIFT descriptors was decreased to 36.

3. Visual codebook generation – *described in Section 2*.

4. Create histogram of visual words occurrence for images from the whole dataset.

5. Find the most similar images by euclidean distance between histograms of visual words.

4 Clustering of Local Features

The most common algorithm for clustering of local features is k-means [9]. For codebook generation was k-means used in many works [1, 8, 12]. The drawback of codebook generation by k-means algorithm can be that cluster centers are located in high density areas. Dute to this problem, there were introduced Bag of Visual Words models with different clustering algorithms i.e. mean-shift algorithm [4] or hierarchical k-means [10]. In this paper we evaluate Bag Of Visual Words model in image retrieval task. We try several Self-organizing maps as alternative to k-means and we evaluate how this algorithm impact on the image retrieval system. There are a lot of approaches for training of Self-organizing map. We investigate Kohonen's algorithm and we compare it with k-means. Self-organizing map should converges slower than k-means but it should be far more robust than k-means.

4.1 K-means

K-means algorithm [7] has become a frequently used algorithms in the process of creation of visual codebooks. It is also known as Lloyd's algorithm, according to the author of this algorithm. The aim of the algorithm is to split n instances into clusters, where each instance is assigned to the nearest centroid.

The algorithm converges to a local minimum, so it may not find the optimal solution. The algorithm is sensitive to initialization, so it is necessary to run it several times. The steps of the algorithm are described briefly as follows:

1. From the set of clustering objects randomly choose k instances. These randomly selected instances represent centroids.

2. All objects x_i are assigned to the clusters. The instance x_i is added into the cluster, which has the nearest centroid c_k. Thus, x_i is assigned to the cluster C_j, where:

$$j = arg \min_{k} \ d(x_i, c_k) \tag{1}$$

3. New centers c_j of clusters are calculated using the mean value, while minimizing the distance between the centroid of the cluster and the objects belonging to the cluster:

$$c_j = \frac{1}{N_j} \sum_{i=1}^{N_j} x_i, \forall x_i \in \ C_j \tag{2}$$

   N_j is number of instances which belong to cluster C_j.

4. Repeat step 2 and 3 until the stop condition is false.

4.2 Self-organizing maps

Self-organizing map is type of a neural network which can be used for data clustering. These maps are widely used in many areas i.e. collaborative filtering [11] or in gene analyses [14]. The goal of self-organizing map is to find a set of centroids and assign each object from the dataset to the closest centroid. In neural network terminology neurons are analogous to centroids.

Self-organizing map consists of two layers. An input layer projects vectors from training dataset to the grid of neurons. Similar input vectors cause a response on neurons which are physically near in the grid [6].

Widely used algorithm for training of Self-organizing map is Kohonen's algorithm [5]. The steps of the algorithm are described briefly as follows:

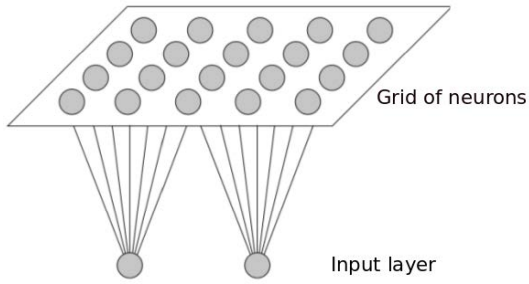


Figure 2. Example of Self-organizing map. Taken from [13].

1. Randomly initialize neuron's weight vectors w.

2. Randomly select an input vector x, and find the winner neuron i^* for selected input vector:

$$i^* = arg \min_i d_E(x, w_i) \tag{3}$$

   In Equation 3 d_E is distance measure function.

3. Update weight w_i vector of winner neuron and his neighbors:

$$w_i(t+1) = w_i(t) + \alpha(t).h(i^*, i).[x(t) - w_i(t)] \tag{4}$$

   In Equation 4 t is a parameter identifying the epoch. Parameter α is learning rate and $h(i^*, i)$ is neighborhood function, we can calculate it by:

$$h(i^*, i) = exp(\frac{d_E^2(i^*, i)}{\gamma^2(t)}) \tag{5}$$

   Learning rate α and size of neighborhood γ is gradually decreased.

4. Repeat step 2 and 3 until the stop condition is false.

5 Experiment

For experiments presented in this paper UK-Bench dataset [10] created at the University of Kentucky has been used. We use first 1000 images from this dataset. Dataset consists of photos of different objects. Each object has four photos taken from different angle. Our measure of performance is to count how many of the first four resulted images belong to a query image from that set of four images. Score for query is done by equation:

$$score = \sum_{i=0}^{3} x(i), x(i) = \begin{cases} 1 & \text{if query image and resulted image[i] are photos of the same object} \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

Three different sizes of codebook have been tested in our experiments: 800, 1000 and 1800. Each codebook was generated ten times with the same parameters. For comparison of algorithms were selected all codebooks. The Table 1 shows average results of testing algorithms and Figure 3 and Figure 4 show data from a Table 1.

Table 1. Results comparison. Learning rate parameter for all configurations of SOM was set to 0.005. Score was calculated according to the proposed Equation 6.

Number of clusters	score: SOM (topology) / Standard deviation	score: K-means / Standard deviation
800	2.54 (20x40) / 0.054	2.47 / 0.037
1000	2.64 (20x50) / 0.044	2.52 / 0.031
1800	2.45 (30*60) / 0.038	2.43 / 0.046

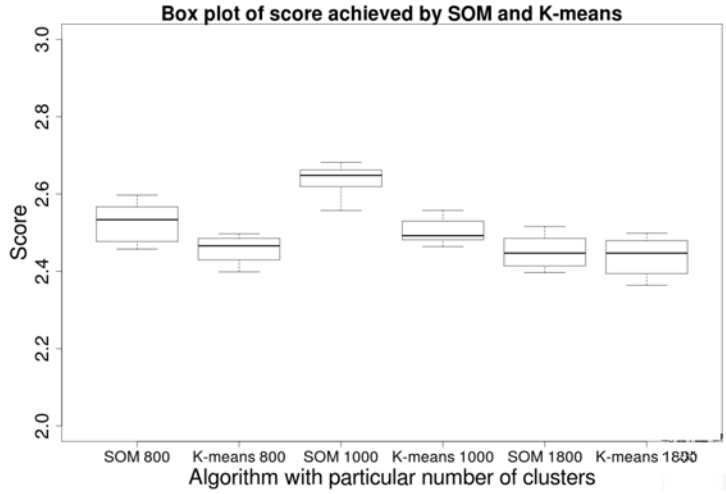


Figure 3. Box plot shows results after ten codebook generations for every size.

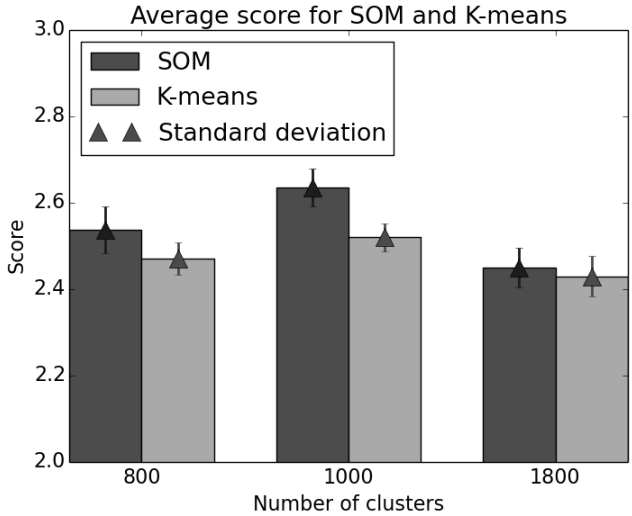


Figure 4. Plot shows average score after ten codebook generations for every size.

6 Conclusion

Two methods for visual codebook generation were presented and experimentally evaluated at image retrieval system. The best results were obtained with self-organizing map. The advantage of k-means method is that it only needs to know number of clusters in advance, whereas SOM needs to know topology, learning rate and number of iterations. On the other hand SOM gives better results in image retrieval system than k-meas. The advantage of SOM is that if there new images arrive we can run some training iterations with new images, whereas k-means must be trained on whole dataset again.

The best results were achieved with 1000 words in visual codebook for k-means and self-organizing map too. The time needed to codebook generation by K-means algorithm was much smaller in compare to time needed to generation by Self-organizing map. Reason of this difference can be, that Self-organizing map is our own implementation whereas K-means is a library algorithm.

References

[1] Benesova, W., Kottman, M., Sidla, O.: Hazardous sign detection for safety applications in traffic monitoring. In: *Proc. SPIE*. Volume 8301., 2012, pp. 830106–830109.

[2] Bouachir, W., Kardouchi, M., Belacel, N.: Improving Bag of Visual Words Image Retrieval: A Fuzzy Weighting Scheme for Efficient Indexation. In: *Signal-Image Technology Internet-Based Systems (SITIS), 2009 Fifth International Conference on*, 2009, pp. 215–220.

[3] Grauman, K., Leibe, B.: *Visual Object Recognition*. Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan & Claypool Publishers, 2011.

[4] Jurie, F., Triggs, B.: Creating Efficient Codebooks for Visual Recognition. In: *Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1 – Volume 01*. ICCV '05, Washington, DC, USA, IEEE Computer Society, 2005, pp. 604–610.

[5] Kohonen, T.: An Adaptive Associative Memory Principle. *IEEE Trans. Comput.*, 1974, vol. 23, no. 4, pp. 444–445.

[6] Kvasnička, V., Pospíchal, J., et al.: *Úvod do teórie neurónovỳch sietí*. Iris, 1997.

[7] Le Cam, Lucien M. Marie , et. all: Proceedings of the sixth Berkeley Symposium on mathematical statistics and probability. : probability theory, Le Cam, Lucien M. Marie , et. all, Berkeley, Ca. Los Angeles, Ca. University of California Press.

[8] Lowe, D.G.: Object Recognition from Local Scale-Invariant Features. In: *Proceedings of the International Conference on Computer Vision-Volume 2 – Volume 2*. ICCV '99, 1999.

[9] Moosmann, F., Triggs, B., Jurie, F.: Fast Discriminative Visual Codebooks using Randomized Clustering Forests. In Schölkopf, B., Platt, J., Hoffman, T., eds.: *Advances in Neural Information Processing Systems 19*. MIT Press, Cambridge, MA, 2006, pp. 985–992.

[10] Nistér, D., Stewénius, H.: Scalable Recognition with a Vocabulary Tree. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Volume 2., 2006, pp. 2161–2168.

[11] Polčicová, G., Tiňo, P.: Making sense of sparse rating data in collaborative filtering via topographic organization of user preference patterns. In: *Neural Networks*, Elsevier, 2004, pp. 1183–1199.

[12] Rashid, M., Abu-Bakar, S.A.R., Mokji, M.M.: Human emotion recognition from videos using spatio-temporal and audio features. *The Visual Computer*, 2013, vol. 29, no. 12, pp. 1269–1275.

[13] Sinčák, P., Andrejková, G.: *Neurónové siete. Inžiniersky prístup: Dopredné siete*. Elfa, 1996.

[14] Törönen, P., Kolehmainen, M., Wong, G., Castrén, E.: Analysis of gene expression data using self-organizing maps. {*FEBS*} *Letters*, 1999, vol. 451, no. 2, pp. 142–146.

Partial Digest Problem: A Simplified Heuristic Approach

Daniel Soós*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
soos.dano@gmail.com

Abstract. This paper deals with a heuristic proposal for a DNA mapping problem called partial digest (PDP). In information theory, this problem is also known as the turnpike problem. A backtracking algorithm solves most of its examples in reasonable time, but a class of instances is known, where it takes exponential time for the algorithm to solve the problem. A genetic algorithm (GA) was constructed to overcome this issue. Our work contributes to the PDP state-of-the-art with a simplified heuristic approach that clearly outperforms the GA. However, we examine the backtracking algorithm as well, and based on our results, it is clear that these heuristics are still not so efficient on the given problem in comparison with the backtracking. A critical review of the paper that presents the GA is also given.

1 Introduction

Information retrieved from molecular biology with the increasing number of powerful tools in the hands of computer scientists makes the topic of DNA processing more relevant than ever. Thanks to the progress in DNA sequencing, mapping of the DNA was pushed more into the background in the last two decades, but digesting the molecule by enzymes still remains a commonly used technique. Also, PDP and the double digest problem (DDP) is an interesting algorithmic task [5]. PDP can be found in literature also as the turnpike problem [2], and unlike DDP [1], it is still not proved whether it belongs to the class of NP-problems, or not.

A physical map of the DNA is a map that constitutes of actual distances between fragments measured in base pairs. Restriction enzymes cut the molecule at some points as they create restriction sites. PDP is considered as a restriction mapping problem, and it is an application of only one enzyme on the molecule, but the experiment takes place many times on the DNA copy. Various restriction sites will be recognized by giving various periods of time to the enzyme to act on the DNA. Even if we disregard the errors that occur in this experiment, the task remains computationally very complex.

---

* Master study programme in field: Ĩnformation Systems
Supervisor: Assoc. Professor Mária Lucká, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 103–108.

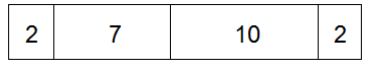


Figure 1. Partial digest problem: a restriction map when only one enzyme is applied.

Ideally, biologists give the input of PDP in a form of all the pairwise distances that construct the map. Thus, the input is a multiset of all the distances between any two sites, and its size, explicitly. The multiset is given as ordered because of a process called gel electrophoresis, which recognizes the fragments by their length. Now, some notation is given for the problem based on [5].

In Figure 1, a restriction map to a simple PDP is shown, i.e., the solution can be immediately derived from the illustration: it is the set of the location of all restriction sites $\{0, 2, 9, 19, 21\}$. Let us assume that the fragment sizes are meant to be constructed from a thousands of base pairs. An input for such problem results in the following fragment lengths: *2*, *9*, *19*, *21*, then *7*, *17*, *19* , and *10*, *12*, and finally, *2*. The first four numbers define the fragments from the left endpoint to all other restriction sites, the next three are the fragments between the right border of the first restriction site and the others to its right, etc.

If X is a set of n points in increasing order and determines the location of restriction sites, then ΔX denotes the multiset of all $\binom{n}{2}$ pairwise distances between points in X. Thus, the instance in Figure 1, has $X = \{0, 2, 9, 19, 21\}$ and $\Delta X = \{2, 2, 7, 9, 10, 12, 17, 19, 19, 21\}$, where $n = 5$. In PDP, multiset L is always given and the optimal solution is such X,that the constructed $\Delta X = L$.

Section 3 gives a detailed insight into our heuristic approach. Experimental results with explanation are shown in the following section. The last section discusses the obtained results and observations. Also, further work is considered in this part of the work, by taking into account the current state of either PDP, and DDP, while the latter being our main research interest in the future.

2 Related work

Skiena's backtracking algorithm [6] is a very efficient method designed for PDP. The algorithm solves "random" instances in polynomial time, and is being considered as the benchmark algorithm for PDP. In this method, firstly there is the removal of the largest element in the multiset and along with 0, which is always the first location in the restriction map, it forms the initial X. That means we already know x_1 and x_n, where n is the cardinality of the set. Now, without the loss of generality we can determine either x_2 or x_{n-1}, because the two cases are symmetric. Whether we calculate with x_2 or x_{n-1} (the selected value is noted as y from now on), we get a set of two values as it is the distance between each occurence in X, which we know that initially consists of x_1 and x_n. If that set is also a subset of L, then we add element y to X and remove this subset from L.

Then the same action is being repeated, thus we grab the largest remaining element in L as well as its complement (subtracting the value from the original width) and generate a possible subset of L (the distances between y and X). If the subset is not in L, we do not explore this branch. Otherwise, we do, and if L is empty, an optimal solution X was found.

In theory, this backtracking algorithm should have at the worst case a recursion tree of depth n with the degree of 2, so the recursive function is called $O(2^n)$ times. Then, new distances have to be computed along with a binary search to check whether these new distances are in L, too. That makes the worst case performance $O(2^n n \log(n))$. However, mostly only one branch is viable, and that makes the algorithm very effective. Zhang's class of instances [8], though, show that the worst case running time for this algorithm indeed might occur. These instances also have the characteristic of having only one unique solution (discounting the mirror image of it). For our experimental tests, we created several instances of this class as there is no reason to confront metaheuristic approaches on average instances with the backtracking algorithm, because the latter will always be superior in these cases.

In order to overcome this drawback of Skiena's algorithm, a GA solution was proposed by Ganjtabesh et al. in [3]. GA is a metaheuristic that behaves as a metaphor to the natural selection process. Its main purpose is to solve search problems and optimization tasks, while selection, crossover, and mutation being the mainly used operations that are responsible for faster convergence. Ganjtabesh et al. in their paper use the roulette wheel as a selection scheme, along with the simple one-point crossover. Mutation can be considered only as a minor feature in this solution as it only mutates one individual with a very low probability at exactly one element of the solution.

Here, we must remark some misconceptions or incorrectnesses that we found in the cited paper that presents the GA for PDP. Firstly, the work constantly refers to the cost function as fitness function, when in fact it is indeed a cost function, because a valid solution equals 0, thus the objective function is to be minimized. Furthermore, in theory, GA should reckon with maximization, i.e., with an actual fitness function. Arguably, an even worse statement is their mentioning of the Hamming distance; originally invented by Hamming [4]. In their GA, they calculate it as $\sum_{i=1}^{k} |L_i - (\Delta X_i)|$, where k is the size of the multiset containing the pairwise distances, as L is the given multiset to PDP, and ΔX is the calculated multiset from a current solution. Now, based on the actual Hamming distance, this is wrong: a Hamming distance value returns the number of different symbols for the corresponding indices for objects of equal length. Other details that we derived and improved from this method are discussed in the next section. The results of this GA are shown in Section 4; these were obtained after our implementation of the GA for PDP, and it is important to note that they are not coincident with the results published in [3] by Ganjtabesh et al.

3 Our algorithm

As stated above, the input of PDP is the multiset of pairwise distances L with the size of $\binom{n}{2}$. The output is set X of n elements, such that $\Delta X = L$. Now, the search space can be reduced as follows [3]. If we have a solution $X = \{x_1, x_2, ..., x_n\}$, where $x_1 = 0$ and $x_n = width$, and $width$ is the maximum value in multiset $L = \{d_1, d_2, ..., width\}$, then $x_i \in L$ and also, $width - x_i \in L$. Because we know that the two endpoints on the restriction map must be 0 and $width$, respectively, this assumption is true in any case. Now, the search space simply consists of those elements, where this predicate is true. That is, the set of the candidate solutions can be described as $\Delta = \{\delta_i : \delta_i = (d_{i_1}, d_{i_2}), d_{i_1} \in L, d_{i_2} \in L, d_{i_1} + d_{i_2} = width, 1 \leq i_1 \leq i_2 < k\}$. As Δ contains ordered pairs, it is assumed that $d_{i_1} \leq d_{i_2}$ for each i.

Ganjtabesh et al. state that Δ must contain at most two equal elements. While this is a valid interpretation of the input, we claim that in the case of Zhang's instances, it would be a needless representation as the properties of these instances imply that any element from the candidate solution will occur in L more than once as well as its complement. Therefore, if we dismiss one element from the ordered pair, and only construct the search space from the distinct elements subtracted from them, the search space is even more reduced. Although the paper does not state it explicitly, it is assumed that duplicate elements are not allowed in a solution, otherwise the results would not be too efficient, so this reduction should not play a big role in practice.

We point out that the crossover operation in the GA for PDP checks whether an element that will be potentially swapped was already contained in the subset, or not; and this is another time consuming operation. Especially in Zhang's class, where the majority of Δ is involved in one candidate solution (e.g., an instance where $|X| = 52$, there will be only 10 elements missing from X that are in Δ). The roulette wheel selection strikes down as another less impressive decision, mainly because of the abovementioned cost function. It results in reversing the cost values of each solution, otherwise a selection would end up in a counterproductive output. In our implementation of their GA, we normalized these values to a selected range, and then reverted it, but that only adds another $O(n^2)$ time complexity to the algorithm. We claim that if a cost function seems to be the best decision, but one looks for a selection scheme, a tournament would be a better choice, and would probably give similar results (and surely, in definitely less time).

It should be clear by now why do we not consider crossover, nor roulette wheel operation as considerable improvements in our algorithm. Instead, our method contains a more robust mutation operation that should mean a faster convergence towards the optimum. Also, an actual Hamming distance of ΔX from L is considered, which behaves more transparently as we show below. Of course, an optimum is still found when the cost function equals 0. The skeleton of the heuristic is shown in a form of a pseudocode in Algorithm 1. Here, obviously the input is the multiset L with all

Algorithm 1 Our heuristic for PDP

```
1: function PDP_HEURISTIC(L, p_mut, popSize, t, t_max)
2:     Create search space from L and store into Δ
3:     Initiate popSize solutions as random XSize permutations of Δ
4:     Calculate the cost values of candidate solutions
5:     while t < t_max do
6:         t = t + 1
7:         for all population do
8:             Mutate each element in a candidate solution with probability of p_mut
9:             if t reaches threshold value then
10:                Apply partially deterministic local mutation on best solutions
11:            end if
12:        end for
13:    end while
14:    return
15: end function
```

the pairwise distances. The size of the population is denoted as $popSize$, while the maximal number of iterations is t_{max}. Similar to the GA, the algorithm starts with the construction of the search space, which we already discussed above. Random solutions are given to all the candidate solutions in the population. This is obtained by a random permutation of Δ, but only $XSize - 2$ elements are taken into consideration each iteration, because both endpoints are already clear in any candidate solution. Then, we iterate through the generations and the only operation which we maintain from the GA is the mutation, although we define two forms of it. Algorithm 2 shows how the basic mutation works. The probability of an element being mutated is p_{mut}, and the given solution to the operation is given also. A random number is generated for each element in the solution, and if p_{mut} is bigger, then we swap this element with a randomly generated that is not contained in the solution, but is in Δ. We might accept the solution if it has a lower cost value than the original, but this is not always the case. As Sur-Kolay et al. state in their GA solution for DDP [7], the cost function is far from ideal, but it is hard to construct a better one, given the nature of these problems. Thus, we invoke an acceptance threshold value, so if the cost value is too low, we reject the newly found solution. It is because of the features of Zhang's instances, that if the Hamming distance is close to the optimum (i.e., to 0), it actually means that the two mirror image solutions are more "mixed up". Our experiments later showed that indeed this application of a threshold value evolves into faster convergence.

Local search is only performed after a large portion of generations, and its task is to explore the potentially near-optimal areas more widely, although it adds to the time complexity quite strongly. For a given smaller range (starting at a random point in the current solution), similarly to the basic mutation, a random element is chosen from Δ and this area is explored. However, to keep the number of instructions in the algorithm low, this local search does not update any solutions, unless it is the optimum. Any time the optimal solution is found, the algorithm returns it, otherwise no "best solution" is returned because of the mentioned imperfection of the cost function. It can be said that this heuristic is closely related to both hill climbing and local search optimization methods.

Algorithm 2 Mutation_PDP

```
1: function MUTATION(solution, p_mut, L)
2:     for all solution_{X_i} do
3:         random = random< 0, 1 >
4:         if random < p_mut then
5:             Store random element from Δ that is missing in solution to randElement
6:             Swap X_i element with randElement and calculate cost value
7:             if Optimal solution was found then
8:                 return Optimal solution
9:             end if
10:            if newCost < cost(solution) & newCost > acceptance threshold then
11:                Update the solution and its cost
12:            else
13:                Reject new solution
14:            end if
15:        end if
16:    end for
17: end function
```

4 Experimental results

Our comparison consisted of testing our heuristic with the GA solution for PDP, and also with Skiena's backtracking. While we implemented the latter in its original form, we already discussed the changes that needed to be made in order to make the GA comparative with the other two. Also, as Ganjtabesh et al. state, $p_{mut} = 0.05$ and $p_{cross} = 0.75$, while the $popSize = 100$, and $t_{max} = |\Delta|$. In our heuristic, we kept these values for both t_{max} and $popSize$. However, our algorithm has $p_{mut} = 0.02$, but there is a difference in its application. We iterate through all the elements for all solutions with the given probability, whilst their GA mutates exactly one element with a given probability for each solution. The local search is applied after half of the maximal generation number is reached, and the acceptance threshold value is at $\frac{|L|}{40}$, thus if the cost value is below that, the solution is discarded. Unlike the GA, our algorithm uses the real Hamming distance as a cost function.

We constructed instances of Zhang's class with various sizes. Table 1 shows the experimental results we obtained in our tests. It is clear that on these sizes of Zhang's instances, the backtracking

Table 1. Comparative results between Skiena's bactracking, Ganjtabesh's GA and our heuristic for Zhang's instances.

time if optimum found (or the whole running time, if not)				success rate in %		
Size of X	Skiena	Ganjtabesh	Our algorithm	Skiena	Ganjtabesh	Our algorithm
12	$0.013s$	$0.032s$	0.031s	100	100	100
32	$0.031s$	$0.500s$	0.375s	100	50	85
52	$1.000s$	$14.647s$	14.751s	100	10	50
57	$6.016s$	$22.625s$	7.803s	100	0	100
62	$32.181s$	$31.530s$	15.165s	100	0	18

algorithm still manages to provide the best results. In fact, our results prove that the GA constructed for PDP in [3] has incorrect measurements. The results also show that we should look for further improvement as we could not match Skiena's outcomes. However, our solution clearly outperforms the GA. We do not have enough stats about how near the unsuccessful test were to the optimum in terms of the number of elements missing from X, but we assume that when $|X| = 52$, our experimental outputs were probably very close to the global optimum. Especially when we consider

that a bigger instance was solved optimally by our heuristic every time. We plan to cover this issue also, but more importantly, to improve the results, too. When $|X| = 77$, even Skiena's backtracking did not found the solution in our tests because of the algorithm's exponential time complexity, so some further improvement on our results would be a considerable progress.

All the implementation was programmed in language C and executed on an Intel Core i5-4210U at 1.70 Ghz. Both heuristics were tested at least 30 times.

5 Conclusion

This paper presented a simplified heuristic to the DNA mapping problem called PDP. We evaluated our method in comparison with Skiena's backtracking algorithm which can be considered as the benchmark solution for PDP. Although our results could not match it, we also reviewed the GA proposed for PDP [3], and our algorithm not only evidently exceeded the GA, but it also became clear that the results published in the cited paper are incorrect.

It is important to note that Dakić [2] handled Zhang's instances with a quadratic program in polynomial time, thus specifically PDP does not need to be optimized currently as it is presumable that the problem does not belong to set of NP problems. Another reason that it is not solved frequently is that real instances always contain erroneous inputs, and these are usually solvable in polynomial time by the backtracking algorithm.

However, DDP belongs to the class of NP-complete problems, and as the concept is similar, furthermore a GA solution for it already exists in [7], our main research interest in the future will consist of solving the DDP and possibly improving the results of GA, again.

Acknowledgement: This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG 1/0752/14.

References

[1] Cieliebak, M., Eidenbenz, S., Woeginger, G.J.: Double Digest Revisited: Complexity and Approximability in the Presence of Noisy Data. In Warnow, T., Zhu, B., eds.: *Computing and Combinatorics*. Volume 2697 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2003, pp. 519–527.

[2] Dakić, T.: *On the Turnpike Problem.* PhD thesis, Simon Fraser University, 2000.

[3] Ganjtabesh, M., Ahrabian, H., Nowzari-Dalini, A., Razaghi-Moghadam-Kashani, Z.: Genetic Algorithm Solution for Partial Digest Problem. *Int. J. Bioinformatics Res. Appl.*, 2013, vol. 9, no. 6, pp. 584–594.

[4] Hamming, R.: Error Detecting and Error Correcting Codes. *Bell System Technical Journal*, 1950, vol. 26, no. 2, pp. 147–160.

[5] Jones, N., Pevzner, P.: *An Introduction to Bioinformatics Algorithms*. A Bradford book. MIT Press, 2004.

[6] Skiena, S., Sundaram, G.: A partial digest approach to restriction site mapping. *Bulletin of Mathematical Biology*, 1994, vol. 56, no. 2, pp. 275–294.

[7] Sur-Kolay, S., Banerjee, S., Mukhopadhyaya, S., Murthy, C.A.: The Double Digest Problem: Finding All Solutions. *Int. J. Bioinformatics Res. Appl.*, 2009, vol. 5, no. 5, pp. 570–592.

[8] Zhang, Z.: An Exponential Example for a Partial Digest Mapping Algorithm. *Journal of Computational Biology*, 1994, vol. 1, no. 3, pp. 235–243.

Processing and Comparing of Eye-Tracking Data Using Machine Learning

Miroslav ŠIMEK*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
mirzekdt@gmail.com

Abstract[1]

Eye-tracking technology is getting more and more attention when evaluating adaptive systems. Current approaches rely upon manually defined areas of interest and optionally also manually designed features fed as an input to some machine learning algorithm. We present a generic way how to learn an efficient abstraction of a user session, which is especially suitable for comparing, clustering and categorizing user sessions.

Our approach is based on unsupervised machine learning, specifically Restricted Boltzmann Machine (RBM), which is presented with fragments of user sessions in form of heat maps capturing spatial (pixel coordinates) and time (pixel intensity) information. RBM is able to find its own features to make efficient session abstraction in the context of other user sessions, which is especially suitable for comparing, clustering and categorizing user eye-tracking sessions.

We developed a set of applications that preprocess the eye-tracking data, train the RBM, acquire data abstractions from trained RBM and process them into most common patterns over time for each individual user session. Subsequently we are able to cluster sessions into groups and find those, where users did not behave as expected.

Our goal is to improve possibilities of automatized evaluation of eye-tracking data sequences by comparing user sessions to each other, comparing user sessions with expected usage of application captured by eye tracker or detecting outliers for closer manual inspection.

Acknowledgement: This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG 1/0752/14.

---

* Master degree study programme in field: Information Systems
Supervisor: Dr. Michal Barla, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

[1] Full paper available in printed proceedings, pages 109–114.

Application of Genetic Algorithm on Model-Based Cluster Analysis

Peter LAURINEC*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
peter.laurinec@stuba.sk

Abstract. This work deals with model-based Gaussian clustering. We define multivariate normal distribution, describe its geometric interpretation and derive optimization criteria for clustering under various conditions on the covariance matrices of clusters. We propose a new optimization criterion for creation of clusters and solve the optimization problem using a generalization of the binary genetic algorithm. We compare the achieved results with the real classification into clusters and classification made by EM-algorithm, which was implemented in the software R. The analyzed data come from botany, astronomy, medicine and other sciences.

1 Introduction

Modern multivariate statistical methods are increasingly used nowadays. The reasons are rapid progress of computer technologies and increasing demands on obtaining more and more information from big data. Scientific teams and commercial companies process the big data and try to get useful information. One possibility how to reveal structured information in the data is clustering of "related" objects into groups. Herewith deals cluster analysis, which will be discussed in this paper. We will describe the analysis of clusters based on probabilistic models. More specifically, the normal model, i.e. we assume that observations come from multivariate normal distribution. This is a big advantage against centroid-based methods of clustering (K-means, K-medoids), because the model-based methods can find clusters of elliptic shapes, smaller clusters embedded in larger ones and partially recognize also overlapping clusters.

The determination of the correct grouping of objects into clusters is quite difficult optimization problem. We will solve it using genetic algorithms (GA), that are widely used in for clustering. Genetic algorithms are a class of stochastic optimization algorithms based on principles of genetic recombination and natural selection. In papers [8, 12, 13] the authors use genetic algorithms for to overcome limitations of K-means or fuzzy c-means algorithm, because GA can find better local optima in a centroid-based clustering situation. Genetic algorithm could be used too in identification

* Doctoral study programme in field: Software Engineering
Supervisor: Assoc. Professor Mária Lucká, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 115–122.

of optimal number of clusters. This approach was used to improve initialization of centroids in the K-means [11]. We will compare our GA with the EM-algorithm, which is searching maximal likelihood estimation of the parameters in a statistical model.

In the next section we will describe model-based normal cluster analysis methodology, then we introduce our GA and the used methods. In the last section – Data analysis and experiments – we evaluate and compare effectiveness and correctness of clustering on four data sets from real world.

2 Cluster Analysis Methodology

2.1 Multivariate normal distribution

If we want to work with model-based normal cluster analysis, we must define multivariate normal distribution [7]. Consider a random vector $\boldsymbol{X} = (X_1, X_2, \ldots, X_p)^T$, a given vector $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_p)^T$ and a positive symmetric definite matrix of type $p \times p$ $\boldsymbol{\Sigma}$. We say, that p-dimensional random vector $\boldsymbol{X}$ has a regular p-dimensional normal distribution with parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, if $\boldsymbol{X}$ is absolute continuous with density

$$f(\boldsymbol{x}) = f(x_1, \ldots, x_p) = \frac{1}{(2\pi)^{\frac{p}{2}} \det(\boldsymbol{\Sigma})^{\frac{1}{2}}} \exp\Big\{ -\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) \Big\},$$

where $\boldsymbol{\mu}$ is a vector of expected values and $\boldsymbol{\Sigma}$ is a covariance matrix of vector $\boldsymbol{X}$. This fact is denoted as $\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

Now, with help of matrix calculus, we can interpret multivariate normal distribution geometrically. This help us better imagine and understand model-based normal cluster analysis. Consider that $f(\boldsymbol{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma})$ is the density of a random vector $\boldsymbol{x}$ with distribution $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. If the level set is non-empty and is defined with the following expression:

$$L_c(f) = \{\boldsymbol{x} \in \mathbb{R}^p : f(\boldsymbol{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) \geq c\},$$

then this set is an ellipsoid. The direction of principal axes of ellipsoids are given by eigenvalues of vectors of covariance matrix $\boldsymbol{\Sigma}$. Relative length of major axis is given by roots of eigenvalues of the matrix $\boldsymbol{\Sigma}$.

If $\boldsymbol{\Sigma} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^T = \boldsymbol{U}\boldsymbol{\Lambda}^{\frac{1}{2}}(\boldsymbol{U}\boldsymbol{\Lambda}^{\frac{1}{2}})^T$ is a decomposition on the basis of eigenvalues, where $\boldsymbol{U}$ is a orthogonal matrix of eigenvectors and $\boldsymbol{\Lambda}$ is diagonal matrix of eigenvalues, then we can write

$$\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \Leftrightarrow \boldsymbol{X} \sim \boldsymbol{\mu} + \boldsymbol{U}\boldsymbol{\Lambda}^{\frac{1}{2}} N_p(0, \boldsymbol{I}) \Leftrightarrow \boldsymbol{X} \sim \boldsymbol{\mu} + \boldsymbol{U} N_p(0, \boldsymbol{\Lambda}),$$

where $\boldsymbol{U}$ we can name as a matrix of rotation. So distribution $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ can be understood like "sferic" distribution $N_p(0, \boldsymbol{I})$, which is scaled by $\boldsymbol{\Lambda}^{\frac{1}{2}}$, rotated by matrix $\boldsymbol{U}$ and shifted by $\boldsymbol{\mu}$.

2.1.1 Model-based normal cluster analysis

Generally, cluster analysis is the task of grouping a set of objects (observations) in such a way that objects in the same group (called a cluster) are more similar (in some sense or another) to each other than to those in other groups (clusters). At first we define denotations, which we will use in cluster analysis.

Let n is a number of objects and c denotes the number of clusters. Let $\Gamma_{n,c}$ is the set of vectors of length n, those components belongs to set $\{1, \ldots, c\}$ and each value $1, \ldots, c$ contains at least once. By classification of n objects to c clusters we denote each vector γ from the set $\Gamma_{n,c}$. For $j \in \{1, \ldots, c\}$ and $i \in \{1, \ldots, n\}$ is equation interpreted $\gamma_i = j$ as classification γ, which assigns i-th observation to j-th cluster. Set $C_j(\gamma) = \{i \in \{1, \ldots, n\} : \gamma_i = j\}$ is denoted as j-th cluster for classification γ. The number of objects in a set $C_j(\gamma)$ is denoted as $n_j(\gamma)$. Optimal

Table 1. Summarization of optimization criteria.

#	Optimization criterion	Condition on $\mathbf{\Sigma}_j$	Distribution	Orientation	Shape	Volume
1.	$\sum_{j=1}^{c} n_j(\gamma) \ln \det(\boldsymbol{S}_j(\gamma))$	none	Ellipsoidal	Variable	Variable	Variable
2.	$\sum_{j=1}^{c} n_j(\gamma) \ln \mathrm{tr}(\boldsymbol{S}_j(\gamma))$	$\lambda_j \boldsymbol{I}$	Spherical	NA	Equal	Variable
3.	$\ln \det(\boldsymbol{S}(\gamma))$	Σ	Ellipsoidal	Equal	Equal	Equal
4.	$\ln \mathrm{tr}(\boldsymbol{S}(\gamma))$	$\lambda \boldsymbol{I}$	Spherical	NA	Equal	Equal

classification to clusters will be denoted as $\hat{\gamma}$. As covariance matrix of a cluster $C_j(\gamma)$ we denote the matrix $\boldsymbol{S}_j(\gamma) = n_j^{-1}(\gamma) \sum_{i \in\, C_j(\gamma)} (\boldsymbol{x}_i - \overline{\boldsymbol{x}}_j(\gamma))(\boldsymbol{x}_i - \overline{\boldsymbol{x}}_j(\gamma))^T$. Overall covariance matrix depended on clustering is denoted as matrix $\boldsymbol{S}(\gamma) = \frac{1}{n} \sum_{j=1}^{c} n_j(\gamma) \boldsymbol{S}_j(\gamma)$. Notice, that this is the weighted average of covariance matrices $\boldsymbol{S}_j(\gamma)$.

Main assumption at model-based normal cluster analysis is that vectors of observations $\boldsymbol{x}_1$, $\boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$, which is characterizing objects, are realizations of random vectors $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots \boldsymbol{X}_n$, and these have p-dimensional normal distribution $\boldsymbol{X}_i \sim N_p(\mu_{\gamma_i}, \mathbf{\Sigma}_{\gamma_i})$. This assumption allows us to use classical known statistical methods, which is a big advantage.

Our goal is to find optimal classification to clusters $\hat{\gamma}$. We find it by method of maximum likelihood [1]. So we maximize likelihood:

$$L(\gamma, \boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_c, \mathbf{\Sigma}_1, \ldots, \mathbf{\Sigma}_c | \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = \prod_{i=1}^{n} f(\boldsymbol{x}_i, \boldsymbol{\mu}_{\gamma_i}, \mathbf{\Sigma}_{\gamma_i}).$$

Another big advantage of this method is, that we can change (vary) all features of distribution between clusters. Under a feature of distribution of cluster we understand its orientation, shape and volume. The key to this is re-parametrization of covariance matrices $\mathbf{\Sigma}_j$, in terms of the decomposition on the basis of eigenvalues [2]. So again we have decomposition $\mathbf{\Sigma}_j = \boldsymbol{U}_j \mathbf{\Lambda}_j \boldsymbol{U}_j^T$. Orientation of principal components $\mathbf{\Sigma}_j$ is given by matrix $\boldsymbol{U}_j$, while $\mathbf{\Lambda}_j$ determine volume and shape of cluster. We can overwrite the matrix $\mathbf{\Lambda}_j$ to $\mathbf{\Lambda}_j = \lambda_j \boldsymbol{A}_j$, where λ_j is first eigenvalue of $\mathbf{\Sigma}_j$, $\boldsymbol{A}_j = diag\{\alpha_{1j}, \ldots, \alpha_{pj}\}$, and $1 = \alpha_{1j} \geq \alpha_{2j} \geq \cdots \geq \alpha_{pj} > 0$. So $\boldsymbol{U}_j$ determines orientation of j-th cluster, λ_j its volume and $\boldsymbol{A}_j$ its shape.

We have derived four optimization criteria to clustering on basis of four conditions of covariance matrices $\mathbf{\Sigma}_j$. These optimization criteria are summarized in a Table 1. Each criterion is argument of the minimum.

2.1.2 Relationship with optimal design of experiments

During deriving optimization criteria for clustering, we have noticed interesting relationship with optimal design of experiments. Consider so-called Kiefer's class of optimization criteria for design of experiments [10]: $\phi_r(\boldsymbol{M}) = \left[\frac{1}{p}\mathrm{tr}(\boldsymbol{M}^r)\right]^{1/r}$, where $\boldsymbol{M}$ is a positively defined matrix of type $p \times p$ and r is an arbitrary real number. For $r = 0$ the criterion is called criterion of D-optimality, if $r = -1$, the criterion is called A-optimality and for $r = -\infty$ it is an E-optimality criterion. For these values r, we consider $\boldsymbol{M} = \boldsymbol{S^{-1}}$, then we can class the criteria as follows:

$$\phi_r(\boldsymbol{S}^{-1}) = \begin{cases} (\det(\boldsymbol{S}^{-1}))^{1/p} & \text{ak } r = 0 \\ \left[\frac{1}{p}\mathrm{tr}(\boldsymbol{S})\right]^{-1} & \text{ak } r = -1 \\ \lambda_{min}(\boldsymbol{S}^{-1}) = \frac{1}{\lambda_{max}(\boldsymbol{S})} & \text{ak } r = -\infty \end{cases}$$

[1] Harman, R.: *Mnohorozmerné štatistické analýzy.* http://www.iam.fmph.uniba.sk/ospm/Harman/vsa.pdf

Notice, that in our optimization criteria on clustering from Table 1 is in each case located determinant or trace of covariance matrix. For first two derived criteria its form:

$$\hat{\gamma} = \arg \min_{\gamma \in \Gamma^R_{n,c}} \sum_{j=1}^{c} n_j(\gamma) \ln[\dots],$$

respectively for another two criteria:

$$\hat{\gamma} = \arg \min_{\gamma \in \Gamma^R_{n,c}} \ln[\dots].$$

We prove that, if for $\dots$ we appoint different criteria $\phi_r(\boldsymbol{S}^{-1})$, then for $r = 0$ we get criteria 1. and 3. from Table 1. For $r = -1$ we get criteria 2. and 4.. In condition $r = -\infty$ we come up with "new" optimization criteria:

$$\hat{\gamma} = \arg \min_{\gamma \in \Gamma^R_{n,c}} \sum_{j=1}^{c} n_j(\gamma) \ln \lambda_{max}(\boldsymbol{S}_j(\gamma)) \tag{1}$$

and

$$\hat{\gamma} = \arg \min_{\gamma \in \Gamma^R_{n,c}} \ln \lambda_{max}(\boldsymbol{S}(\gamma)). \tag{2}$$

It means, that after these findings we could first customary type of clustering called D-optimal clustering and second type A-optimal clustering. Third derived type of clustering we called E-optimal clustering. So we could our optimization criteria on clustering overwrite to general form, which exploits criteria for optimal design of experiments:

$$\hat{\gamma} = \arg \max_{\gamma \in \Gamma^R_{n,c}} \sum_{j=1}^{c} n_j(\gamma) \ln(\phi_r(\boldsymbol{S}_j^{-1}(\gamma)))$$

or

$$\hat{\gamma} = \arg \max_{\gamma \in \Gamma^R_{n,c}} \ln(\phi_r(\boldsymbol{S}^{-1}(\gamma))).$$

Our criteria could be geometrically interpret also by concentration ellipsoid, which is exploit in design of experiments [9]. In criteria of D-optimality is minimizing volume of concentration ellipsoid. In A-optimality is minimizing average of ellipsoid axes and in E-optimality is minimized greatest half-axis of concentration ellipsoid. Thus we can say, that in our optimization criteria for clustering, we minimize scaled parameters (features) of criteria for optimal experimental design.

3 Genetic Algorithm

To solve optimization problem for model-based clustering, like we defined above, is NP-hard. We have chosen genetic algorithm to tackle this problem [6]. Genetic algorithms are modern class of heuristic optimization techniques, inspired by principles of genetic recombination and natural selection. They are based on populations composed by many individuals, which represents acceptable solutions of optimization problem. They allow to maximize fitness of population at specific selection rules. For the needs of clustering (specifically, in the case of at least three clusters) we had to develop generalization of binary genetic algorithm, because binary genetic algorithm work only with two integers. The steps of our algorithm are as follows:

1. Define chromosome, population and fitness function.

2. Generate initial population.

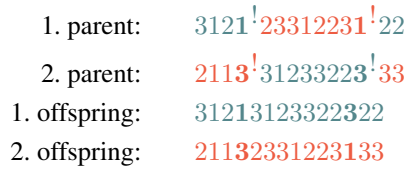


Figure 1. Double-crossover.

3. Compute fitness of each chromosome. Order population. Natural selection.
4. Select mates.
5. Mating.
6. Mutation.
7. Convergence check. If stopping rule is not satisfied , go to step 3.

Let us describe the steps and chosen methods in this algorithm. To each of our observations $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ we assign number of the cluster. Set of possible clusters is $\{1, 2, 3, \ldots, c\}$. For simplicity, we will have always given the number of clusters. Chromosome is defined as classification of observations $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ to clusters, labelling γ. Population Pop is a matrix $N_{pop} \times n$ large, where N_{pop} is number of chromosomes and n is number of observations. Fitness function in our case is a optimization criterion on clustering $\hat{\gamma}$. Fitness is a value of this criterion. The first computational step is randomly generated initial population. Next procedure is natural selection, where we select first half of chromosomes from ordered population. Population is ordered by value of fitness. Next is the selection of parents to mating. We have chosen rank-selection method, so parent with better fitness has larger probability to mating. Method for mating (recombination of chromosomes) we use double-crossover as is shown in Figure 1.

At mutation is randomly changed classification of a observation to cluster. In our algorithm we use random mutation, where is changed α observations in each chromosome. Mutation is not done, as a rule, in best ongoing chromosome, so in chromosome with best fitness. This "activity" in algorithm is called elitism. In various cases it may be preferred, if number of "elit" chromosomes is greater than 1. After mutation we created a new population, in which we must again calculate values of a fitness function. Another step is to test, whether new population is already optimal. We have chosen exceeding number of iterations as stopping rule. If stopping rule isn't fulfilled in iteration, then we continue in natural selection.

4 Data Analysis and Results

We perform cluster analysis and evaluation with help of our genetic algorithm (called *GenClust*) on real data. Cluster analysis is generally used on classification of data, for which we do not know "right" classification. In case of data sets in this paper however we know natural classification, designated by experts, and we can use it to assess the quality of clustering methods. Data analysis is performed in statistical software R.

We will compare our results with implemented function *Mclust*, which is part of library *mclust* for interface R [4]. This function mainly solve model-based normal cluster analysis. Function *Mclust* use optimization method called EM-algorithm (Expectation-maximization) [5]. EM-algorithm is iterative method for searching maximal likelihood estimation of selected parameter in statistical model. This function has 10 different models, in other words 10 different conditions on covariance

Table 2. Mclust and conditions to covariance matrices.

Conditions on $\mathbf{\Sigma}_j$	Identifier	Distribution	Volume	Shape	Orientation
None	VVV	Ellipsoidal	Variable	Variable	Variable
$\lambda_j \boldsymbol{I}$	VII	Spherical	Variable	Equal	Undefined
Σ	EEE	Spherical	Equal	Equal	Equal
$\lambda \boldsymbol{I}$	EII	Ellipsoidal	Equal	Equal	Undefined

Table 3. Data sets used in the experiments.

dataset	# Observations	# Dimensions	# Classes
IRIS	150	4	3
SDSS	1084	4	3
Diagnostic	569	30	2
Wine	178	13	3

matrices. In four cases is condition equal like with our first four criteria on clustering. Since rule of E-optimal clustering we do not know express as solution of searching of parameters with maximal likelihood, so is not possible use EM-algorithm, neither function *Mclust*. In Table 2 are broken down identifiers, like it understand *Mclust*.

To evaluation we have used these 4 data sets: 1. IRIS Flower Data set [1], 2. SDSS point sources [3], 3. Breast Cancer Wisconsin (Diagnostic), 4. Wine. Breast Cancer Diagnostic and Wine datasets are available at `http://archive.ics.uci.edu/ml/datasets/`. In IRIS datas set we have 3 species of iris, in SDSS 3 types of stars, in Breast Cancer Diagnostic 2 types of tumors and in Wine data set 3 types of cultivars. Summarization of the data sets are in Table 3 and visualization in Figure 2.

In Table 4 are results of the experiments. Numbers in the first row are referring to our optimization criteria. Numbers from 1. to 4. are same as in Table 1 and numbers 5. and 6. are criteria 1 and 2 from Section 2.1.2. In the fields of the table are percentages of correctness of clustering. These values are calculated from clustering, which has lowest value of the optimization criteria from several runs of algorithms. In the SDDS data set was results with most criteria incorrect, it means that shape, orientation and volume of clusters were absolutely different from a real, so we can not calculate correctness of a clustering. Except our two criteria performs good clustering. In Wine data set we were again better than EM-algorithm and in IRIS we were equal good. In general we can say that the best clustering was performing with 3. criterion, which is in condition on covariance matrices $\mathbf{\Sigma}_j = \Sigma$.

5 Conclusions

In this paper we have described implementation of one of the most modern methods of multivariate statistical analysis. This method was model-based normal cluster analysis. In the first section we defined multivariate normal distribution and described its geometric interpretation, which helped us better understand the problem of cluster analysis. Then we described model-based normal cluster analysis and showed optimization criteria in the different conditions on covariance matrices of clusters. These conditions allowed us to vary different geometric features of clusters. We

Table 4. Results of the experiments. Results are in percentages (%) of correctness of clustering.

	1.	VVV	2.	VII	3.	EEE	4.	EII	5.	6.
IRIS	86.0	96.0	91.3	89.3	**98.0**	**98.0**	89.3	89.3	87.3	87.3
SDSS	86.5	-	-	-	**91.3**	-	-	-	-	-
Diagnostic	89.9	**95.0**	90.8	90.8	77.7	83.1	85.4	85.2	90.8	85.4
Wine	87.0	80.9	71.9	71.9	**98.8**	96.6	70.2	70.2	71.9	70.2

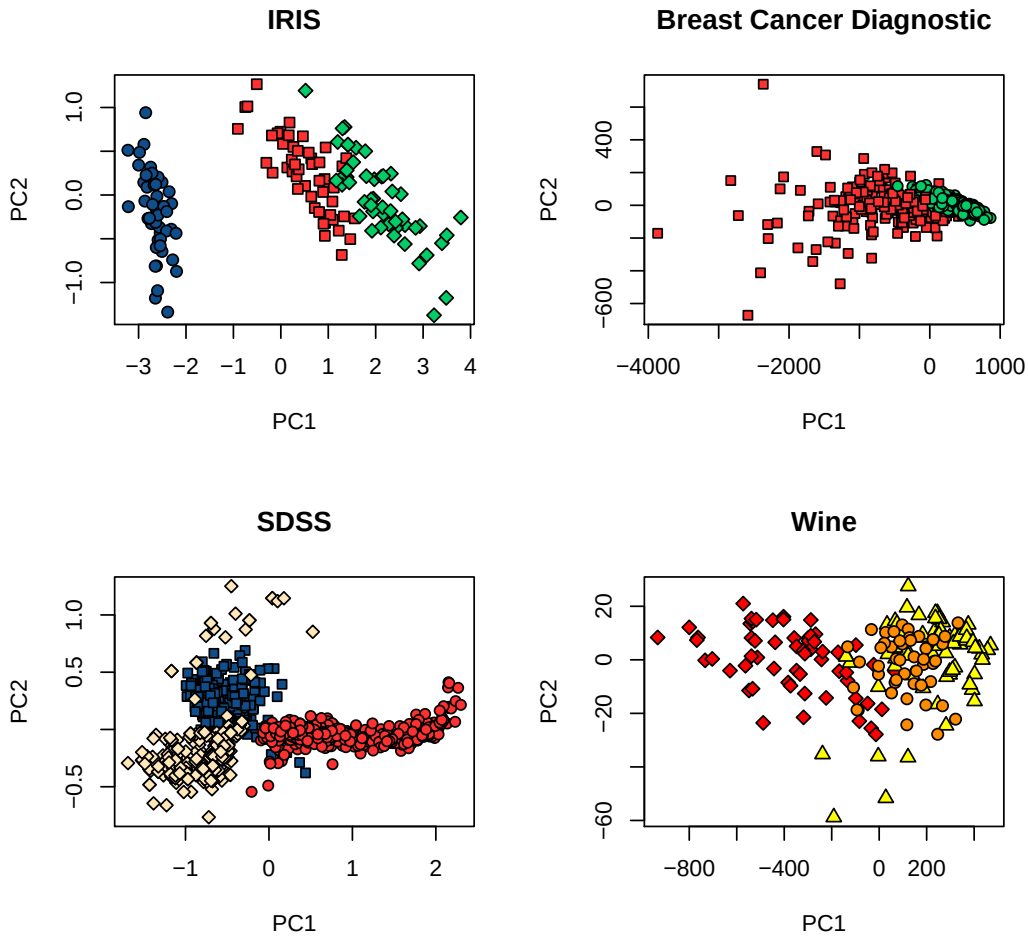


Figure 2. Data sets.

have presented interesting relationship of optimization criteria of clustering with criteria of optimal experimental design. For E-optimality we came up with entirely new optimization criteria, which represents interesting alternative to common used optimization criteria on clustering. In the next section we described genetic algorithm, which we programmed on cluster analysis. In the end we performed data analysis, in which we compared our clustering with real classification and clustering, which performed EM-algorithm implemented on the function *Mclust* in the software R. Our results were comparable to the results achieved by *Mclust*. In astronomical SDSS data set we achieved comparatively proper clustering with two optimization criteria. In contrast, the function *Mclust* classified the objects for each model not correctly.

Our hypothesis is, that these two criteria in our algorithm can find better local minima than the EM algorithm used in the function *Mclust*. In the future we can investigate, why we found better results exactly on these two criteria.

Acknowledgement: This paper was written on basis of our diploma thesis, which supervised Radoslav Harman.

References

[1] Anderson, E.: *The species problem in Iris, Annals of the Missouri Botanical Garden, No.23*. Biodiversity Heritage Library, 1936.

[2] Banfield, J. D., Raftery, A. E.: *Model-Based Gaussian and Non-Gaussian Clustering*. Biometrics, Vol. 49, No. 3, International Biometric Society, 1993.

[3] Feigelson, E. D., Jogesh Babu, G.: *Modern Statistical Methods for Astronomy: With R Applications.* Cambridge University Press, 2012.

[4] Fraley, Ch., Raftery, A. E., Murphy, T. B., Scrucca, L.: *mclust Version 4 for R: Normal Mixture Modeling for Model-Based Clustering, Classification, and Density Estimation Technical Report No. 597.* Department of Statistics, University of Washington, 2012.

[5] Fraley, Ch., Raftery, A. E.: *Model-based Methods of Classification: Using the Mclust Software in Chemometrics.* Journal of Statistical Software, Vol. 18, American Statistical Association, 2007.

[6] Haupt, R. L., Haupt, S. E.: *Practical Genetic Algorithms – second edition.* John Wiley & Sons, 2004.

[7] Lamoš, F., Potocký, R.: *Pravdepodobnost'a matematická štatistika, Štatistické analýzy.* Univerzita Komenského, 1988.

[8] Maulik, U., Bandyopadhyay, S.: *Genetic algorithm-based clustering technique.* Journal of the Pattern Recognition Society, Vol. 33, pp. 1455–1465, 2000.

[9] Pázman, A., Lacko, V.: *Prednášky z regresných modelov*. Univerzita Komenského, 2012.

[10] Pukelsheim, F.: *Optimal Design of Experiments.* Society for Industrial and Applied Mathematic, 2006.

[11] Rahman, A., Islam, Z.: *A hybrid clustering technique combining a novel genetic algorithm with K-Means.* Knowledge-Based Systems, Volume 71, pp. 345–365, 2014.

[12] Raposo, C., Antunes, C. H., Barreto, J. P.: *Automatic Clustering using a Genetic Algorithm with New Solution Encoding and Operators.* Computational Science and Its Applications — ICCSA, Lecture Notes in Computer Science Volume 8580, pp. 92–103, 2014.

[13] Wikaisuksakul, S.: *A multi-objective genetic algorithm with fuzzy c-means for automatic data clustering.* Applied Soft Computing, Volume 24, pp. 679–691, 2014.

Repeating Sequence Length Estimation in Time Series

Jakub ŠEVCECH*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
jakub.sevcech@stuba.sk

Abstract[1]

Over the past decades many different time series representations were proposed. These representations have to satisfy several requirements such as reduction of the data dimensionality or emphasis on fundamental shape characteristics present in the time series data. With the growing amount of data processed and specialized applications, additional requirements emerge. In our work, we focus on the need for incremental transformation of the data into the representation and the interpretability of the transformed data. We are looking for a time series representation suitable for stream data processing and straightforward interpretation of repeating sequences for tasks such as anomaly detection or stream classification. Based on previous works, a representation grouping repeating sequences into symbols seems to be a viable option for such task. However such representation would be limited to iterative processing of data due to the iterative nature of most commonly used clustering algorithms such as K-means. To overcome this drawback we use an incremental clustering algorithm not limiting the number of clusters but rather their radius.

As multiple works on applicability of data mining methods suggested, one of the main limitations for using various algorithms are their parameters and their necessity to be set or trained before the method application. As the proposed method has three parameters (window size, step between windows and cluster radius) we have to search for a methods to limit their negative impact on the ease of use of the method. As a mean for such simplification we search for a reliable method for window length setting. In this work we compare two methods for detection of lengths of repeating subsequences in the time series to determine the optimal length of the symbols (window) formed during the transformation. We compare autocorrelation and a method based on repetitiveness metric on synthetic and real world datasets.

As the results suggest, the autocorrelation function clearly outperforms the second compared method when processing synthetic datasets with different settings. The autocorrelation obtained significantly higher accuracy values and its computational complexity is much lower. These two results suggest the autocorrelation to be a clear choice. However on real world datasets the repetitiveness based method showed other interesting properties as it was able to identify not only the dominant pattern but also other minor patterns. The choice of the method thus remains to be determined based on the nature of the proposed data.

---

* Doctoral degree study programme in field: Software Engineering
Supervisor: Professor Mária Bieliková, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

[1] Full paper available in printed proceedings, pages 123–130.

Power Demand Forecasting from Stream Data with Concept Drifts

Petra VRABLECOVÁ*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
petra.vrablecova@stuba.sk

Extended Abstract

The directive 2009/72/EC orders all the member states of the European Union to measure at least 80 % of their power demand with the smart meters by 2020. A smart meter can send information about the electricity consumption every couple of minutes (e.g., quarter-hourly, half-hourly). The hundreds of smart meters from all over a country form a new source of big data. The numerous streams of data that simultaneously flow into a repository are characterized by the two main properties of big data – volume and velocity [3]. Since the power distributor has to fulfil the customers' demands and the power authorities' regulations, it is his task to exploit the data about consumption to predict the future demand and schedule the distribution accordingly.

The stream processing differs from traditional batch processing. The main restrictions of the method for the stream mining are the limited constant time and memory, a single scan over data, the ability to adapt to the changes in stream over time, i.e. concept drift and the necessity to provide similar results as the batch learner in spite of the aforementioned restrictions [2].

The concept drift is defined as a change in input variables or a change in the relationship between inputs and the target (predicted) variable [1]. The possible changes in a stream of data are depicted in the Figure 1. The examples of situations when changes can occur are a power demand during holidays, sport events, etc.

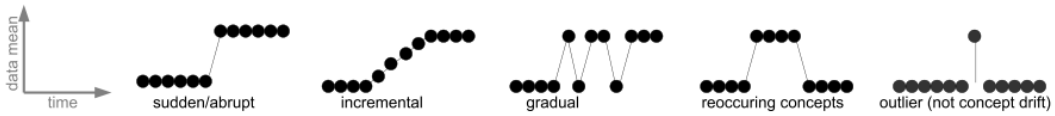


Figure 1. Four patterns of concept drift in a data stream. The x-axis is the time and the y-axis the mean value of the currently monitored portion of the stream. The outliers are not a concept drift [1].

The drifts make the prediction from stream data more difficult. It is necessary to detect the drift as soon as possible and *adapt* the prediction model to it if needed; to correctly differ it from noise; and to perform the detection in less than data arrival time and within the fixed size of memory [1]. *The online adaptive learning*, i.e. the process of changing of the predictive model over time,

* Doctoral degree study programme in field: Software Systems
Supervisor: Assoc. Professor Viera Rozinajová, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

consists of three steps [1]: (1) predict, (2) *diagnose* and (3) update. The diagnostics' part of the process concerns the concept drift detection.

In the power demand forecasting domain there are currently only a few methods that consider the concept drift, i.e. are *adaptive*. However, we have not met with any method that performs the *informed adaptation*, i.e. adapts the prediction model based on information from a change detection mechanism. Usually the methods adapt the prediction model *blindly*, i.e. regularly in the determined time intervals. Therefore it is our goal to find out whether the informed adaptation improves the power demand prediction and at the same time to devise a method that meets two conditions that emerges from stream processing restrictions we mentioned earlier. The method has to be:

- *incremental* – use a fixed amount of memory and time for computation of prediction for the next period on the arrival of new data and exploit the previous calculations to make it,
- *adaptive* – adapt the prediction model when a change in stream is detected, without the retraining the prediction model on all historical data (only single scan over data is possible).

In our work we studied the existing prediction methods and time series analysis methods. We analyzed the possibilities of modifying them to meet the two conditions. We chose as our underlying method the exponential smoothing because of its recursive calculation, which enables the incrementality and because of its simplicity, robustness and existing application in the power demand forecasting. We modified it so it adapts its smoothing parameter on concept drift detected by a change detection mechanism, which detects changes in the mean value of one-step ahead error. We compare the precision of this approach with the precision of the underlying method, which makes predictions without the change detection mechanism.

In future we want try out different underlying methods and different change detection mechanisms. The goal is to consider and incorporate in the concept drift detection the domain specifics (seasonality, weather conditions, etc.) that affects the electricity consumption, and to design the incremental and adaptive method with optimal precision, time and memory usage. We perform the evaluation on the Slovak smart-metering data (see sample in Figure 2).

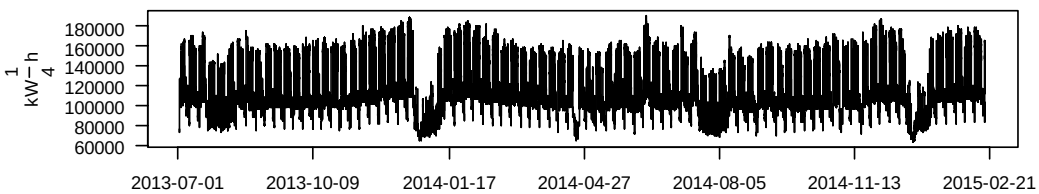


Figure 2. Electricity consumption in the Bratislava region.

Acknowledgement: This contribution was created with the support of the Research and Development Operational Programme for the project International centre of excellence for research of intelligent and secure information-communication technologies and systems, ITMS 26240120039.

References

[1] Gama, J. et al.: A survey on concept drift adaptation. *ACM Computing Surveys*, (2014), vol. 46, no. 4, pp. 1–37.

[2] Gama, J.: Data Stream Mining: the Bounded Rationality. *Informatica*, (2013), vol. 37, no. 1, pp. 21–25.

[3] Laney, D.: *3D data management: Controlling data volume, velocity and variety*. Stamford, CT: META Group, (2001), 3 p.

Web Science and Engineering

TV Program Guide Metadata Acquisition

Peter DUBEC*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
peter.dubec93@gmail.com

Abstract. Even today, there is still many people watching television and using TV program guide when deciding what to watch. It is thus important to describe entities of TV program with relevant and high quality metadata, which can later be used for recommendation of these items. In this paper we propose a method for automatic acquisition of metadata for TV program guide from online movie databases and Linked Open Data space. The method is designed to be able to extract metadata in English language and map them on entities of tv program guide described in Slovak language. Part of our method is automatic classification of documentary movies for the purpose of enriching tv program guide with our own metadata.

1 Introduction

Today we can find huge amounts of metadata for multimedia items published online in different forms. Quality of services aimed on multimedia directly depends on quality of metadata available for these multimedia items. Currently main problem is that items are published online in different languages, and we are trying to map these items on Slovak entities. We see a big problem in how titles of items of TV program guide are translated, because many times some words are translated differently or whole local title is completely different than original title so then it is a problem to find adequate entity published on online sources.

Our goal is to focus our work on TV program guide metadata enrichment aimed at mapping entities of Slovak TV program guide on entities published online for the purpose of improving quality of these services and also to enable personalized recommendation of TV program guide items [6]. This work is a part of bigger project aimed at TV program recommendation [10].

Metadata are being published online on different sources in structured or unstructured form. Related work in this area is described in Section 2. Our method focus on finding relevant, high quality metadata and on extracting them in structured form and also classification of documentary movies for the purpose of generating new metadata. We describe our method in Section 3. Evaluation and conclusions of our work can be found in Section 4.

---

* Bachelor study programme in field: Informatics
Supervisor: Professor Mária Bieliková, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 135–139.

2 Related Work

On the Internet we can find huge amount of different multimedia items and metadata are key factor for multimedia management. It is thus very important to describe multimedia items with relevant and high quality metadata. Linked Data seems as good possibility for finding these necessary metadata, but as they mention in [9] studies that focus on relationships between multimedia and Linked Data are really rare and there is lack of Linked Data publication frameworks for multimedia management solutions. Also as was mentioned in [3] key problem that has to be solved is development of metadata, because for managing multimedia data, metadata are much more important than in case of managing traditional text-only data.

Also as we can see in [4] the authors used only metadata extracted from Linked Data space to create fully working content-based recommender system. For this work were used three big datasets from Linked Data space: Dbpedia, Freebase and LinkedMDB[1]. So as we can see metadata extracted from Linked Data are sufficient for constructing whole recommender-system, and that is why we decided to use Linked Data as one of our sources for enriching TV program guide but later we found out that it is not sufficient for Slovak entities.

3 Metadata Acquisition

It is very important to find relevant and high quality metadata to describe items of TV program guide. Thus as a source of metadata we are using two big online movie databases, which can be considered reliable, because publishing of metadata is supervised by administrators. First is CSFD[2] where data are published in Czech or Slovak language and also have good structure so it is easy to work with them and we are also using CSFD as a source for our TV program guide.

Second database which we used as a source for metadata extraction is IMDB[3] which is international movie database where data are published in English language [8]. We use Linked Data [1, 2] as a next source for metadata. But we are extracting metadata only from two datasets DBpedia[4] and Freebase[5]. We chose DBpedia mainly because it is extracting data from Wikipedia [5] and it contains huge amount of data. On the other hand Freebase also contains enough high quality and nice structured data. As a final source of metadata we take our own automatically generated metadata, obtained from documentary movies classification.

Our method of metadata acquisition consists of three steps:

1. TV program guide extraction and basic metadata acquisition.
2. mapping items on Linked Data space and metadata extraction.
3. generating new metadata based on documentary movies classification.

3.1 Basic Metadata Acquisition

First we parse TV program guide from CSFD database. This process is fully automatic and we are able to parse TV program for specified attributes as different day or channel. During parsing of TV program, when item is extracted it is automatically mapped on it's representing entity in CSFD

[1] http://www.linkedmdb.org/
[2] http://www.csfd.cz/
[3] http://www.imdb.com/
[4] http://dbpedia.org/
[5] http://www.freebase.com/

```
SELECT DISTINCT ?film
        WHERE {
        ?film a dbpedia-owl:Work ;
        rdfs:label ?label .
        filter contains( ?label, "The Terminator" )
    }
    LIMIT 100
```

Figure 1. Basic SPARQL query example.

movie database. Since as a source program guide we use CSFD program guide, mapping items on their representing entities in this database is only a simple process of extracting representing links.

Next we map extracted entities on it's representing entities in International Movie Database (IMDB). Since in first step entities were mapped on their representing entities in CSFD and these entities contains links to their representing entities in IMDB this process of mapping is again only extraction of links. After these two steps, for each item of TV program guide, we have it's representing entity in CSFD and IMDB databases (if it exists in these databases).

Now we are able to enrich items of TV program guide with metadata which can be extracted from these online movie databases. We are not extracting everything, but only metadata which we consider relevant and important for users. We are extracting basic metadata such as item local and original title, year of distribution, item genres and some important persons connected with item such as directors or lead actors. We are also extracting metadata that we think might be interesting for users such as filming locations for each item, if they are available.

3.2 Extraction from Linked Data

We extract metadata from entities in Linked Data space. Since entities published in Linked Data are described in English language and items of TV program are described in Slovak, we need English equivalent for each item which we want to map on Linked Data space. So it means we can map only those entities, which in previous step were found in IMDB because then we have available metadata in English language for these items. We also tried automatic translation for items that were not found in IMDB so that next we would be able to map them on entities in Linked Data space, but we found out that items that were not present in IMDB were not present in Linked Data space too.

Process of mapping items on Linked Data space sends specified queries in SPARQL (see Figure 1) for DBpedia endpoint[6] and in MQL (see Figure 2) for Freebase endpoint[7]. These queries are constructed from item attributes, which we have available in English and on beginning these queries are very strict, what means they contains as many attributes as possible. When there is no match found, some attributes are deleted from query to make it more general and to increase possibility of finding match. When a match is found and an entity is found in Linked Data space and the item of TV program guide contains some identical attributes, we consider this entity as it represents the entity so we can parse additional metadata from this entity.

3.3 Documentary Movies Classification

Next we enrich TV program guide with our own automatically generated metadata. We decided to focus only on documentary movies, because we found out that classification is not appropriate and

[6] http://dbpedia.org/sparql

[7] https://www.freebase.com/query

```
[{
  "id": null,
  "name": "The Terminator",
  "limit": 100,
  "type": "/film/film"
}]
```

Figure 2. Basic MQL query example.

necessary for every genre of TV program guide item. Our aim is to achieve automatic classification of items with documentary genre. Documentary movies are classified into smaller categories based on their content description and labels of these categories are added to items as new metadata. Our aim is to acquire such metadata that can serve for more accurate recommendation of documentary movies.

First we create labels for subcategories (nature, science, history etc.) into which items are divided. Next we tag items with representing keywords generated from item descriptions in English using AlchemyAPI[8]. When item description is not present in English, but in another language, Google Translator API[9] is used for translation into English. We use translation into English because of keyword extraction API limitations. It is able to properly extract keywords only from texts in English. Final step is to classificate items into correct subcategory based on generated keywords with use of SVM (Support Vector Machines) [7]. When item is classified into some specified subcategory, label of this category is added to item as new additional metada. With these metadata, it will be much more easier to recommend items to users based on their genre and also it will be easier to find similar items which can be recommended to user.

4 Evaluation and Conclusions

In our work we enrich Slovak TV program guide with metadata extracted from online movie databases CSFD, IMDB and also from Linked Data space. Moreover, we concentrate on one specific category of TV program items – documentary movies for which we acquire additional metadata based on documentary movies classification. This work is a part of larger project aimed at personalized recommendation of TV program for Slovak televiewers. This purpose primarily influenced our focus of metadata enrichment.

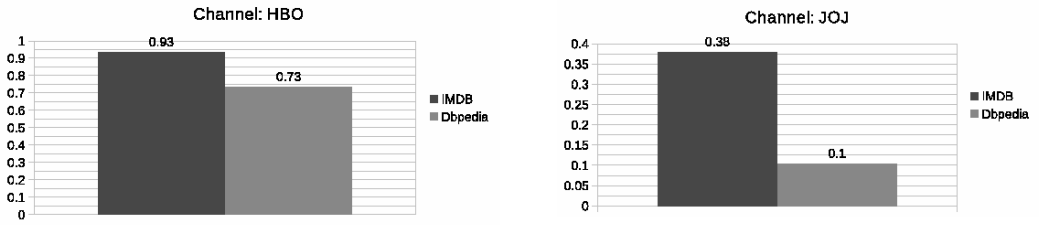


Figure 3. HBO and JOJ channel items statistics

During the process of implementing and testing our method for automatic metadata acquisition, we found out that it is very difficult to enrich Slovak or Czech entities of TV program guide, because

[8] http://www.alchemyapi.com/

[9] https://cloud.google.com/translate/

these entities are not present in international movie databases and neither in the Linked Data space. We tested our method on different TV channel items for a period of one day. Figure 3 illustrates situation: more than 90 % of items from channel HBO was found in IMDB and more than 70 % was found in DBpedia dataset. On the other hand, situation with Slovak entities is much worse. Only 38 % of items was present in IMDB and only 10 % in DBpedia dataset. In further work we plan to evaluate our proposed approach for automatic metadata enrichment of TV program with focus to documentary movies classification. We also plan to look for new ways how to generate useful and interesting metadata for Slovak or Czech entities of TV program guide.

Acknowledgement: This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG 1/0646/15.

References

[1] Berners-Lee, T.: Linked Data. Available at `http://www.w3.org/DesignIssues/LinkedData.html`, 2006, [Accessed 01.12.2014].

[2] Bizer, C., Heath, T., Berners-Lee, T.: Linked Data – the story so far. *International Journal on Semantic Web and Information Systems*, 2009, vol. 5, no. 3, pp. 1–22.

[3] Boll, S., Klas, W., Sheth, A.P.: Overview on Using Metadata to Manage Multimedia Data. In: *Multimedia Data Management: Using Metadata to Integrate and Apply Digital Media*, 1998, pp. 1–23.

[4] Di Noia, T., Mirizzi, R., Ostuni, V.C., Romito, D., Zanker, M.: Linked Open Data to Support Content-based Recommender Systems. In: *Proceedings of the 8th International Conference on Semantic Systems*. I-SEMANTICS '12, New York, NY, USA, ACM, 2012, pp. 1–8.

[5] Exner, P., Nugues, P.: Entity extraction: From unstructured text to DBpedia RDF triples. In: *The Web of Linked Entities Workshop (WoLE 2012)*, Boston, USA, 2012.

[6] Hu, Y., Wang, Z., Wu, W., Guo, J., Zhang, M.: Recommendation for Movies and Stars Using YAGO and IMDB. In: *Web Conference (APWEB), 2010 12th International Asia-Pacific*, Busan, Korea, IEEE, 6–8 April 2010, pp. 123–129.

[7] Joachims, T.: Text Categorization with Suport Vector Machines: Learning with Many Relevant Features. In: *Proceedings of the 10th European Conference on Machine Learning*. ECML '98, London, UK, UK, Springer-Verlag, 1998, pp. 137–142.

[8] Peralta, V.: Extraction and integration of movielens and imdb data. Technical report, Technical Report, Laboratoire PRiSM, Université de Versailles, France, 2007.

[9] Schandl, B., Haslhofer, B., Bürger, T., Langegger, A., Halb, W.: Linked Data and Multimedia: The State of Affairs. *Multimedia Tools and Applications*, 2012, vol. 59, no. 2, pp. 523–556.

[10] Šimko J., Michal, K., Dusan, Z., Maria, B.: Tool for Social Fazeted Multimedia Exploration. In: *8th Workshop on Intelligent and Knowledge Oriented Technologies (in Slovak)*, Slovaka, 2013.

Slovak Web-Based Encyclopedia

Patrik GAJDOŠÍK*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
gajdosik.patrikk@gmail.com

Abstract. The Semantic Web is a big topic these days. Structuralized data give much better way for finding relations that would be very hard to spot before. Slovak Web also contains certain amount of information and the goal is to make them accessible through the Semantic Web as well. One of the big players in the Web of Data is DBpedia that already has prepared a solid ground that allows the international extraction of information. We propose a method that allows DBpedia to extract information in other language mutations of Wikipedia by automatically creating mappings that DBpedia uses for this purpose. First, our method creates mappings of infoboxes to the classes and mappings of individual attributes found in the infoboxes to the predicates in the DBpedia ontology. For this task we use the existing mappings of different language mutations. By using different metrics we focus on finding more qualitative than quantitative results.

1 Introduction

The Web contains a huge amount of unstructured or semi-structured information of different qualities floating around. In order to search through them and use them effectively, the concept of Semantic Web comes to aid with transforming this chaos into structured machine-readable data. After creating these blocks of knowledge and finding the right relations in between them the cloud of Linked Data is growing, allowing to search for the data in more specific and precise way.

Nowadays, one of the great sources of information on the Web is Wikipedia. Although it contains huge amounts of data, they come in a form that has lost its structure comprehensible for computers. Existing solutions such as DBpedia [5] or YAGO [7] (in its later version YAGO2 [3]) solve this problem by extracting the offered information and giving them their meaning and also finding the right relations between them.

For Slovak web and Slovak language being not the most used in the world comes lots of unstructured data (e.g. Slovak Wikipedia contains around 195 000 articles that are in no way processed) that, when appropriately filtered, could actually enrich the Web of data. This issue is applicable to all other languages with a smaller number of speakers like Slovak.

In order to start the extraction, additional metadata are needed for the DBpedia to be capable of doing so. The process of creation of this metadata can be automated thanks to the already existing

* Bachelor study programme in field: Informatics
Supervisor: Michal Holub, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

variations of these metadata for different language versions. A method that would be capable of performing a task like this would allow for the information in different language that are not part of the DBpedia to be included in the cloud of Linked Data.

2 Related work

The great reference for this work is DBpedia [5], the big ontology that in its current state is the beacon of the cloud of Linked Data. DBpedia extracts information from Wikipedia, adding semantics to it. It is semi-automatic method that needs mappings (which are created manually) that will create relations between Wikipedia infoboxes and classes in the DBpedia ontology. The works done in [2,4] have added support for different language versions of Wikipedia to the DBpedia. The problem is that even though the DBpedia is internationally oriented, infoboxes are different in each language and the mappings need to be created for the specific language anew.

Yet Another Great Ontology is YAGO [7]. It works on a different principle than DBpedia, by utilizing Wikipedia's category pages. It does not focus on all the data on Wikipedia. There is a set of rules that extract few particular information but this extraction is done well and with high quality. YAGO uses WordNet and GeoNames to maximize the quality of extracted data. The problem with YAGO is that it is solely focused on English, even though the last version – YAGO 3 – should add the possibility of multilingual extraction.

Airpedia project [6] trains a supervised classifier using classes connected through interlinks as training data. They add more than 1 million entities with high precision and recall. Finally, a similar work that focuses on creating automatic mappings was done in [1]. We base our own method upon this research, we focus on how to start suggesting mappings for infoboxes and their attributes. The work has high precision results.

3 Semantification of the Slovak Web

As mentioned before, the environment around gathering knowledge from the Web was from the beginning solely centered around English language. DBpedia and YAGO, they both were built in the English environment and intended to mainly work with that language. However, nowadays there is certain progress that allows other language mutations to be added to the facts composing the Web of Data. One of the ways that helps the cloud grow is internationalization of DBpedia which combined with the robustness of Wikipedia makes it a luxurious destination.

3.1 DBpedia

As of today, DBpedia includes extraction of data from almost every language mutation of Wikipedia. However it is only the general information such as pages used to disambiguate other pages, finding links to other sources etc. The point of interest for us are infoboxes that contain much more information. What is needed are mappings for Slovak language so that the DBpedia then can take a Slovak infobox and extract structured data from it. The mappings consist of two parts:

- mappings of an infoboxes to the classes in DBpedia ontology (e.g. *Infobox_štát* is mapped to *dbpedia-owl:Country*),
- mappings of attributes of a certain infobox to the predicates in the DBpedia ontology (e.g. *hlavné_mesto* is mapped to *dbpedia-owl:capital*).

Our method tries to find these mappings automatically by using mappings of different languages that already exist. First, we assign each infobox a class and then we map individual attributes to the predicates. The foundations of our method are based on the work in [1] that covers the area of mapping an infobox to a class in the DBpedia ontology.

3.1.1 Mapping of an infobox to a class

Mapping of infoboxes to classes is done with the help from other language mutations of DBpedia that already have their own mappings created. Our method looks at all possible classes to which a certain infobox could be mapped and decides which one it will pick for a Slovak mapping. To decide the mapping of one infobox, all the resources that use it need to be gathered. This information is already extracted by DBpedia.

Then, for each Slovak resource, resources in different language mutations are obtained. RDF property *type* shows to which class are the different language mutations of the resource mapped.

Our method proposes a mapping for Slovak resource based on two rules adapted from [1]:

1. the most specific common class of resources of all language mutations is found (as shown in Figure 1, if we only consider the first two branches, the last common class is chosen),

2. if some branch ends but so far it was part of the hierarchy with other possible classes, the process continues (looking at the third branch in Figure 1, the process does not stop at *Place* because so far all the classes matched).

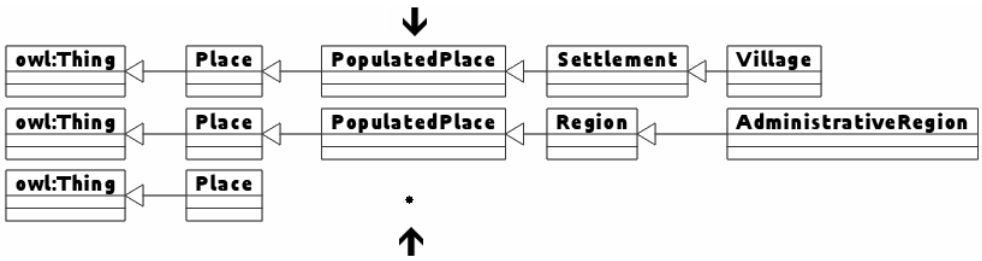


Figure 1. Selecting the class for a mapping suggestion.

After all the resources have a suggestion for a mapping, a histogram is created which will represent possible mappings of the infobox having the number of appearances as score. This is in the process of infobox mapping considered to be the basic scoring metric.

We propose an extension of this scoring metric which focuses on picking the most specific class as possible. If an ancestor of a class is also present in the histogram, this class takes the score of the ancestor as its own and the ancestor can be removed from the possible mappings. The process starts at the top level class and step by step goes to the lower levels of the inheritance tree transferring points from higher classes to their descendants. This way *Village* also gets points from *Settlement* which got points from *PopulatedPlace* etc. In the end, only leaf classes are left to choose from.

In comparison with the extension that focuses on finding the most specific class, we also propose a metric that takes generalizing approach. The result of the basic scoring metric is taken (the classes from the histogram) and the leaf classes are picked. Then their common ancestor is found. It will be considered the only possible mapping. This method takes a safer approach, and will result in the most correct results with lower quality (e. g. a village infobox will be assigned the class *Place* and not the more accurate *Village*).

It's up to the user to pick which metric to use, whether the basic scoring metric or the other two extensions to it. After the application of the scoring metrics we get all the possible classes to which an infobox could be mapped. Each has its score that was given to it and the class with the highest number of points is chosen. After this, the mapping for Slovak language is generated in a form used by DBpedia[1] and can be used in the extraction process of DBpedia.

[1] The database of language mappings is in form of a wiki that is placed at `http://mappings.dbpedia.org`.

3.1.2 Mapping of attributes to predicates

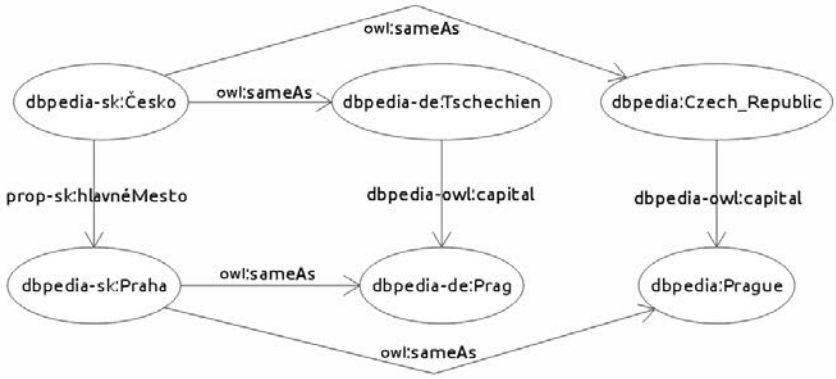


Figure 2. Mapping of properties which have DBpedia resource as a domain.

The mapping of attributes of infoboxes to predicates in the DBpedia ontology is a harder task to perform. In this case, what we can do is to compare the data in between multiple language versions and if it matches, the mappings of infobox attributes to predicates can be inspected.

If a value is tied to a Slovak resource through a property (DBpedia extracts raw properties that have the same name as an attribute of the infobox and are in separate namespace), our method looks, whether in other language mutations there also exists a property that ties it to the same value. Hopefully, this property will be part of the DBpedia ontology which is what are we looking for. System of points is put to work here and for every found match, n-th of a point is added to the found possible mapping, n representing the number of the language mutations we use for the search of mappings. The possible mapping with the most points (and having at least one point) is selected to be the suggested mapping of an attribute to a predicate.

In Figure 2, the example of the process described above is shown. We are looking for possible mapping of *prop-sk:hlavnéMesto*. The domain of this property is another resource in DBpedia. We use interlinks to find the equivalent of the resource (Česko) in other language mutations (through predicate owl:sameAs) of DBpedia. In this case however, we also have to find other variations of the object (Praha) since it's also a resource. As it shows, two other language mutations for this relation use a property from the DBpedia ontology – *dbpedia-owl:capital*. A mapping from *prop-sk:hlavnéMesto* to *dbpedia-owl:capital* gains one point (half a point for German version and half a point for English version of DBpedia).

3.2 Tool for managing the mappings

Part of our solution is a tool for managing the automatically generated mappings and for verifying the correctness of data extracted using the mappings. Some of its main capabilities are listed below:

- allows users to begin the process of extraction, review the generated mappings, change them, add or remove mappings;
- allows users to start the process of extraction using DBpedia and the generated mappings, through all the extracted data, see them in comparison with the real infobox;
- gives the users the possibility to mark whether the extracted information is correct or not which will be remembered by the system that then learns from it.

4 Evaluation

First thing that is evaluated is the result of automatic creation of infobox mappings – how many mappings have been created and with what precision. For the mappings of infoboxes to the classes in DBpedia ontology, we executed our method to create automatic mappings for 40 infoboxes. The results can be seen in Table 1. We use three metrics of deciding the mapping as described in Section 3.1.1 with different score ratio being required. The number of infoboxes for which a mapping was found is shown in the column Quantity, and column Precision showing ratio of correct mappings. Since we are mapping all the infoboxes, the recall is not valid in our case.

The basic metric gives us good results. Most of the results are correct and even at 0.5 ratio being set, the results are accurate. The specialization metric shows its weak spots with giving average results. The wrong mappings are mostly caused by classes that have a very small score in the beginning but accidentally fit into the hierarchy. This metric could be improved by eliminating low scored possible mappings so they do not cause errors in the results. The last generalizing metric does not give us very good results. The score ratio does not affect this metric and the results are the same. However, what cannot be seen from the table is that this metric gives us correct results for the few infoboxes that have many possible mappings (e.g. *Infobox živočíchy*). These infoboxes are problematic for the two previous metrics, therefore there is place for improvement of our method by wisely deciding when to use which metric.

Table 1. Success rate of an automatic mapping of 20 infoboxes in DBpedia.

Required score ratio	Basic metric		Specialization metric		Generalizing metric	
	Quantity	Precision	Quantity	Precision	Quantity	Precision
0.5	29	0.86	26	0.77		
0.7	24	0.96	23	0.78	23	0.56
0.9	17	1.00	19	0.79		

The evaluation of infobox mapping is strict since only the results that match an expected result are marked correct. A different technique could be used that would count how far from the expected mapping we are. If we map the infobox Scientist to the class Person, it is not exactly wrong, because a scientist is a person. For now, we calculate the correctness of a mapping as a ratio of how deep in the inheritance tree is the class we have created mapping to to how deep the expected class is.

Table 2. Results of the automatic mapping of infobox attributes.

Infobox name	# of infobox attributes	# of created mapping	Precision of created mappings
Infobox štát	225	19	0.73
Infobox Film	83	16	1.00
Infobox Kniha	121	8	1.00
Infobox Softvér	49	10	0.90

Next part of the process is also mapping of infobox attributes. We evaluated attributes from 21 infoboxes with the precision being at around 85 %. The few examples can be seen in Table 2. For each infobox, the process of automatic mapping is separate. Since our method only creates mappings for numeric values or values that are resources in DBpedia, the number of created mappings is not very high compared to the number of all attributes for an infobox[2], however, the precision is quite good. Since we create mappings for all attributes, recall is not valid.

[2] The number of attributes for an infobox is obtained from the statistics page of DBpedia mappings wiki at `http://mappings.dbpedia.org/server/statistics/sk/`.

For a future evaluation, the evaluation of mapping all the infoboxes could show overall quality instead of just assessing a few picked infoboxes. Also, improvement of the described metrics will hopefully bring better results. Further evaluation for property mapping is also needed to show, how would our method work if it would consider properties throughout all the infoboxes instead of creating separate environment for each infobox.

5 Conclusions and future work

We proposed a method for automatic generation of mappings of infoboxes. The proposed method also automatically generates mappings for infobox attributes, which is the most important part for extracting data in DBpedia. For our purposes we also developed a tool that helps with the evaluation and later on could be used for overall assesment of DBpedia mappings.

The main goal of our work is to improve the correctness of the automatically created mappings. For the mapping of infoboxes, trying to make the three described metrics work together could create interesting results. Automatically setting the required score ratio according to the underlying dataset or automatic picking of the right metric for individual infobox might improve quality of our results.

To solve the problem of creating better mappings of attributes, the system of Wikipedia comes into question since it also has some information about the data stored inside of it and knows when to put the right units after some number. However, the problem lies with the need to still manually map this value to the predicate in DBpedia or create a table of common types beforehand.

Another way that could improve the results is to use NLP, which would be capable of deciding whether two words in two different languages have the same lexical meaning. The name of the attribute can be compared with the name of predicate in the DBpedia, too. There is no need for manual intervention in this case. Combining two of these methods could bring better results with higher quality and better precision.

Acknowledgement: This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG 1/0646/15.

References

[1] Aprosio, A.P., et al.: Automatic Mapping of Wikipedia Templates for Fast Deployment of Localised DBpedia Datasets. In: *Proc. of the 13th Int. Conf. on Knowledge Management and Knowledge Technologies – i-Know '13*, New York, New York, USA, ACM Press, 2013, pp. 1–8.

[2] Hellmann, S., Stadler, C., Lehmann, J.: The German DBpedia: A Sense Repository for Linking Entities. In: *Linked Data in Linguistics*. Springer Berlin Heidelberg, 2012, pp. 181–190.

[3] Hoffart, J., Suchanek, F.M., Berberich, K., Weikum, G.: YAGO2: A spatially and temporally enhanced knowledge base from Wikipedia. *Artificial Intelligence*, 2013, vol. 194, pp. 28–61.

[4] Kontokostas, D., Bratsas, C., Auer, S., Hellmann, S., Antoniou, I., Metakides, G.: Internationalization of Linked Data: The case of the Greek DBpedia edition. *Web Semantics: Science, Services and Agents on the World Wide Web*, 2012, vol. 15, no. Ld, pp. 51–61.

[5] Lehmann, J., Isele, R., Jakob, M., Jentzsch, A., Kontokostas, D., Mendes, P., Hellmann, S., Morsey, M., van Kleef, P., Auer, S., Bizer, C.: DBpedia – A Large-scale, Multilingual Knowledge Base Extracted from Wikipedia. *Semantic Web Journal*, 2014.

[6] Palmero Aprosio, A., Giuliano, C., Lavelli, A.: Automatic Expansion of DBpedia Exploiting Wikipedia Cross-Language Information. In Cimiano, P., Corcho, O., Presutti, V., Hollink, L., Rudolph, S., eds.: *The Semantic Web: Semantics and Big Data*. Volume 7882 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2013, pp. 397–411.

[7] Suchanek, F.M., Kasneci, G., Weikum, G.: YAGO: A Core of Semantic Knowledge Unifying WordNet and Wikipedia. In: *Proc. of the 16th int. conf. on World Wide Web - WWW '07*, New York, New York, USA, ACM Press, 2007, p. 697.

User Reputation in Community Question Answering

Adrián HUŇA*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
huna.adrian@gmail.com

Abstract. Community Question Answering systems (e.g. StackOverflow) have gained growing popularity in the last years. With the increasing amount of user generated content, a problem of identifying skilled and reliable users arose. The existing approaches estimate user reputation especially according to the overall user activity in the system, while the quality of user's contributions is considered only secondary. To address this drawback, our main goal is to estimate user reputation with focus on the quality of user's contributions. We proposed a reputation schema that takes the amount of carried out activity as well as expertise into consideration. The calculated reputation was consequently compared with four baseline methods. The experimental results showed a higher precision achieved by our approach, and confirm an important role of contribution quality in estimation of user reputation.

1 Introduction

The Internet is an enormous source of information which helps lots of people every day. Despite the amount of information available, there are still situations in which it is difficult to find specific information, or to answer a question that is too complex to be understood by a search engine. These types of situations led to creation of online communities whose members are focused on helping each other in a specific area. In the past years, especially many Community Question Answering (CQA) systems have appeared and gained popularity among users. Their principle is simple – everybody can ask a question, and anyone can answer it. Other members of the community then vote on the provided answers with the aim to select the most valuable answer according to the community. Consequently, the asker can pick an answer and mark it as the best answer, what also serves as an expression of quality. All questions and answers are publicly available, and thus the CQA systems serve as centers of community knowledge. We can observe two types of CQA systems. Universal systems consisting of categories from physics, to love or psychology (e.g. Yahoo! Answers), and specialized systems which focus only on a specific area (e.g. StackOverflow that concerns with programming).

* Bachelor study programme in field: Informatics
Supervisor: Ivan Srba, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 146–151.

A traditional problem in systems that employ user generated knowledge is recognition of expert users. Their identification is important in order to extend their rights in managing the community, to mentor them for better engagement with the site, or to route hard questions. Expert identification is also an important part in the process of estimating user reputation. In general, reputation is an indicator of a value of a user to the community, and it should reflect a combination of user expertise and activity. In other words, the more expert answers a user can provide, and the more frequently he/she participates in the question answering process, the more reputation he/she should have.

Most CQA systems include some kind of method to calculate and visualize user reputation. These methods, however, often employ very simple principles, what leads to an inaccurate reflection of user expertise. There are many actions in CQA systems which can be analyzed in order to calculate users' reputation more accurately. It is possible to observe user's behavior in terms of asking and answering questions, look at the feedback from community voting, or study a social graph between askers and answerers. We suppose, that by focusing on the quality of users' contributions, we will be able to better estimate user reputation than approaches in the previous works.

2 Related work

Reputation in CQA systems has been intensively studied in the previous years. We can broadly divide the methods into graph-based and feature-based approaches. The graph-based approaches work with a social graph underlying users' interactions in CQA systems (mainly between askers and answerers). Algorithms developed to rank websites, such as PageRank and HITS, are then applied on these graphs in order to identify experts in the community. Feature-based approaches extract features about users and content which are available in CQA systems. We can identify two important types of features – features representing user activity and expertise.

An early attempt in expert identification in CQA systems was made by Zhang et al. [7] who studied users' expertise in a system called Java Forum. The biggest influence for further research in this area comes from proposal of a new feature-based expertise measure called Z-score, which works with number of answers and questions a user contributed. They also provided comparison between graph-based and feature-based approaches. Z-score came out as the best method.

Jurczyk et al. [2] studied performance of HITS algorithm on two types of graphs and different categories in Yahoo! Answers. They used Degree (number of connections entering a node minus number of connections outgoing) as baseline, and also examined performance differences for different structures of categories. Liu et al. [3] proposed to emphasize the relationship between best answerer and other answerers, and named this approach as a competition based. In comparison with the previous graph-based approaches, they used methods for ranking players for pairwise comparison between the users. The effectiveness of these ranking methods was compared with graph-based algorithms PageRank and HITS, and also with feature-based approaches – Number of Answers, Number of Best Answers, Best Answer Ratio and Smoothed Best Answer Ratio. The results showed that the proposed graph-based approach achieved very similar or even worse performance as much simpler feature-based metric Best Answer Ratio. Pal et al. [4] modelled a question selection bias in order to identify experts. They hypothesized that experts provide answers which will be considered helpful, and choose those questions which have no or only unsatisfying answers.

Yang et al. focused on the quality of users' contributions for expert identification. They identified two important types of users in CQA systems – very active users (called sparrows) who generate most of the content, and users (called owls) who provide valuable answers to questions that are perceived as important by the community [6]. Users are labeled by a metric called Mean Expertise Contribution which takes question debatableness and answer utility into calculation. Furthermore, they provide a comparative study between sparrows and owls. Question difficulty was taken into consideration also by Hanrahan et al. [1]. They decided to use duration between the time when the question was asked and the time when an answer was marked as the best answer as the measure for question difficulty and consequently for user expertise. They, however, did not propose any method for reputation estimation, only observed correlation between StackOverflow reputation and Z-score.

The state-of-the-art approaches to reputation estimation provide directions for a proposal of our method. At first, feature-based approaches perform better than graph-based ones. Secondly, in feature-based approaches it is essential to distinguish between user activity and expertise. In spite of that, the most of existing approaches give a priority on the amount of user activity, while just a few approaches prefer quality of user contributions (e.g. [1,6]). The final challenge is to measure quality and difficulty of users' contributions. Even though a method to measure the quality was already proposed [6], this method does not take asking questions into consideration, and precision of the method was not evaluated.

3 Estimating reputation in CQA systems

Our main goal is to model users' reputation with accentuation on the quality of users' contributions, not their activity as it is done in the popular CQA systems and the previous studies. As user reputation is a combination of user activity and expertise, there will be still a possibility for less expert users to gain the same amount of reputation as an expert user may have got for their high-quality contributions. However, it will be necessary to carry out much more activity.

In our approach, reputation of a user consists of reputation gained for answering questions as well as for asking them. A user gains greater reputation for asking difficult questions and for providing useful answers on other difficult questions. The gained reputation for such actions is added to previously earned reputation. Final reputation R of a user u can thus be expressed as a sum of reputation gained for asking questions R_q, summed up with a sum of reputation gained for answering questions R_a. Formula 1 represents the formal expression of the final reputation:

$$R(u) = \sum R_q(q) + \sum R_a(a, q) \tag{1}$$

We also propose an alternative formula in order to completely eliminate an influence of amount of users' activity:

$$R(u) = \frac{\sum R_q(q) + \sum R_a(a, q)}{|q| + |a|} \tag{2}$$

where $|q|$ is the number of questions a user asked and $|a|$ is the number of answers he/she provided.

3.1 Reputation for asking questions

Inspired by the previous work [1], we decided to calculate reputation for asking questions based on question difficulty in combination with question utility. We can expect that the longer it takes for the first answer to be added (time to answer a question $q - TTA(q)$), the more difficult the question is. In order to take into account differences between various topics in CQA systems, we decided to normalize this time by maximum time to add the first answer for questions assigned to the same topic t ($TTA_{max}(t)$). If a question belongs to more topics, $TTA_{max}(t)$ values for each topic are averaged and used in the place of $TTS_{max}(t)$. Question difficulty D_q for a question q can be formally formulated as:

$$D_q(q) = \frac{TTA(q)}{TTA_{max}(t)} \tag{3}$$

The second factor for calculating reputation for asking questions is question utility QU. Our formula for question utility is an adaptation of an idea in the previous work [6]. We calculate question utility as a score $Score$ (number of positive votes minus number of negative votes) normalized by a maximum value of scores – $MaxScore(t)$ on questions in the same topic t to reflect differences between topics in CQA systems. Similarly as for question difficulty, if a question belongs to more than one topic, we use the average of $MaxScore(t)$ for all topics of the question.

$$QU(q) = \frac{Score(q)}{MaxScore(t)} \tag{4}$$

If a question receives negative score, question utility will be negative too. To calculate negative utility more accurately, we use absolute value of minimum question score for a topic t in the place of $MaxScore(t)$.

We decided to use sum of question difficulty and question utility for the final form of formula for reputation obtained for asking questions. Formula 5 displays the final relationship for calculating reputation R_q for asking a question q:

$$R_q(q) = D_q(q) + QU(q) \tag{5}$$

3.2 Reputation for answering questions

The algorithm for calculating reputation for answering questions utilizes question difficulty (Formula 3) as described in the previous section, and combines it with answer utility which adapts an idea from the previous work [6]. Answer utility $AU(a, q)$ for an answer a in a question q is calculated as:

$$AU(a, q) = \frac{Score(a)}{MaxAnswerScore(q)} \tag{6}$$

where $Score(a)$ is a score of an answer a, and $MaxAnswerScore(q)$ represents a maximum score from all answers provided for a question q. If an answer receives negative score, answer utility will be negative too, as the same approach as for Question Utility is used. As well as in Formula 5, we have decided to use sum between question difficulty and answer utility for calculating reputation gained for answering a question:

$$R_a(a, q) = D_q(q) + AU(a) \tag{7}$$

4 Evaluation

We are not aware of any standardized gold standard that could be used to evaluate our results against. Nevertheless, many different approaches have been already employed in the previous works. The most objective way to evaluate the performance of user reputation estimation without manual data labelling is, probably, a comparison of two sorted lists of users for each question. The first list is sorted according to calculated reputation, while the second one is sorted according to the score of answers as accumulated in the CQA system. This gives us the ability to see how many users are in their correct position as well as examine the difference in ranking between the two lists. Standard information retrieval metrics are applied in order to compare performance of our method and baselines.

As the baselines for comparison, we chose four feature-based approaches. Firstly, we tried to reconstruct the original user reputation based on StackOverflow reputation rules[1]. As the second method for comparison, we chose Best Answer Ratio for each user (BAR), which performed as the best in the previous works. As the third method, we chose Z-score, as proposed by Zhang et al. [7]. Finally, we employed a number of previously posted answers, which reflects only user activity and totally ignores quality of provided contributions.

The experiment was performed on data from CQA system Programmers[2], which is a part of StackExchange network. The data is publicly available to download on archive.org[3]. We adapted standard information retrieval metrics to evaluate the performance of the methods:

[1] `http://stackoverflow.com/help/whats-reputation`

[2] `http://programmers.stackexchange.com/`

[3] `https://archive.org/details/stackexchange`

– Precision at N (P@n): The proportion of top N users who are ranked at the correct position.

$$P@N = \frac{r}{N} \tag{8}$$

where r is the number of users in the correct position.

– Mean Reciprocal Rank (MRR): The reciprocal rank is the inverse of position (according to the ground truth) for the user with highest reputation (evaluated by the proposed method). The mean reciprocal rank is the average of reciprocal ranks for all questions evaluated:

$$MRR = \frac{1}{|Q|} \sum_{i=1}^{|Q|} \frac{1}{rank_i} \tag{9}$$

where $|Q|$ is the number of questions, and $rank_i$ is the position of the user.

– Normalized Discounted Cumulative Gain (nDCG): A method which uses graded relevance as a measure of usefulness. Positions of users in the beginning of the list are more important then positions in the end of the list. The formula stands as follows:

$$nDCG = \frac{DCG_p}{IDCG_p} \tag{10}$$

where DCG_p is Discounted Cumulative Gain, and $IDCG_p$ is the ideal possible DCG – it is DCG of the ground truth, while DCG_p is Discounted Cumulative Gain of users sorted according a method being evaluated. We use alternative formulation of DCG:

$$DCG_p = \sum_{i=1}^{p} \frac{2^{rel_i} - 1}{log_2(i+1)} \tag{11}$$

where p is a rank position evaluated, rel_i is relevance of a user at a position i.

As our method works with question difficulty, which is based on time to answer a question, we can take into consideration only those questions that have at least one answer. Moreover, we evaluated the performance of all methods for only those questions which have at least two answerers with calculated reputation, so we could perform a comparison between the lists of users (users with unknown reputation were left out from the comparison). For this reasons, we report our results on almost 20 000 questions even though there are 33 052 questions in the 09/2014 dataset.

The evaluation was performed employing an experimental infrastructure, a part of CQA system Askalot [5] which is being developed at Faculty of Informatics and Information Technologies at Slovak University of Technology. The infrastructure enables us to reconstruct events as they happened in time, thus allows us to perform chronological evaluation process. Table 1 reports the results of our experiments. We present performance of Precision@1 (P@1), Precision@2 (P@2), Mean Reciprocal Rank (MRR) and Normalized Discounted Cumulative Gain (nDCG). The last column displays the number of questions which were evaluated.

In order to evaluate how individual components of the proposed method contribute to user reputation, we evaluated the method's performance in two steps. Firstly we worked only with reputation gained for answering questions (labeled as *Answers only*). Secondly we employed also reputation for asking questions (i.e. full variant of the proposed method). We also tried two configurations of our method in order to completely eliminate activity factor (Formula 2 labeled as *average*), and Formula 1 labeled as *sum* in the results.

The results show that our method outperformed all the baseline methods. The interesting observation is, however, the fact that the variant which completely eliminates user activity performed as the best. This suggests that we should focus only on the quality of user contributions, as even a small factor of activity can degrade precision of reputation estimation. In addition, we observe little difference between full and partial variants of our method. The full variant, however, reflects user reputation better because it captures reputation gained from answering as well as asking questions.

Table 1. Comparison of the performance of the methods.

	P@1 (%)	P@2 (%)	MRR (%)	nDCG (%)	Questions
Full variant *(sum)*	41.256	38.744	64.319	83.592	19861
Full variant *(average)*	**43.829**	40.801	**66.271**	84.533	19861
Answers only *(sum)*	41.331	38.916	64.428	83.662	19651
Answers only *(average)*	43.677	**40.842**	66.232	**84.548**	19651
StackOverflow Reputation	40.360	38.218	63.707	83.239	19866
Best Answer Ratio	**41.936**	**40.115**	**64.571**	**83.698**	19651
Z-score	38.019	36.703	62.062	82.394	19866
Number of answers	38.176	36.998	62.204	82.500	19651

5 Conclusions

In this paper, we presented a method for estimating user reputation in CQA systems. Our main goal was to strengthen the importance of quality of user's contributions when calculating reputation. It is done by employing question difficulty and utility of questions and answers. The performance of our method was compared with other feature-based baseline approaches. Our method outperformed all the baselines, and thus we can confirm our assumption that consideration of content quality (user expertise) plays an important role in estimation of user reputation. For future work we plan to apply our method for reputation estimation in Askalot and display reputation in a friendly form to its users.

Acknowledgement: This work was partially supported by the Cultural and Educational Grant Agency of the Slovak Republic, grant No. 009STU-4/2014.

References

[1] Hanrahan, B.V., Convertino, G., Nelson, L.: Modeling Problem Difficulty and Expertise in Stackoverflow. In: *Proc. of the ACM 2012 Conference on Computer Supported Cooperative Work Companion*. CSCW '12, New York, NY, USA, ACM, 2012, pp. 91–94.

[2] Jurczyk, P., Agichtein, E.: Discovering Authorities in Question Answer Communities by Using Link Analysis. In: *Proc. of the Sixteenth ACM Conference on Conference on Information and Knowledge Management*. CIKM '07, New York, NY, USA, ACM, 2007, pp. 919–922.

[3] Liu, J., Song, Y.I., Lin, C.Y.: Competition-based User Expertise Score Estimation. In: *Proc. of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*. SIGIR '11, New York, NY, USA, ACM, 2011, pp. 425–434.

[4] Pal, A., Konstan, J.A.: Expert Identification in Community Question Answering: Exploring Question Selection Bias. In: *Proc. of the 19th ACM International Conference on Information and Knowledge Management*. CIKM '10, New York, NY, USA, ACM, 2010, pp. 1505–1508.

[5] Srba, I., Bielikova, M.: Askalot: Community Question Answering As a Means for Knowledge Sharing in an Educational Organization. In: *Proc. of the 18th ACM Conf. Comp. on Computer Supported Cooperative Work*. CSCW'15, New York, NY, USA, ACM, 2015, pp. 179–182.

[6] Yang, J., Tao, K., Bozzon, A., Houben, G.J.: Sparrows and Owls: Characterisation of Expert Behaviour in StackOverflow. In: *User Modeling, Adaptation, and Personalization*. Volume 8538 of *Lecture Notes in Computer Science*. Springer International Publishing, 2014, pp. 266–277.

[7] Zhang, J., Ackerman, M.S., Adamic, L.: Expertise Networks in Online Communities: Structure and Algorithms. In: *Proc. of the 16th International Conference on World Wide Web*. WWW '07, New York, NY, USA, ACM, 2007, pp. 221–230.

Game for Connection of Metadata and Sources in the Domain of Multimedia

Aleš MÄSIAR*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
alesmasiar@gmail.com

Abstract. In the present, gathering metadata is necessary because of its use for effective information processing on the Web. There are three particular approaches of metadata acquisition, which are expert work, crowd work and automated machine-based methods. In our work we focus on crowdsourcing, because it can produce high quantities of results, which are still in reasonable quality. In this paper we present a game with a purpose, which acquires metadata for video data already existing on the Web. The goal of the game for the player is to use provided search engine to find video, which is playing on the screen without any describing information displayed. Our goal is to make the game challenging and entertaining for the player, while producing quality metadata for the videos.

1 Introduction

Amount of information on the Web is growing rapidly every day, therefore it is important to have the information organized using metadata. Especially, when considering multimedia sharing services, it is desirable for the user to be able to search or browse through the media effectively. This is achieved by constructing metadata layer over the media.

There are many video sharing sites and some of them are based on sharing the video via mobile phone, where user usually does not have much time for properly describing his data. A lot of users also do not consider proper video description to be important. This can lead to videos being lost amongst others, because they are very hard to find. Therefore we decided to focus our method on providing metadata for short videos obtained from video sharing service Vine[1], where users share their videos from mobile devices and generally do not pay much attention to providing useful metadata.

There are several methods, which try to describe the media better. Approaches can be divided to expert work, crowd work and machine-based methods [1]. Expert work means hiring one or more experts in desired field, who will then provide metadata in exchange for reward or wage. This approach produces high quality metadata, but due to the financial dependence,

* Bachelor degree study programme in field: Informatics
Supervisor: Dr. Jakub Šimko, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

[1] www.vine.co

it cannot produce high quantity of results. On the other hand, the machine-based methods can obtain high amount of data, but there is no guarantee of quality.

Crowdsourcing is important part of metadata acquisition using the work of large groups of people. With the possibilities of Web, it can produce high quantities of results, which are in sufficient quality. The most interesting technique for metadata acquisition in the field of crowdsourcing is human-based computation game or game with a purpose. The main principle of games with a purpose is that they can be freely played by large group of people and as a by-product, they produce metadata.

Games with a purpose are mostly very unique and the used methods are different for every area of use. Many games have been made for image and audio data description, however there is lack of games focused on gathering data for videos, so this field is open for new ideas.

2 Related work

Several different methods have been invented in field of automated video description. An interesting one of them is video tagging by mining user search behaviour [2]. This method analyses users' click-through data, when they are browsing video hosting website. Based on this analysis, it is possible to propose similarity between two videos and then mathematically assume tag assignment. This way, it is possible to get sufficient results, but this method works only when there is already some metadata layer created and it is based only on assumption that users always click on related videos, which is not definitive.

In the field of crowdsourcing, ESP Game [3] was first and one of the most famous games with a purpose created by Louis von Ahn. This game was used by Google for annotation of images, where it gained great popularity and successfully obtained lot of quality metadata. The game was multi-player and it was based on mutual output agreement. Two players were given the same image and with no possible means of communication, they had to agree on a word. The best strategy to win the game for both players was to write the words describing the image. This method worked very well, because it became popular, but in similar multi-player games there is always the risk of cold start problem. This means that when the game needs more than one player, especially in the early stages of deployment, it can be difficult to find an opponent and therefore the game cannot start.

Single-player games are generally harder to create, because it can be difficult to check correctness of player output. In these cases, metadata are usually acquired indirectly from player's behaviour in the game. Very interesting single-player approach is used in the game Akinator [4]. In this game player thinks of famous character or person and then computer in role of web genie is asking the player yes-or-no questions, trying to guess the name. In the end, the Akinator asks, if the guess is correct or not. In the case of incorrect guess, player writes the name and the game then processes the new data. During the game, it also asks some new questions, so when the guess is correct, it can still acquire new data. This approach is very innovative, but it requires very good starting data and lot of played games to give satisfactory results.

Filmillion[2] is similar game based on same principle as Akinator, but it is focused on movies. However it does not work as well, probably because of poor starter data set and lesser popularity.

Little search game [5] is a single-player game, which interestingly uses search engine for indirect validation of player output. At the beginning of the game, the player is given a term and his task is to lower the number of results returned by the search engine. This can be achieved by entering terms that have high co-occurrences with the specified term on the Web. The game is then able to construct term relationship network from acquired game logs. This approach satisfyingly reduces the cold start problem, because the only artefact used for output validation is result returned from search engine, which is independent from the game. Using search engine

[2] www.filmillion.com

for validation is innovative and interesting way of constructing games with a purpose and we are also using modification of this approach in our method.

3 Method overview

We present a game with a purpose, which is used for obtaining keywords for video data on the Web. Our game specifies on videos obtained from video sharing service Vine, where users upload six-second-long videos mostly from mobile devices. The game output is one or more keywords relevantly describing the corresponding video.

As shown in Figure 1, the game consists of three main parts:

- video, that is being played,
- search engine and
- results returned from the search engine.

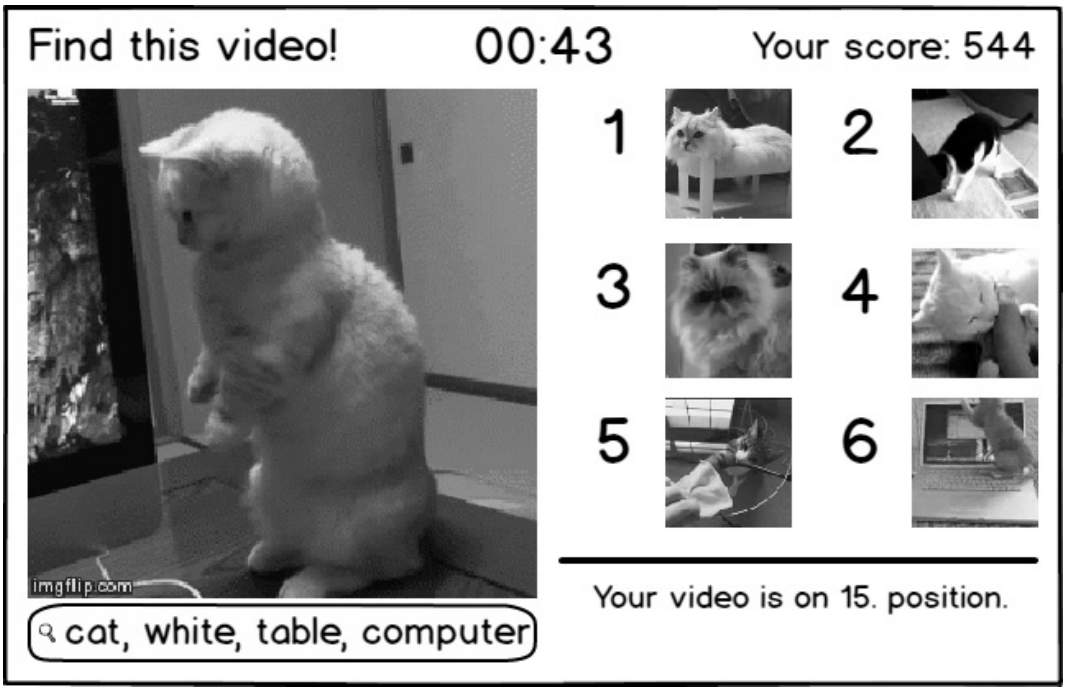


Figure 1. Graphical user interface of our game.

Amongst these, search engine is the most vital part, because that is what generates action in the game.

The game mechanics from player's point of view is based on searching the video that is playing on the screen using the search engine. Player is given no other information than the video itself. The best way to find the video then is to use keywords that he considers to be relevant to what he sees. Although, the keywords usually are not used in the source video description, the player's first search queries are actually the ones we are interested in.

Search engine returns videos in order relevant to player's search query. If the search engine returns searched video on a position lower than selected threshold, player wins the game and receives points. If not, player is given as many tries as he needs to move the video upwards

in the ladder. During the game player is helped by showing the video position in the overall ladder, but he is showed only videos up to threshold position.

Overall, our method can be described as follows:

1. Player is shown the video.
2. Player types query consisting of one or more keywords into the search engine.
3. Depending on the results of query, player tries to refine the query to move the video higher in the ladder, if needed.
4. After the game, keywords are extracted from player's queries.

Actual finding of the video is possible, because of the data that is acquired from the website with the video, which is then processed by full text search engine. We assume that some words relevant to the video are already mentioned in the description, but they can be hidden in the context or there is just lack of useful information.

Challenge of the game is strengthened by using time limit for one game. This forces player to concentrate on the video and quickly focus on the most important points, which he immediately tries to describe using keywords. This is great advantage against automated methods, because humans can not only quickly and with great precision describe what they see, but they can also consider the context and identify actions happening in the video.

After the game all game logs are saved to the database for later use and analysis. The most important log data are the search queries, which contain the desired keywords. The other log data include order of the search queries, results of the specific query that are shown to player, game score, duration of the game or the count of how many times player played the video. Considering these logs we can later improve our metadata extraction method or the game mechanics.

Video annotations are obtained from the search queries. If one keyword is mentioned specified number of times, it will be considered to be relevant to the video and will be used in the output data of the game. It is also interesting to focus on behaviour of single player in a specific game, assuming, that the first one or two queries in the game will be the best describing, we can then consider them to be more significant.

Source videos for the game are acquired from the Web only with basic describing data that consists from name of the video, user description, user tags and video comments. These data are most commonly present with the videos and in most cases they contain some relevant information.

4 Evaluation

Hypotheses and questions. We have set up the following hypothesis:

- Our method is able to acquire relevant and correct keywords for videos, which in the beginning contain only basic description.

We also aim to answer following research questions:

- How many of the acquired metadata represent new keywords and how many of them confirm or filter already existing metadata.
- How specific are the acquired metadata?

To confirm our hypothesis, we first plan to perform a closed experiment in a controlled environment.

Participants. Players of our game will be students or young people without any specified selection criteria. They will play the game voluntarily and in their free time. The latter evaluation will be performed by experts familiar with the domain.

Data. Source data for our experiment will consist of specified amount of videos acquired from the video sharing service Vine. Videos will be obtained and used only with the description and data that comes from the Vine App application programming interface.

Methodology. Players will become familiarized with the rules of the game and then they will play the game given amount of times. All the game information will be logged, as well as player behavior. After the game players will be asked questions about the game mechanics, entertainment and learnability, which will then help us understand the game from a player's perspective.

Process. After finishing all the games, the keyword extraction will be performed and the resulting keywords will be evaluated by experts.

In the next phase, we plan to perform a live experiment, where we deploy our game on the Web, where it can be publicly played. With this experiment, we can evaluate, whether our game is interesting and enjoyable enough to be played without any external motivation. After satisfactory amount of played games, we will perform data correctness evaluation as with the closed experiment.

5 Conclusions

Due to the constantly growing amount of data on the Web, the importance of metadata is still increasing. There are many services where person uploading the data does not feel the need to include metadata, which are sufficient for satisfactory video processing.

We presented a game with a purpose, which is able to acquire new keywords for video files. Our game is based on full-text search engine, which player uses to find videos playing on the screen. Keywords he uses for finding the video are then used for obtaining new metadata.

The advantage of our method is, that it is a single-player game, therefore it is not affected by cold start problem, which can happen with multi-player games. Also, the keyword metadata are many times used for efficient video searching and our game mechanics practically simulates this process, so we implicitly get exactly those annotations that user would use to search the specific video.

In the future we plan to perform the experiments and depending on the results adjust our method to be more efficient. It would also be interesting to improve our method to be more universal and able to perform well also on other types of data.

Acknowledgement: This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG 1/0646/15.

References

[1] Šimko, J., Bieliková, M.: *Semantic Acquisition Games*. Springer International Publishing, (2014).

[2] Yao, T., Mei, T., Ngo, C., et al.: Annotation for free. In: *Proceedings of the 21st ACM international conference on Multimedia – MM '13*, ACM Press, (2013), pp. 977–986.

[3] Ahn, L. von, Dabbish, L.: Labeling images with a computer game. In: *Proceedings of the 2004 conference on Human factors in computing systems – CHI '04*, ACM Press, (2004), pp. 319–326.

[4] Šimko, J., Bieliková, M.: State-of-the-Art: Semantics Acquisition Games. In: *Semantic Acquisition Games SE – 3*, Springer International Publishing, (2014), pp. 35–50.

[5] Šimko, J., Bieliková, M.: Little Search Game: Lightweight Domain Modeling. In: *Semantic Acquisition Games SE – 4*, Springer International Publishing, (2014), pp. 51–65.

Question Routing Based on Non-QA Data in Community Question Answering

Marek GRZNÁR*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
marek.grznar@gmail.com

Abstract. With the development of Web 2.0, there is a novel option to obtain required information by asking a community in Community Question Answering (CQA) systems, such as Yahoo! Answers and Stack Overflow. The existing CQA systems, despite of their increasing popularity, fail to answer a significant number of questions in required time. One option for supporting cooperation in CQA systems, is a recommendation of questions to users who are suitable candidates for providing correct answers (so called question routing). Various methods have been proposed so far to find answerers for questions in CQA systems, but almost all studies heavily depend on previous users' activities in the system (QA-data). In our work, we focus on utilizing users' context as a way of support question routing. We proposed a question routing method which analyses users' non-QA activities from external services and platforms, such as blogs, micro-blogs, social networks, in order to better identify suitable users for answering new questions. This solution allows us to involve also users with no or minimal previous activity in the system.

1 Introduction and motivation

The current web systems use a variety of search engines to offer theirs users the ability to effectively identify and obtain valuable information. Despite the fact that search engines in the last decade have significantly improved their efficiency and effectiveness, there are still some aspects that search engines cannot meet. This is due to several reasons. Often it happens that search engines fail because the provided results are not sufficient and necessary information is fragmented among several sources and therefore obtaining the full answers can be very challenging. In addition in many cases, required information cannot be simply described as a search query by keywords that search engines know to find.

Web 2.0 offers an opportunity that helps to solve the mentioned problems. This option is to obtain necessary information by asking a community [1]. Such information systems typically include various web forums. However, in the recent years, Community Questions Answering

* Master degree study programme in field: Information Systems
Supervisor: Ivan Srba, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 157–162.

(CQA) more and more come to the fore. Popular CQA systems include Stack Overflow, Yahoo! Answers or Quora, to name a few.

CQA systems have recently become an increasingly popular place not only for knowledge seeking, but also knowledge sharing. The user can insert a question into system and other users by answering this question can share their knowledge (e.g. [5]). In comparison with conventional systems which receive the information, CQA systems are based on a community, which constitutes a form of social network. Searching for information using wisdom of a crowd can provide the most accurate answers. Thereby CQA systems also use many social features (e.g. following users, commenting answers).

Currently in the CQA domain, one of the most serious problems is that CQA systems are quite often failing to answer questions in significant time. Authors in [2] reviewed the newly published 3000 questions in Yahoo! Answers and found out that only 17.6 % of questions received satisfactory answer within 48 hours. Authors in [3], two years later, similarly explored data from the same system of Yahoo! Answers. They discovered that only 11.95 % of the questions were answered in one day and only 19.95 % of the questions were answered in two days. From the results of these studies, we can observe ongoing seriousness of the problem that CQA systems cannot solve users' questions efficiently.

2 Related works

Currently, CQA systems represent a very interesting field of research in the domain of informational systems as a number of open problems has already become a subject of previous research studies. In the CQA systems, different adaptive approaches have been proposed to support the process of question answering. These approaches largely deal with four life-cycle phases of questions.

Question searching brings together approaches that try to identify in the already answered questions high-quality answers that would provide the same knowledge as is required to answer a new question which a user put into the system.

Question creating main goal is to classify a newly created question and assign a correct topic (category) to which the question relates.

Question answering. Approaches which are included in question answering and concern with the answer ranking.

Question routing includes a group of approaches that aim to a recommendation of questions to users that would be able to provide answers on new questions posted to the system.

From the above groups of approaches, we focus on question routing, because this group of approaches has the best chance to solve the problem of low number of answered questions. There are many existing research approaches concerning with question routing. Authors in [6] proposed a segmented topic model for the representation of user expertise. Based on user history in the system, Rhiahi et al. built a user model and matched it with newly posted questions. Their results were compared with TF-IDF and LDA. They found out that although LDA achieved better results than TF-IDF but segmented topic model has the best results.

Different user properties are considered in study [2]. Except user expertise, they take into consideration also user activity. First, they built a user profile from previous answered questions. Then they use the user activity as a measure of his availability in the system. They aim was to estimate whether the user log on into the system in several days after the routed question is posted. From the results, they find out that user availability can also boost the performance of question routing.

There are still large gaps and drawbacks in the existing approaches for question routing. Almost all approaches rely only on data from CQA system (answers, questions etc.). When a new user sign up in the system, we have no QA data about this user (cold start problem). Besides cold start problem, there are a lot of users called lurkers who would be good candidates to answer questions, but as they have only minimal interaction with system, system do not have sufficient

information about their expertise and therefore it cannot route any question to them. Authors in [9] report that only 15.67 % of all users answered more than four questions. These sparse data may lead to fail to achieve satisfactory results of question routing.

Based on successful results achieved in [7] and [8], we decided to focus on non-QA data in our work. By taking into account the results of these researches, we believe that the use of non-QA data can help with the question routing issues (i.e. cold start problem, lurkers, sparse data). More specifically, we suppose that non-QA data will provide us with correct information for routing questions to users about which we have little or no information.

3 Method for question routing using non-QA data

The whole process of question routing is illustrated in Figure 1. Our solution consists of four main processes. First we extract and process all available non-QA data for all users in system (Step 1).

Except of the user non-QA context, we use also historical data from the previous question answering process. Then we create user profiles from obtained information. Each potential answerer in the system is represented by two profiles (non-QA and QA) (Step 3). For a newly posted question, we also create a question profile (Step 2). Finally when all profiles are created, we rank all the answerers based on the similarity of their user profiles with the question profile of the question to be routed (Step 4).

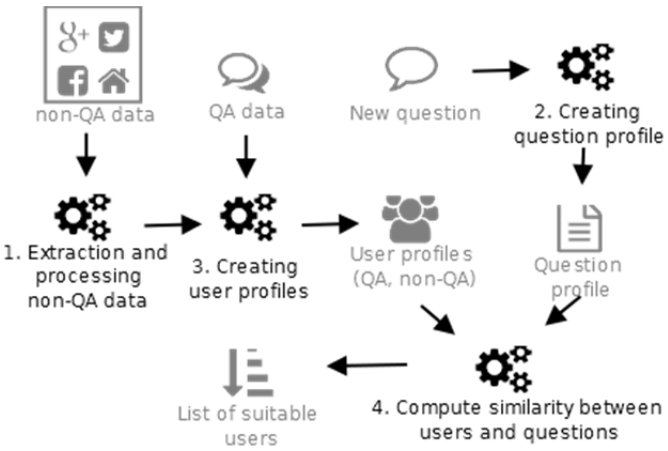


Figure 1. The framework of question routing.

3.1 Extraction and processing non-QA data

Non-QA data: For successful operation of the method, primary non-QA data about the user are needed. All non-QA data can be divided into external and internal, the ones that come from the CQA system, but are not related to the question answering process itself and those that do not come from the CQA system including a variety of information about the user publicly spread through external systems and social networks. Together, these data form the context of the user.

Number of sources of external non-QA data is quite large. The actual social networks (Facebook, Twitter, LinkedIn, etc.) provide a great amount of information about the users. If we consider only non-QA data on the social network Facebook, we can learn what the user likes and what interests him (pages that has been marked as "like"), or we may receive information about his/her expertise directly from his profile (which schools he/she attended, where he/she works). However, we can also obtain information indirectly from people that he/she has added as a friend. Another increasingly common phenomenon are home pages and blogs, where users write short articles, for example from their professional life.

Internal non-QA data include personal profile data. The current CQA systems commonly require a registration for user interactions with the system, during which users should enter a few details about themselves. However, non-QA data are not only explicitly specified profile information by user but also information related to the existence of the account, such as a number of days from its creation, an average time spent in the system etc.

Extraction and processing of non-QA data consists of three steps: (1) obtaining user context from external sources using web crawler; (2) content identification; (3) and finally, text analysis.

3.2 Creating question profile

We represent each question by a question profile, which is one of the building blocks in our method. When the question is posted, the amount of available information is limited. Thus, when creating the question profile, we only consider the question tags and the question textual content (i.e. title and body of the question). We ignore any additional question data that may be later added, for example comments to the question. Therefore, the question profile is built only once at the time when the question is added to system. The main motivation for utilizing only the referred kind of data source is that we do not need any updates of the question profile later. This presents clear benefits in terms of scalability [4] .

Before building the question profile, we make one document from all data by concatenation of title, body and tags of the question. Secondly, we apply some basic text processing methods on this document (tokenization, lemmatization and stop-word removal). From the processed document, we built the question profile using latent Dirichlet allocation (LDA). The question profile is represented by identifier of topic, topic probability (represents probability that latent topic is related to question) and word vector (word vector describes words distribution within topic).

3.3 Creating user profiles

Every user is represented by two user profiles. The first one is QA user profile which captures the user interest form QA data. The second one is non-QA user profile which use non-QA data to find user interests.

User profile from QA data: In most question recommendation methods, users are represented by their interactions with the questions they answered or asked in the past. These interactions include answers which user posts, question which user asks etc. Like [4], we also derived QA user profiles from question profiles. Each question, with which the user interacted, represents a separate document. The final QA user profile is represented like average of the derived QA profiles belongs to user.

User profile from non-QA data: In this part, we introduce how we can build a non-QA user profile from the processed source of non-QA data. In Section 3.1, we introduced how to get relevant data from various sources of non-QA data. From these processed data, we create one document for each source. We build a partial non-QA user profile for each document. When creating the final non-QA profile, it is necessary to take into account the diversity of non-QA data sources. Therefore each partial non-QA profile must has its own weight. The final non-QA user profile is computed as a linear combination of individual non-QA profiles.

3.4 Compute similarity between users and questions

The core idea behind the question routing is to return to a new question a list of suitable users ranked by a relevance score, which is calculated for the pair (question profile, user profile). Measuring similarity between users and questions runs in two phases:

1. Computing similarity between QA user profile and question profile
2. Computing similarity between non-QA user profile and question profile.

Considering to representation of profiles, first we need to flatten both question and user profiles into two-dimensional vectors. After that, we compute a cosine similarity between these vectors:

$$sim(P_{question}, P_{user}) = \frac{P_{question} \cdot P_{user}}{\|P_{question}\| \|P_{user}\|} \quad (1)$$

where P_{user} is the user profile and $P_{question}$ is the question profile. Finally, the resulting values of similarity between user and question are obtained by a linear combination of QA and non-QA similarity (Equation 2). From the obtained overall similarities, a ranked list is created where the first record will represent the most similar user to the given question.

$$sim(Q, P) = \alpha * sim(P_{question}, P_{QA}) + (1 - \alpha) * sim(P_{question}, P_{non-QA}) \quad (2)$$

4 Evaluation

The experimental dataset is collected from a community-based answering site Android Enthusiasts. It contains data from May 2009 to December 2014. During this period, about 26,000 questions and 33,000 answers were posted. From the total number of questions, about 18,900 questions contains at least one answer.

Given the diversity of sources of non-QA data and on the basis of an analysis of the experimental dataset, we decided to choose three the most numerous sources: About me, Homepage, Twitter (see Table 1).

Table 1. Distribution of non-QA data sources in the dataset.

Number of users	Source of non-QA data
525	About me only
236	Homepage only
27	Twitter only
10 073	About me, Homepage
275	About me, Twitter
726	Homepage, Twitter
10 668	About me, Homepage, Twitter

Equation 3 show how we compute final non-QA user profile, where P are partial user profiles and $\alpha + \beta + \gamma = 1$.

$$P_{non-QA}(u) = \alpha * P_{AboutMe} + \beta * P_{HomePage} + \gamma * P_{Twitter} \quad (3)$$

We compare our method to a gold standard. The gold standard represents users who answered to questions in the system. However, we realize that our chosen gold standard does not completely correspond to the real interest from users about question. Interest from users does not have to be expressed only by answering the question but also by voting for answer, which was added to the system for the given question. However, due to the incompleteness of data in the selected dataset, we do not know to clearly identify who gave the vote for a specific answer. Therefore, we decided to use a set of users that provided the first answer to a question for comparison with the results from our method. We also evaluate the proposed methods on user who have a different amount of previous activities in the system (no activity – 0 answers by user in system, a few answers – 1–5 answers by user in system, a large amount of answers – more than 5 answers by user in system). The preliminary results show that using non-QA data can boost the performance of question routing.

In evaluation of our method, we used an experimental infrastructure which can simulate the events depending on time when they occurred. This infrastructure is implemented as a part of the CQA system Askalot. In addition, we have already implemented a Facebook crawler that collect non-QA data about users' friendships. It will allow us also to make also an online experiment with several participants.

5 Conclusion

In this paper, we proposed a new method which uses non-QA data for routing question. One of the goal of this method is routing question to users which are not so actively involved in interaction with a system (asking or answering questions). Using of different sources of non-QA data may help to engage users who are new or totally passive in the system (lurkers). It means that we can reduce a cold start problem because users spread a lot of information about themselves. The proposed method in addition to non-QA data also involves previous question-answering history which is widely used by the most of existing approaches. Our method should improve question answering process what will definitely contribute to successful CQA systems.

In the future work, we plan to finish evaluation of our method and experimentation with its parameters. To evaluate its overall successfulness, we will use different evaluation metrics to compare our method to other research approaches.

Acknowledgement: This work was partially supported by the Cultural and Educational Grant Agency of the Slovak Republic, grant No. 009STU-4/2014.

References

[1] Liu, Q., Agichtein, E., Dror, G., Maarek, Y., Szpektor, I.: When web search fails, searchers become askers. In: *Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval – SIGIR '12*, ACM Press, (2012), pp. 801–810.

[2] Li, B., King, I.: Routing questions to appropriate answerers in community question answering services. In: *Proceedings of the 19th ACM international conference on Information and knowledge management – CIKM '10*, ACM Press, (2010), pp. 1585–1588.

[3] Zhou, T. C., Lyu, M. R., King, I.: A classification-based approach to question routing in community question answering. In: *Proceedings of the 21st international conference companion on World Wide Web – WWW '12 Companion*, ACM Press, (2012), pp. 783–790.

[4] Szpektor, I., Maarek, Y., Pelleg, D.: When Relevance is not Enough: Promoting Diversity and Freshness in Personalized Question Recommendation. In: *Proceedings of the 22nd international conference on World Wide Web – WWW '13*, International World Wide Web Conferences Steering Committee, (2013), pp. 1249–1260.

[5] Xu, F., Ji, Z., Wang, B.: Dual role model for question recommendation in community question answering. In: *Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval – SIGIR '12*, ACM Press, (2012), pp. 771–779.

[6] Riahi, F., Zolaktaf, Z., Shafiei, M., Milios, E.: Finding expert users in community question answering. In: *Proceedings of the 21st international conference companion on World Wide Web - WWW '12 Companion*, ACM Press, (2012), pp. 791–798

[7] Pan, Y., Luo, L., Chi, C., Liao, Q.: To answer or not. In: *Proceedings of the 2013 conference on Computer supported cooperative work – CSCW '13*, ACM Press, (2013), pp. 1253–1262.

[8] Liu, Z., Jansen, B. J.: Predicting potential responders in social Q&A based on non-QA features. In: *Proceedings of the extended abstracts of the 32nd annual ACM conference on Human factors in computing systems – CHI EA '14*, ACM Press, (2014), pp. 2131–2136.

[9] Zhou, G., Liu, K., Zhao, J.: Joint relevance and answer quality learning for question routing in community QA. In: *Proceedings of the 21st ACM international conference on Information and knowledge management – CIKM '12*, ACM Press, (2012), pp. 1492–1496.

Analyzing User Gaze on the Web

Patrik HLAVÁČ*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
patrik@hlavac.sk

Abstract. In this paper, we propose a method of reading detection from gaze data. Eye tracking devices provide irreplaceable information about a user's gaze. This work deals with the possibilities of identifying user interaction in the educational system. Our algorithm takes into account user's fixation data and maps their coordinates onto single word elements. These are then processed with respect to their relative word distance. Rule-based solution works by considering the sequences in the order of their occurrence. Unlike studies that calculate distance in points that eyes moved around the screen, we consider the distance of words in the vector.

1 Introduction

Automatic analysis of user behavior on the Web and interaction with a web browser via computer is a nontrivial matter, where new solutions have recently opened with better availability of technologies and devices for tracking the user, such as eye tracking.

Whilst existing solutions are based on tracking user behavior using peripheral devices such as mouse and keyboard [1], actual options offer us direct tracking of user gaze and focus block of content on the website. Thus, it can improve the accuracy of the information in the user model.

An important part when working with web education system is the information, about the extent with which the user is interacting with the environment. We can distinguish the following conditions:

- whether the user is present,
- whether the user is watching,
- whether the user is reading,
- whether the user is learning.

These four defined tasks relate to each other and are gradually shifting the level of complexity to the problem of detecting reading and determine the status of learning, which are still at a low level of investment research in the world. The aim of this work is to provide a method for collecting and evaluating implicit feedback to detect reading, potentially to the level of detection if the user is learning. The solution is based on proposing a rule-based algorithm for reading detection.

* Master degree study programme in field: Information Systems
Supervisor: Dr. Marián Šimko, Institute of Applied Informatics, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 163–168.

2 Related Work

Despite the fact that technology of eye tracking is already known in the research and used for many years [5], approaches for the detection of reading the text appeared just at the beginning of the new millennium. Current approaches include tracking of available interaction with a web browser, such as the movement of the mouse, clicking and scrolling, from which they derive the behavior of the user's gaze. One of the relevant works is a comparison of correlation mouse movement and perspective when interacting on the page [4]. The result was 78.7 % correct determination of the read paragraph of text using the interaction with the browser.

One of the first uses of reading detection was tracking the reader's interest in levels: high, medium and low, with the recorded read word and according to them were adapted by searching the locations of the local files [3]. Number of readings of the word suggests its importance in understanding the text. Tokens that correspond to individual actions were defined during detection, such as "short move to the right", "long move higher," etc. Due to gaze movement and direction, where each token has appropriate value of points, the performed actions were rated and the amount of the accumulated points determined whether it is read or only scrolled through the text. Compared results of detection rate showed that this algorithm "pooled evidence" estimated reading in 99 % of cases faster than the original, matching algorithm, which takes into account only fixations to the right. Thus, we can conclude that the higher number of gaze features allow to define more precisely the way we interact with the text.

Another proposed method of detection was reading by modeling the sequences of fixations and saccades with Hidden Markov Model (HMM) [6]. It proposes HMM without linking to regions of interest (ROI). Datasets gained from interaction were labeled by expert according to whether they represent a sequence read (base was again three fixation) or not. This method is fast enough to process during the interaction, but in offline processing it produces better results. With the results of the accuracy = 0.8 and precision = 0.8 using directly during online interactions compared with the 0.88 / 0.87, obtained by additional offline processing.

One of the newest methods applied also in the additional software for equipment Tobii is reading detection method independent on content called "A Robust Reading Realtime-Skimming Classifier" [2]. The detection is performed on the basis of the characteristics observed saccades of the eye. It enriches the original method of spatial perception problem with the mapping of the text lying under a layer coordinates and further classification [3]. It compares the information about the speed and direction of eye saccade. It determines whether it is a reading comprehension, or passing the text by using SVM classifier. Author informs about the difference in the detection of training data as it is an aligned text area, or on the contrary, well-structured text displayed in columns and bullets, which is not applicable pattern reading by the horizontal direction. Provided accuracy results of 83 % are attributed to samples with smooth, clear identification. For each solution was present preprocessing, always above the raw data obtained from devices, being used primarily for grouping fixation, verification of the mapping words onto the background, or discarding the outliers [2].

Experiments were carried out in a prepared environment, often free of interfering elements (side menu bar), focusing on the perception of the text. Sometimes it is concerning about prepared articles shown in rectangular shape on a white background [3]; other times about selection of the area of the page labeled as area of interest [6]. Therefore, applying the algorithm to the web of the education system should have different results and different application possibilities.

3 Our method for detecting genuinely read parts of text

We record user interaction (especially gaze) focused on obtaining accurate coordinates. Specialized equipment called “eye tracker” stores the obtained data in a relational database or file. It is also necessary to obtain information from the system, especially the position of individual words, or parent elements, such as HTML tags, with which we pair obtained coordinates.

Assuming that we get the right words, clusters of words that a user saw, we also consider different ways of passing the text: reading, searching, skimming. These must be included in the process of recording information in the interaction. Along with information of interaction with the words we create a vector model for user content within a domain. By viewing the text as a continuous vector of words, which the reader is trying to understand, we can still analyze fixation, eye saccades and their derivatives, and it allows us to separate from the spatial information (note: of course, the spatial information will be used, but only to the state of element identification).

3.1 Document Reading Model

We use information captured from user interaction in the system to build a model. Through the available devices for gaze monitoring, we accurately capture the process of behavior, meaning the time and type of interaction on the Web and also deduce some basic features and user interest in the displayed information.

The user model can be treated as a model composed of two layers. The bottom layer (Evidence Layer) maintains the information obtained in the process of collecting data. It contains data from different sources, such as the basic document properties (title, link) and the feedback from eye tracker. Raw data from the eye tracker contain information about the time stamp, positions of the eyes and coordinates of the center point of view. This information is stored in a predefined structure. Collected data are processed in the upper layer (Inference Layer). Document Reading Model represents all instances of scanned words with records of properties and characteristics of the current session. These contain the calculated value of the reading interest for single words along with the total number of fixations and length of viewing words. The obtained information is used to analyze user's reading and possible adaptation.

Table 1. Specification of the Document Reading Model.

Word position	Content	String length	No. of fixations	Fixations duration	Reading Interest	... +
...						
40	zaujíma	7	2	0.35	1.10	
41	sa	2	2	0.28	1.50	
42	o	1	1	0.15	0.70	
43	rozpočet	8	3	1.13	1.90	
44	sprostredkovateľ	16	5	0.82	2.30	
45	medzi	5	4	0.57	2.20	
46	manažérmi	9	3	0.60	2.00	
...						

3.2 Algorithm

The basic assumption of reading detection at the level of words is the ability to detect words as unique elements of the code. By pre-processing the source code before the detection we can change the structure of the source code so that each word represents a separate element with a unique label. With this modification we get a number of ROI, which are approximate of the spatial size of a given word. Regions of individual words are modified to individual and unique elements.

A lot of research was undertaken at the spatial level below what we understand the algorithm that was based on gaze motion recorded in field of screen coordinates. Our proposed solution takes into account the measurement and evaluation of data from the gaze interaction directly to the content on the level of words, so we are going to talk about the content plane.

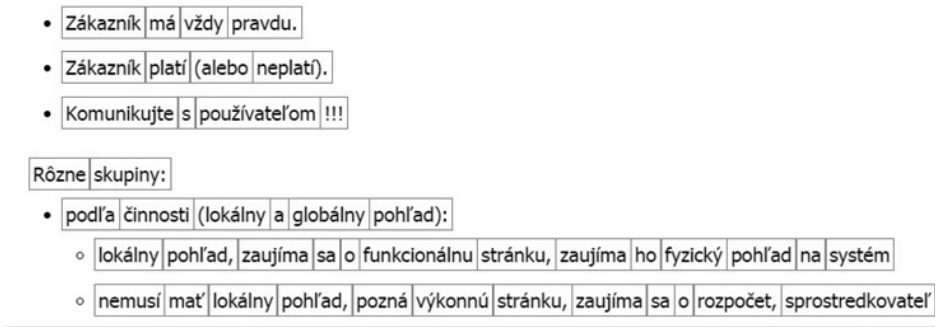


Figure 1. Visualization of document content after converting text into regions of interest.

By moving the spatial plane to the plane of the content represented by a vector, we get a different view on the distance of words on which reader's gaze was fixed. Unlike studies that calculated distance in pixels / points that eye moved around the screen, we are considering the distance of words in the vector. Each word (the word is considered string separated by spaces or end of line signs) is stored in its original sequence in the list with other attributes for each word, creating a model.

Basic, rule-based algorithm is based on the time point of view, so a solution consists of order of sequences in the order of their occurrence. We consider the following variable of reading interests (RI), which maintain the current state of reading intensity. This variable will take values greater than zero. The sequence of fixation determines the variable growth, or penalty. While reading nearby words, RI is increasing and marking passed words with actual RI value. During next passing, the word value of RI and actual RI are summed for this word.

Variable storing Updating value (UV) indicates how the reading interest will be increased due to activity among words. The base value will be further selected according to results from experiments, but it is based on the level of words approach, so it is subtracting the position index of the current word and the previous word in the vector, to determine the specific distance between words.

During initializing variables we consider also the complexity of the content. The more complex the text is to understand, the longer the processing time and reading is. Thus we introduce a variable that reflects the complexity of the text - that limits what the system will consider as read. There are number of ways, how to determine the complexity of the text. Quick and easy way represents the Flesch-Kincaid Reading Level formula, which determines the complexity of the text based on the length and number of words.

Since reading is a sequence of successive fixations, the proposed algorithm is based on increasing the variable of user interest. Based on the pre-processed data from the user's gaze, the cycle passes sequentially recorded data and according to their characteristics – mapping position coordinate to elements in the background and assigns them the weight with which the user has watched them. After weighting, a first simple metric e.g. "words per minute" could be applied.

4 Evaluation

Using rule-based algorithm we detect the reading in user sessions during the experiments. To obtain information about document we use the programmable features of web browsers, the current solution is *Google Chrome Extensions*. In order to evaluate our detection method, a user study with volunteers was made. Participants are on bachelor or master study with age from 18 to 24 years.

4.1 Data

To obtain a dataset of user sessions we chose eight articles published online. Half are written in encyclopedic style and come from the Wikipedia and the second half is written in popular science style containing news from the scientific environment. The shortest article contains 289 words and the longest 369 words; the average article has 313 words. Articles are not strictly focused on one topic, but include information from everyday life, such as sports, animals, actors, space or internet. We gathered these articles on one page containing horizontal menu on the top of a window. Articles are drawn in paragraphs with or without title, where from each article we deduce five issues mapping the content of paragraphs.

We use Tobii TX300 Eye Tracker with Tobii Studio software for collecting gaze data. This device provides us with three hundred coordinates per second, which results in approx. 850 000 data rows per user session. Every row contains additional information about timestamp, media title, whether it is fixation, saccade or none and detailed coordinates information for both eyes.

4.2 Experimental Design

Based on data from the eye tracker we apply the model, which will be verified explicitly by evaluation questionnaire for each user. Sitting with each user includes solving of the tasks, once prepared and uniformly interpreted to all participants. The participant is asked to read the articles prepared in the specified order with the intention of learning the parts of his choice. Then, after reading the article he answers five questions related to clarification whether he knows the answer or whether he at least saw information. We set up experiment in Tobii Studio with webpage articles content in rotation with instruction sheets, so we have made 16 screens (8 for reading and 8 non-disturbing while respond questionnaire). User session took slightly less than one hour.

We have prepared experimental infrastructure for whole experiment. At the moment we have gathered information from 15 user sessions. About 16 millions of coordinates will after filtration reduce to half.

5 Conclusions

This paper provides insight into the current state of research and using of the eye tracker to detect reading. Detecting genuine reading is a challenge and a prerequisite for accurate modelling of the user, e.g. in adaptive systems. We propose a new method of detection based on eye fixation mapped to specific words in a document. The main contribution consists of providing a method which not only shows where the user is looking, but detects the reading or the degree of accuracy of learning.

Acknowledgement: This contribution was created with the support of the Ministry of Education, Science, Research and Sport of the Slovak Republic within the Research and Development Operational Programme for the project University Science Park of STU Bratislava, ITMS 26240220084, co-funded by the ERDF.

References

[1] Atterer, R., Schmidt, A.: Tracking the interaction of users with AJAX applications for usability testing. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, (2007). pp. 1347–1350.

[2] Biedert, R., et al.: A robust realtime reading-skimming classifier. In: *Proceedings of the Symposium on Eye Tracking Research and Applications*, ACM, (2012), pp. 123–130.

[3] Campbell, C.S., Maglio, P. P.: A robust algorithm for reading detection. In: *Proceedings of the 2001 workshop on Perceptive user interfaces*, ACM, (2001), pp. 1–7.

[4] Hauger, D., Paramythis, A.; Weibelzahl, S.: Using browser interaction data to determine page reading behavior. In: *User Modeling, Adaption and Personalization*, LNCS 6787. Springer, Berlin, (2011), pp. 147–158.

[5] Jacob, R.J.K.: What you look at is what you get: eye movement-based interaction techniques. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, (1990), pp. 11–18.

[6] Kollmorgen, S., Holmqvist, K.: *Automatically detecting reading in eye tracking data.* Lund University Cognitive Studies, (2007).

Refining Interest of Information Seekers on the Web by Eye-Tracking Feedback and Groupisation

Jozef MARCIN*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
jozef.marcin@gmail.com

Abstract. This paper discusses refining of interest of an information seeker while searching the web. Revelation of interest, respectively, the intent with which the information seeker run a search can bring significant improvement in the searching process, but also in the browsing. Acquisition of the intention is a problem, which cannot be solved by explicit feedback only because browsing the Web takes place in real time. The intention must therefore be generated by means of implicit feedback preferably in the form of data from the eye tracker and mouse. Main part of this work is to propose a method, which through the data collected by eye – tracking device and through the groupisation will lead to the provision of more relevant results for the information seeker's searching and decrease the number of search query reformulations.

1 Introduction

The topic of this paper is mainly using implicit feedback – specifically the trace of information seeker's eye movements acquired by eye tracker device for solving the problem of refining the web seeker's interest during his searching session. Eye tracker provides a very effective feedback almost immediately, without having seekers irritated by answering them some questions or by having seekers do their job differently than it´s natural for them.

In the web search the reformulation of initial query occurs very often. Approximately 1/3 of all created queries are composed by reformulating the initial query [1]. This demonstrates the fact that initial relevance of the documents received is quite low and it is necessary to reformulate query. This phenomenon is a consequence of several factors, such as separating parts of query or web seeker ignorance. Our primary goal is to reduce the need for reformulation of the query and to increase the relevance of resulting documents. We are trying to maintain this by the data obtained from tracking the eye, but we use an explicit feedback too, in particular the groupisation. Although we use the words selection from the elements of the page and not a direct words extraction, we focus on using respectively creation of our own words ontology in Slovak and English.

* Master degree study programme in field: Information Systems
Supervisor: Professor Pavol Návrat, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 169–174.

Our solution is based on the assumption that if we know the web seeker's interest, we can offer him more relevant documents to his query. We focus on the selection of appropriate words in elements of search engine results page (hereinafter SERP) to which web seeker fixated his eyes. In such way we get a possibly strong clue on his interest. We can enhance the query. Consequently, more relevant documents can be found for him.

2 Related work

Ability of refining the interest plays key role in our work. Another important part of the analysis is analyzing of groupisation methods. Due to this fact we tried to analyze current related approaches to solving these problems. More works are focused on the seeker's work with SERP from the view of selection of documents that are relevant for him. One of the main findings was the time of fixation, which represents the minimal time the seeker is looking on relevant information and was experimentally defined as 200 – 300 milliseconds [2]. In [1] researchers used multiple methods involving eye tracker to find the seeker's interest. They tried to create a set of pairs – term and its weight. This way they created some kind of vector of intent. To achieve the weight of the term (1) they used functions below. Function (1) can by calculated by four different methods, specifically by functions (2), (3), (4), (5).

$$TermScore_m(tp) = \tag{1}$$

$$Look(t,p) * TF(t,p), m = MLT \tag{2}$$

$$\frac{Look(t,p)}{max_{t1 \epsilon p} Look(t_1,p)} * \frac{TF(t,p)}{max_{t1 \epsilon p} TF(t_1,p)}, m = nMLT \tag{3}$$

$$\frac{Look(t,p)}{TF(t,p)}, m = DLT \tag{4}$$

$$\frac{Look(t,p)}{max_{t1 \epsilon p} Look(t_1,p)} / \frac{TF(t,p)}{max_{t1 \epsilon p} TF(t_1,p)}, m = nDLT \tag{5}$$

In the above formulas, $Look(t,p)$ is the number of times seeker looks at term t at page p and $TF(t,p)$ is occurrence frequency of t on p. This method assumes that the terms that seeker looked at many times, are connected with seekers interest. The MLT (Multiply Look by Tf) method (2) is based on the idea that the more frequently the term appears, the more relevant it is within seeker's intent. The DLT (Look Divide by Tf) method (4) is based on the argument that even if the term occurs with low frequency, the term is relevant, if seeker looks often at it. nMLT (3) and nDLT (5) methods are standardized versions of MLT and DLT.

A separate part of this work was search the intent in segments – individual elements in DOM tree of the page.

In [4] researchers addressed methods of groupisation. They were looking for seekers with similar features to provide them more relevant links to their queries. Current trends in creating a groups are:

- similarity based on clicks on the link: determined by three ways: match URL clicked on, match domain of URL clicked on, consistency between the categories of topics of URL clicked on,
- syntactic similarity – the similarity of the querying (keywords),
- semantic similarity – even if the queries are not similar based on syntactic similarity, they can be similar by their meanings.

Another problematic area, which researchers are engaged in, is obtaining text by eye tracking. In [3], a number of methods for extracting text is described.

In [5], a possibility of using coincidence between eye fixation and mouse cursor position, considering the observed correlation between the two entities is discussed. Researchers discovered a correlation between looking and cursor position considering various behavioral patterns – reading, hesitation, scrolling, clicking.

3 Our method for refining interest

Our method primarily deals with creating a vector of interest by selection of appropriate words from elements of the page that the seeker is looking at during web search. Our hypothesis can be described as follows: If we can obtain and process the information on seeker´s gaze, we can provide better suggestions to seeker.

3.1 Web seeker's gaze acquisition

To obtain data on seeker's gaze, we use the eye tracker and software developed by a university project team "*Carrots*". The web service receives raw data from the eye tracker. The data includes normalized coordinates of eye position sampled many times in a second (sampling frequency depends on the kind of eye tracker device employed, e.g. 60 or 300 Hz), so that they can be used on any monitors resolution. Subsequently, the data is sent to Google Chrome extension, where it is enriched by several other values, the most fundamental being timestamp, URL and XPath.

3.2 Processing of data obtained from seeker's gaze

If we have available data from seeker's gaze, it is necessary to process it in order to obtain seeker's interest from it. This section is divided into following parts:

- Filtering of unnecessary data – There are some specific data particularly important for us, which must satisfies the conditions: time of gazing at element is greater than lowest possible fixation time, the data are receiving from pages of SERP that within google.com domain and subdomains, data of elements, that contains abstracts of document in SERP.
- Getting a specific element – The individual data records we collected contains XPath, with the help of which we can obtain element and its value.
- Mining the interesting text and creating a vector of interest – first we select keywords from web search query and create the initial vector of interest, which is represented by pair – word and value initialized to 1.

  $$V_z = \{(w_1, 1.0), (w_2, 1.0) \dots, (w_n, 1.0)\}$$

  where V_z is vector of interest and w_n is the n-th word of interest.
- Enrichment of vector of interest – next we need to compare each word from obtained element with words from initial vector V_z so we can next update the vector with interesting words obtained from element, that seeker is gazing on. For example if seeker interests in sport results, enriched V_z might look like:

  $$V_z = \{(sport, 1.0), (Nik\acute{e}, 0.85), (hockey\ results, 0.55), (TV\ program\ sport\ 0.11)\}$$

  where pair $(Nik\acute{e},\ 0.85)$ is built from keyword and value normalized by maximum and minimum time of fixation.

3.3 Determination of seeker's interest

Using vector of interest and TF-IDF analysis we create new vector, vector of search V_v, which holds number of word occurrence from V_z on the page. This is a first estimation of the relevance of documents. The resulting V_v should be for example:

$$V_v = (www.aktuality.sk), (sport\ 1.0), (17)$$

where individual elements in parentheses represent successively – the page on which frequency of word occurrence was searching for, word from V_z with value and frequency of occurrence of the word on page.

Subsequently, we enrich V_v outside the SERP website. It is a process in which for the vector of interest V_z we run a silent background search and helps to TF-IDF analysis we enrich V_v with further relevant documents.

The last step of this method phase is cross-over mutation of vector of interest and vector of search. First we order vector of search according to the number of occurrences of words from founded interest. Next, we regroup vector of search, so that resulting vector will be sorted by weights of words and by number of word occurrences. Subsequently, we use the resulting vector to display relevant documents to the seeker.

3.4 Groupisation method

Generally, grupisation is one of several possible approaches used to improve personalized web search [8].

In our work a groupisation is performed by using various methods, but still including the index of similarity. From options of creating groups we choose those:

- creating classes based on similarities in query (syntactic query similarity),
- creating classes based on similarities obtained intend (syntactic similarity of interests).

Using of groupisation is particularly important due to the fact, that seeker can search for a completely different things that was written in initial query. Because of this fact, refinement of seeker's interest with using only eye tracking can be limited. Thanks to groupisation we can solve situations like this by previous experiences. In context of our method, we use grupisation to provide results to the seeker, that other seekers noticed as relevant for their query before or during search session. We choose this results by methods mentioned above. By this results we can enrich the web search.

4 Experiments

Experiments in our work require a specific approach, because we have found no similar work so far that produces results that we could compare our results with. Experiments that we accomplished so far deal with the verification of our method to refine seeker's interest. Chapter experiment we therefore divide into two parts described below:

4.1 Experiments completed so far

In the current phase of implementation we performed the following experiments. In the first experiment we had a set of 30 queries. These queries and relevant documents were predefined. Set of queries was built by various specific tasks and was given to seeker to be performed. We were observing the necessity of query reformulation. Table 1 shows results that we achieved.

Table 1. Needed reformulation in web search using our system.

	Number of queries	% of relevant documents	% of irrelevant documents	% of reformulation needed
1	*30*	*64.6*	*35.4*	*26.66*

The Table 1 shows that current state of our work decreases need for reformulation of query from 33 % to approximately 27 % in this experiment.

In the next experiment we chose some ambiguous queries and declare the interest, with which seeker's search session started. We also identified, which documents will be relevant and which will not be. By using our methods of gaze tracking we have tried to reveal the real intention. Then we brought a set of relevant documents with numbers of words from revealed interest, which individual documents contained. In Table 2 we show individual queries and defined search interest together with search interest revealed by our method.

Table 2. Correctness of refining interest.

	Query	Intended interest	Interest found	Number of relevant documents	Number of irrelevant documents	Precision
1	*jaguar*	*Jaguar animal*	*jaguar american animal*	*8*	*2*	*80 %*
2	*musical instruments*	*bass guitar equipment*	*musical instruments equipment bass guitar*	*5*	*5*	*50 %*
3	*sport results*	*bookmaker pages*	*sport results Niké*	*7*	*3*	*70 %*
					Avg. precision	*66.67 %*

In Table 2, precision is counted as number of relevant documents divided by number of all documents.

In this experiment we did not use ontology only the WordNet dictionary, by which we selected words that in certain way was similar to primary query. We set the threshold for needed reformulation as 50% of precision, from which we made conclusion, that reformulation was not needed in either case.

4.2 Forthcoming experiments

After completing the ontologies and groupisation module we would like to perform experiments on a group of respondents as follows: using and without using our system, with and without using ontology, with and without using a groupisation module. Consequently, we would like to evaluate these approaches using these metrics:

- Accuracy – represents level of accuracy of suggested results from the seeker's point of view and seeker evaluates individual suggested links by himself.

$$Accuracy = \frac{\sum_1^n O_n * H_n}{O_v}$$

where O_n is an *nth* recommended link from set of offered links, $H_n$is an index of satisfaction with link O_n and has value from interval <0,1>, where 0 is no satisfaction and 1 is complete satisfaction, O_v represents set of all offered links.

- Rfactor – represents the number of query reformulations, that seeker has to execute during web search to obtain relevant results.

$$Rfactor = R$$

where $Rfactor$ is factor of reformulation necessity and R is number of reformulations until achieving relevant results.

- Satisfaction – represents seeker's overall level of satisfaction with results by time needed to obtain relevant results and by his evaluation of results satisfaction.

$$Satisfaction = \frac{1}{t} * s_i$$

where $Satisfaction$ is seeker's satisfaction with results, $\frac{1}{t}$ represents time level to achieve results, s_i represents seeker's satisfaction with results and is from interval <0, 1>.

All of these experiments will be performed as a group experiments. Firstly, we create a set of topics, which seekers will be looking for. This will tell us how our system works for specific queries. Then we let seekers to choose what they want to search for to find out how our system works for queries that seekers are creating by themselves.

5 Conclusions and future work

In this paper we described a method to refine seeker's interest in web search using implicit feedback. We focused mainly on eye tracking and groupisation. We have analyzed the problem area and proposed our own method for refining interest and provision relevant documents to web seekers. We chose the way of extraction possible words of interest from elements in the page of SERP. We have performed simple experiments that confirm our hypothesis in the current phase of implementation.

Further work will focus at creating an ontology with which we could better extract appropriate words. Another part of future work is completion of groupisation module, which will bring quicker reactions for similar sessions. Subsequently, we shall experiment to find out how our solution performs compared to other similar methods.

Acknowledgement: This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG 1/0752/14.

References

[1] Kazutishi, U., Takehiro, Y., Katsumi, T.: Search Intent Estimation from User`s Eye Movements for Supporting Infromation Seeking. In: *AVI´12*, (2012), pp. 349–356.

[2] Granka, A., L., Thorsten, J., Gay, G.: Eye-tracking Analysis of User Behavior in WWW Search. In: *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval*, (2004), pp. 478–479.

[3] Biedert, R., Dengel A., Käding, C.: Universal Eye-tracking Based Text Cursor Warping. In: *ETRA '12 Proceedings of the Symposium on Eye Tracking Research and Application*, (2012), pp. 361–364.

[4] White, R., W., Chu, W., Hassan, A., He, X., Song, Y., Wang, H.: Enhancing Personalized Search by Mining and Modeling Task Behavior. In: *WWW '13 Proceedings of the 22nd international conference on World Wide Web*, (2013), pp. 1411–1420.

[5] White, R., Huang, J., Buscher, G.: User See, User Point: Gaze and Cursor Alignment in Web Search. In: *CHI`12 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (2012), pp. 1341–1350.

[6] Kompan, M.: Group and Single-user Influence Modeling for Personalized Recommendation. In: *Information Sciences and Technologies Bulletin of the ACM Slovakia*, (2014), vol. 6, no. 2, pp. 11–20.

[7] Kovárová, A.: Special Interaction Approaches and their Impact on Usability. In: *Information Sciences and Technologies Bulletin of the ACM Slovakia*, (2011), vol. 3, no. 3, pp. 14–25.

[8] Teevan, J., Morris, M., R., Bush, S.: Discovering and Using Groups to Improve Personalized Search. In: *WSDM '09 Proceedings of the Second ACM International Conference on Web Search and Data Mining*, (2009), pp. 15–24.

Identification of Search Sessions with Use of Semantic Text Models

Martin PETLUŠ*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
martinpetlus@gmail.com

Abstract. Web search engine pages are one of the most visited pages on the Web. They are visited by users who come there to satisfy their informational needs. Users try to satisfy their needs by entering queries to the search engine and browsing the results returned by the search engine as a response to the entered query. These queries are often ambiguous and thus it is very hard for a search engine to retrieve relevant results for the query. One of the possible solutions to the query ambiguity is identification of previous related queries. With the identified previous related queries, we can predict the overall goal of the user and this way we can help the search engine with the ambiguous queries. Goal of this work is the identification of related queries using machine learning and semantic text models. Our proposed approach also considers multitasking, in which user interrupts work on the current task and later returns to the previously abandoned task.

1 Introduction

At present the number of documents on the Web is estimated in billions by most popular web search engines like Google, Bing or Yahoo[1]. The user who comes on the Web with a certain information need he wants to fulfill, does not have a chance to manually scan such number of documents, and these search engines play a major role in providing access to such huge number of documents.

Users interact with web search engine by entering a query (few keywords, also called terms), which describe their intent (search goal) and expect that a web search engine returns a list of documents relevant to the query. This classical approach, where the user interacts with the search engine by entering a single query, encourages the development of various text retrieval methods. However, it has many known disadvantages, such as the query ambiguity. If the user enters a query *jaguar*, web search engine does not know without any other information (i.e. without context) if we would like to see results about a car of this brand, or results about animal of this name. Query

* Master study programme in field: Information Systems
Supervisor: Dr. Tomáš Kramár, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

[1] http://www.worldwidewebsize.com/

ambiguity leads to unclear intent and therefore irrelevant result. Queries are also often too specific or too general [3].

The reality is that web search is more about questioning, learning and reformulating queries. The user often has to perform a number of actions in search engine to satisfy a single informational need [10]. Identification of user actions that have been made in the search engine to satisfy his single informational need is generally referred to as *detection of search sessions*.

The goal of search session detection is to segment the query sequence into segments (also called search sessions or simply sessions), where each segment satisfies the condition that queries in it are related with respect to a single informational need. Such detection of search sessions can greatly help search engine to return relevant results to the query. The term *search session* was never formally defined in the literature and it have various meanings in different works. In this work we use definition from [7], where this term means sequence of search related actions with the single underlying informational intent.

There are several existing approaches to detection of search sessions, but they have various disadvantages. Many of these approaches, e.g. [2, 5, 8], do not use semantics from queries. Such approach considers two queries as part of two different search sessions. These queries are not similar lexically, but they are related semantically. Many existing approaches also do not consider interruptions and multitasking in Web browsing. The user has multiple informational needs at the same time and he is working on them.

2 Related work

One of the simplest and most widely used approach to search session detection is based on comparison of temporal distance of queries. If temporal distance between two consecutive queries is larger than maximal, predefined threshold, it is assumed that user has probably changed informational need and that new search session started at this place containing the latter query. This approach was first described in [2], establishing the threshold of 25.5 minutes.

Different works experimented with different values of threshold. In work [9] threshold was set to have value of 30 minutes. In [6] experimented with different thresholds from 5 minutes to 2 hours. They also trained the best threshold on their data. The threshold may differ for each user. The goal in work [8] was to find the best suitable threshold for each user separately.

Although this approach is very simple to implement it has its drawbacks. One of them is that it can not detect a change of search session, which happened in a very short time – two consecutive queries, that are separated by small temporal difference, are falsely considered as part of the same search session, but they both might have different search intent. Another disadvantage is that the larger temporal difference may not mean that the user's intent changed. It might, for instance, mean that the user is reading longer article [4].

Another often used approach is to compare the lexical similarity of two consecutive queries. If two queries are lexically similar, then are considered as part of the same search session, they share the same search intent. Example of this approach can be found in [5]. This approach based on lexical similarity of queries is again easy to implement, but has its drawbacks. The main drawback is that, two queries may be completely lexically dissimilar, but they can be semantically related. For instance queries *bing* and *yandex* share no common words but they are both web search engines and thus very probably have the same intent.

Previous approaches does not consider *multitasking*. Multitasking of the user when he is browsing the Web can be present in several forms, such as *parallel browsing* or *browsing with interruptions*. Parallel browsing happens when the user has multiple search goals at the same time and he is working on them by having opened multiple browsers or multiple browser tabs. Browsing with interruptions means that the user interrupted work on the current search goal because of other goal or from other reason, and later he returns back to the abandoned search goal.

Approaches that use machine learning can be used to deal with multitasking. Examples of such approaches are presented in [6, 11]. When dealing with multitasking, these approaches basically

compares all queries to each other. Trained binary classifiers are used to detect whether such pair of queries are part of the same search task or not. Classifiers are trained with set of lexical, temporal, query log and web search features. Nevertheless, any of this approaches does not solve problem of query ambiguity.

In [12] semantics is used to identify underlying intent of queries. They use ESA model, which measures semantic similarity of queries using articles from wikipedia. Also LDA model is used to extract hidden topics from queries and URLs.

Our work and algorithm for search session detection is different from this research in one important way. We compare queries directly with our algorithm identified search sessions, not with other queries. If processed query has big similarity with one of the sessions already identified before this query, the query is considered as part of this session. Our approach also considers multitasking and query ambiguity. We use LDA model to extract hidden topics from queries to match semantic related queries.

3 Identification of sessions

The proposed method identifies search sessions by comparing currently issued query with search sessions started before the query. For this purpose we use algorithm described in [7]. Algorithm starts with ordered list of queries. Each query is associated with additional data such as search results, result titles and snippets. When processing the first query from the list, a new session containing this query is automatically created.

Algorithm then proceeds in the following steps:

1. the queries are ordered by the date they were typed into the search engine;
2. session stack T, which stores recent sessions, is initialized to empty stack;
3. oldest query from the ordered list is processed first;
4. stack T is traversed from top, where sits most recent session, to the bottom and each session S_c is examined;
5. the query q is compared with the session S_c and if this similarity is sufficiently high, then the query q is considered as part of the session S_c, the session S_c is removed from stack and added to the top of the stack T and becomes the most recent session, the processing continues with processing of the next query at the step 3;
6. if the query q is not similar enough to any of the sessions from the stack T, the new session S_c is created at that place and added to the top of the stack containing this single query q, again algorithm continues with step 3.

Session stack T is the main part which allows us to consider multitasking of users.

4 Comparing the query and search session

We propose two different algorithms for comparing search sessions and queries to find out, whether there is a sufficient probability that query belongs to the search session. Algorithms differ in the way they extract features from the query and the session. First algorithm extracts features between query and the whole session and second between the constructed query pairs. These features fall within the categories such as lexical, temporal, URL, and very important sematic features.

First algorithm is referred to as A_1. In this algorithm we union all queries from session into a single big query and compute cosine similarity between this big query and our query. In other

feature we compute number of common clicked documents between queries in the session and our query. Number of other features are extracted with this algorithm.

Algorithm referred here as A_2 is more complicated and borrowed from [1], where they compare two sets of queries. First we create query pairs with our query and all the queries in the session. From every such pair we extract features and this way we get a set of feature vectors for every pair. For every feature, we choose its minimum and maximum value, and compute standard deviation and mean value. In this way we get single feature vector that is four times bigger than the vector extracted from query pairs. This algorithm extracts features like cosine and jaccard similarity of pair of queries, number of common words that share queries, common URLs or common clicked URLs.

In our method we are using both those algorithms A_1 and A_2. Feature vectors computed with these algorithms are joined into a single feature vector. This vector is used as an input to a binary classifier in training and for trained binary classifier in identification of search sessions.

*T*raining the classifier. Training set for the classifier is generated from the input data. Training example is a feature vector extracted with algorithms described in previous section. Mean number of queries in session is 6.42 as established by [4]. Therefore it is advantageous to vary the number of queries in training sessions when extracting features to train the classifier also for sessions with lower number of queries in it. And therefore the number of queries in sessions is varied from 1 to original count. This is achieved by randomly removing queries from sessions. Supervised binary classifier needs for training both positive and negative training examples. When generating positive training example, random query from search session is removed and feature vector is extracted between this removed query and session without this query. Negative example is generated between random query from all search sessions excluding session that we are using for generating current training example. Positive and negative examples are generated from all training search sessions.

LibSVM[2] implementation of SVM is used to train binary classifier. Linear kernel and default parameter values were set for training.

*S*emantics. Each search result in the input data consists of document name and text snippet from the document shown in search engine results page. Every result, together with its name and snippet is used as a single training example for a text model. We use LDA[3] text model together with Gibbs sampling to infer hidden topics from this training examples. Number of hidden topics was set to 100 [12], number of iterations to value 300 and other parameters are set to default values[4].

Algorithm A_2 used for feature extraction, uses this trained model for computing the values of semantic features. For example, one feature first inferences topics for every clicked result from both query pairs. Text for topic inference is composed of result name and snippet. Because outcome of inference is vector of probabilities, we compute cosine of angle between every pair of clicked results and divide sum of these values by number of pairs. This way we compute value of this feature.

5 Evaluation

The proposed method is designed to work with search logs originating directly from a Web search engine. However, at present, such type of data is problematic to find. Available sources are suffering from some disadvantages, for example they contain anonymised queries (numeric identifier of the query instead of the original query terms), which are useless for extracting lexical or semantic features. The datasets do not contain enough queries and some datasets do not contain a list of returned documents with their full URLs. Furthermore, to the best of our knowledge, there is no publicly available dataset where the queries are labeled into search sessions, which would provide a useful baseline to compare to.

[2] `http://www.csie.ntu.edu.tw/~cjlin/libsvm`

[3] `http://mallet.cs.umass.edu/download.php`

[4] `http://mallet.cs.umass.edu/topics-devel.php`

For experiments, we use the TREC Session Track data[5]. This data was collected by asking participants to solve simple tasks which required them to use search engines in order to complete the task. This data thus contains an implicit session label – all queries issued in the process of task solving share the same goal and thus are in the same session by definition. Each query in session contains ranked list of search results returned as a response to the query and a ranked list of results clicked by the user. Each result contains URL of the document, title and snippet from the search engine. Each click contains time stamp when user visited document and time stamp when the user leaved document.

The data spans the period of 2011, 2012 and 2013. None of the input data files contains an explicit user identifier and therefore we treat them as originating from a single user. Queries are also not timestamped; for this reason we can not sort queries from oldest to newest, neither can we compute temporal difference between any pair of queries.

All texts in these files were normalized before further processing. Some of them were stored directly as HTML codes and therefore contained for instance HTML entities. Hence we first removed characters written in unicode. Then we removed all HTML entities. We also removed all characters which were not alphanumeric and we have also removed all digits-only words and words which contained no alphanumeric characters. English stop words were also removed, for this purpose we used a list of stop words that is part of a tool for machine learning[6]. Remaining words were stemmed with Porter stemmer algorithm[7].

We divided 170 sessions loaded from TREC data into two subsets. First subset with 25 sessions was intended for testing. The remaining subset of sessions with 584 queries was used for training the SVM and LDA. 205 queries in testing subset were shuffled randomly before identification to obscure original sessions. Identified sessions from shuffled queries were then compared against original sessions in the testing subset according to the methodology in [7]. Results averaged after four runs of implemented method are presented in the Table 1. From these results it is obvious that method performed best when using all of the features.

In evaluation of the trained LDA model we took a closer look at learned hidden topics from search results. One topic contained terms such as: *world*, *cup*, *franc*, *fifa*, *final*, *football*. After examination of the data we found queries that were typed into the search engine to fulfill task of finding when the France won final world cup in football. These terms are therefore semantically similar.

Table 1. Method results achieved by using different features.

Features	precision	recall	f_1-score
all	0.918	0.625	0.738
only lexical and semantic	0.903	0.608	0.726
only lexical	0.909	0.591	0.715

6 Conclusions and future work

In this work we have analyzed different approaches to search session detection. From this analysis, we have proposed and implemented our method. Main difference from the other methods is in the comparison of the query directly with search session and the usage of semantic models. Implemented approach is able to solve multitasking behaviour of users of the search engine. The proposed method

[5] `http://trec.nist.gov/data/session.html`

[6] `http://mallet.cs.umass.edu/download.php`

[7] `http://tartarus.org/martin/PorterStemmer/`

uses binary classifier with lexical, semantic and other categories of features. We also make use of semantic text model to learn hidden topics from the search results. Preliminary experimental results are promising.

In future we want to focus on experimenting with new features, feature selection, optimization of SVM and LDA learning parameters. Promising way is to learn hidden topics only from search results that user deemed useful. We also plan to conduct experiment on data coming from real web search engine and manually labelled queries into search sessions.

Acknowledgement: This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG 1/0646/15.

References

[1] Aiello, L.M., Donato, D., Ozertem, U., Menczer, F.: Behavior-driven Clustering of Queries into Topics. In: *Proceedings of the 20th ACM International Conference on Information and Knowledge Management*. CIKM '11, New York, NY, USA, ACM, 2011, pp. 1373–1382.

[2] Catledge, L.D., Pitkow, J.E.: Characterizing Browsing Strategies in the World-Wide Web. In: *Computer Networks and ISDN Systems*, 1995, pp. 1065–1073.

[3] Downey, D., Dumais, S., Liebling, D., Horvitz, E.: Understanding the Relationship Between Searchers' Queries and Information Goals. In: *Proceedings of the 17th ACM Conference on Information and Knowledge Management*. CIKM '08, New York, NY, USA, ACM, 2008, pp. 449–458.

[4] Hagen, M., Gomoll, J., Beyer, A., Stein, B.: From Search Session Detection to Search Mission Detection. In: *Proceedings of the 10th Conference on Open Research Areas in Information Retrieval*. OAIR '13, Paris, France, France, LE CENTRE DE HAUTES ETUDES INTERNATIONALES D'INFORMATIQUE DOCUMENTAIRE, 2013, pp. 85–92.

[5] Jansen, B.J., Spink, A.H., Blakely, C., Koshman, S.: Defining a session on Web search engines. *Journal of the American Society for Information Science and Technology*, 2007, vol. 58, no. 6, pp. 862–871.

[6] Jones, R., Klinkner, K.L.: Beyond the Session Timeout: Automatic Hierarchical Segmentation of Search Topics in Query Logs. In: *Proceedings of the 17th ACM Conference on Information and Knowledge Management*. CIKM '08, New York, NY, USA, ACM, 2008, pp. 699–708.

[7] Kramár, T., Bieliková, M.: Detecting Search Sessions Using Document Metadata and Implicit Feedback. *Proceedings of the WSCD 2012 Workshop on Web Search Click Data*, 2012.

[8] Mehrzadi, D., Feitelson, D.G.: On Extracting Session Data from Activity Logs. In: *Proceedings of the 5th Annual International Systems and Storage Conference*. SYSTOR '12, New York, NY, USA, ACM, 2012, pp. 3:1–3:7.

[9] Radlinski, F., Joachims, T.: Query Chains: Learning to Rank from Implicit Feedback. In: *Proceedings of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery in Data Mining*. KDD '05, New York, NY, USA, ACM, 2005, pp. 239–248.

[10] Spink, A., Jansen, B.J., Ozmultu, C.H.: Use of query reformulation and relevance feedback by Excite users. *Internet Research: Electronic Networking Applications and Policy*, 2000, vol. 10, no. 4, pp. 317–328.

[11] Ustinovsky, Y., Mazur, A., Serdyukov, P.: Intent-Based Browse Activity Segmentation. In Serdyukov, P., Braslavski, P., Kuznetsov, S.O., Kamps, J., Rüger, S.M., Agichtein, E., Segalovich, I., Yilmaz, E., eds.: *ECIR*. Volume 7814 of *Lecture Notes in Computer Science*., Springer, 2013, pp. 242–253.

[12] Zhang, Z., Sun, L., Han, X.: Learning to detect task boundaries of query session. In: *Proceedings of the 22nd ACM international conference on Conference on information & knowledge management*. CIKM '13, New York, NY, USA, ACM, 2013, pp. 1885–1888.

Implicit Feedback-Based Estimation of Student's Knowledge

Veronika ŠTRBÁKOVÁ*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
veronikastrbakova@gmail.com

Abstract. Nowadays, there is a number of adaptive educational systems, which adapt the content automatically for a student by monitoring his activities. Inaccurate information obtained by the evaluation of implicit feedback based on the student's behavior has an impact on accuracy of the recommendation. With the increasing options of the monitoring activities such as signals from eye camera we can more accurate evaluate the implicit feedback and interpret various activity signals of the user implicit feedback. Our research aims to monitor a student during his study in a web-based educational system and to estimate active time spent on learning objects. Based on this we estimate the level of his knowledge on concepts related to given learning objects.

1 Introduction

With the increasing numbers of web users there is also an increase of information on the Internet. It can be difficult for the user to separate relevant information from the non-essential. In most cases the Web sites treat all users in a same way and show them the same content, even though the users are diverse. Thus it happens that the user gets the information that is too general for him, or too specific, which may be a problem, especially if it is a domain of education [1].

Nowadays, it is more common, that the learning process is being transferred from adaptive educational systems, where students have possibilities to study available materials in different courses in addition to the possibility of elaborating a variety of exercises and questions and, inter alia, collaborative work together to solve these tasks with their classmates or teachers. It is common for such systems to store information about users (students) – create a so-called user model for the purpose of estimating their interests, knowledge of such content and their future behavior, eventually in pursuance of gathered information to include them to the group of users with similar properties [2]. The user model is created based on a feedback from the user or a group of users. There is an explicit and implicit feedback. We deal with the implicit feedback.

The goal of this work is to propose a method for the implicit feedback-based estimation of student's knowledge, therefore we monitor the student activity during the study of the learning objects. Despite the existing methods for user modelling that evaluate implicit feedback in form of

* Master degree study programme in field: Information Systems
Supervisor: Professor Mária Bieliková, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 181–186.

various signals of user activity, which aim to explore its characteristics, there is still a room for improvement. Inaccurate information obtained by the evaluation of the implicit feedback based on the student's behavior, have an impact on the accuracy of user modeling. With the increasing possibilities of monitoring users on the Web, like signals from eye tracking camera, blood pressure, body temperature and pulse sensors, we gain the ability to evaluate implicit feedback with great accuracy, and with that, gain the related interpretation of various signals of activity not only in the domain of education mentioned before.

2 Related work

Nowadays, there are several forms of implicit feedback. For the evaluation of the user's interests on the web page, we can monitor the different signals of user activity such as mouse movements, the number of mouse clicks, scrolling of the page, pressing different keys and selections of the text on the web page [4]. In the domain of the information retrieval and recommendation the user's interests can be modeled after the clicking on the recommended link, the time spent on the web page, adding the tags, adding a page to bookmarks but also based on leaving the page [3].

In the domain of the education, there is a number of adaptive educational systems of which content adapts itself for the student by monitoring his activities during the study. These are in addition to user domain enriched by the domain model, which captures the structure of the content of the education system and the relationships between components (in terms of navigation). Regarding the student work with the learning object, it is possible to monitor the several implicit feedback indicators of interest from the basic - such as the number of mouse clicks, through the actions as the selection of a text on a page and searching the page to the scrolling of the page, respectively from the scrolling derived time of reading the different parts of the web page [6, 7].

In some elearning systems the user characteristics stored in the model result from the monitoring the user interaction with the learning object, and after a certain number of actions it evaluates the user's knowledge, which is subsequently stored to the model. During the reading of the interpretation of the user modeling, they take into consideration the display time and the number of clicks. In questions, the process of evaluating the correctness of the results is either automatic or the student explicitly evaluates his answer when the correctness of this information is not determined. This approach uses for example the Alef system (Adaptive LEarning Framework) which is an adaptive web learning system created at the Faculty of Informatics and Information Technologies [8].

Another approach of the modeling of the user knowledge is used in the KBS-Hyperbook [5]. The user modeling here is based on indexing any information resources situated in the system by using the so-called knowledge items. The knowledge of a user is modeled as a knowledge vector. Each component of the vector is a conditional probability describing the system's estimate, that user has knowledge about the topic, what emerge from all the observations of the user by the system.

In our work we estimate the knowledge of student on the basis of monitoring his activities. During the performing these activities we measure the active time of work. We determine the similarity of the learning objects based on the average time spent with working of all the users. The estimation of the student's knowledge is a result of the evaluation of this time in relation to the learning objects characteristics.

3 Method of estimation of the student's knowledge

Despite the existing methods for user modelling that evaluate implicit feedback in form of various signals of user activity, which aim to explore its characteristics, there is still a room for improvement. Inaccurate information obtained by the evaluation of the implicit feedback based on the student's behavior have an impact on the accuracy of user modeling. The argument that user

has a learning object displayed for some time and he also actively reads it while his level of knowledge increases does not always have to be true. It is because the user does not have to be near the computer or he might even be performing an activity that does not relate to the study.

Our goal is to estimate the student's knowledge based on the active time of working with the learning object. To determine the level of the user's knowledge related to some concept included in a learning object, we propose to take a closer look at some of the learning object characteristics, more specifically:

- understandability, the readability index (we use ARI and LIX metrics [9]),
- concepts connected with the given learning object,
- the average time of all students spent working with a learning object.

Various signals of user's activity play an important role in determining the value of the knowledge level. We monitor these signals when users are working with the educational system while we focus mainly on those that help us detect time of active work with the learning object, e.g.: the display time of the learning object, the reading time of the learning object, mouse clicks, scrolling and pressing different keys.

These characteristics of the learning object along with the student's actions go to the input of our method. Subsequently, the method performs the following steps:

1. Estimates the active time spent working with the learning object
2. Estimates the level of the student's knowledge of the learning object
3. Estimates the level of the student's knowledge of the concept connected to the given learning object

3.1 The characteristics of the learning object

During the estimation of the active time of the work that the student performs with the learning object, it is important to know what is the complexity of the learning object to read and understand. Moreover, if we can be allowed to compare the learning objects with each other, we need to organize them by some of their attributes. In our method we use the Automated Readability Index (ARI) metric to express the readability of the learning object and the LIX formula metric (LIX) to determine the understandability of the learning object. The readability index and the LIX formula generally speak about the difficulty of reading, respectively about the necessary education to understand the text.

Nowadays, there are several readability indexes. We have decided to choose the ARI and the LIX, since those do not require the calculation of the syllables of each word and the calculation process is relatively quick. We can obtain the readability and the understandability with the following formulas according to [9]:

$$\boldsymbol{ARI} = \mathbf{4.7} \times \frac{\boldsymbol{chars}(\boldsymbol{d_i})}{\boldsymbol{words}(\boldsymbol{d_i})} + \mathbf{0.5} \times \frac{\boldsymbol{words}(\boldsymbol{d_i})}{\boldsymbol{sent}(\boldsymbol{d_i})} - \mathbf{21.43} \tag{1}$$

$$\boldsymbol{LIX} = \frac{\boldsymbol{words}(\boldsymbol{d_i})}{\boldsymbol{sent}(\boldsymbol{d_i})} + \mathbf{100} \times \frac{\boldsymbol{longWords}(\boldsymbol{d_i})}{\boldsymbol{words}(\boldsymbol{d_i})} - \mathbf{21.43} \tag{2}$$

where we take into consideration the number of characters , the number of words, the number of sentences and the number of long words in the learning object, which contains more than 6 characters. These values are adjusted by various constants obtained experimentally [9].

If the document is read by more users, these times can be averaged and organized into a sequence. The average time spent working with a document for all users together with understandability, readability and others metrics mentioned before, are used to express the overall complexity of the document. Our aim is to find the time for a single user, by which the user starts to acquire the given knowledge and the time that the user must actively spend working with

a learning object, to gain the greatest knowledge. For the correct settings of the user's knowledge, it is necessary to know which concepts are connected with the learning object.

3.2 Monitoring of student's actions performed in the educational system

We monitor various signals of the user activity while working with the educational system, at which we specialize mostly on the ones that help us to determine the time of an active work with the learning object, such as:

- the display time of the learning object,
- the reading time of the learning object,
- the mouse clicks and scrolls,
- the pressing of different keys.

From these signals the average time of work with the learning object can be derived for a given user. The value of characteristics of the knowledge, interests and goals by itself is represented by a three-dimensional vector [value, certainty, source]. The value of the characteristic is acquired from the enclosed interval of <0, 1>, wherein 1 is the maximum value. Each of these characteristics has an assigned certainty that reflects whether the user really acquires the value of the characteristic. The source determines, how the value and the certainty of the characteristic gets into the user model. One characteristic can be thus determined by various sources with different levels of the certainty. The paper compares the signals of the user's activity and finds out, which source gives the most accurate value of the reading time of the learning object.

3.3 The student's knowledge estimation

The characteristics of the learning object along with the observed student's actions mentioned before go to the input of our method. To estimate the level of a student's knowledge we have to estimate the active time of working with the learning object. We monitor various signals of the student activity while working with the educational system. Therefore it is necessary to evaluate the weight of each signal based on how these signals are related to the active time of the user. Based on these activities we estimate the whole active time of student. We measure the time of performing these activities.

Once we have estimated the active time of working with the learning object, on the basis of similarity of the learning objects in terms of readability, understandability and others metrics mentioned before, we can determine the increase of knowledge for the unread learning objects. If the student reads a given number of learning objects, we can determine, for the unread learning objects, how long will it take on average to read each of them. For example, the learning objects l_1, l_2 and l_3 have been read by other students on average 11 minutes. The given student read l_1 and l_2 in 11 and 12 minutes. If he gets on input the learning object l_3, we can assume that the student will read it in the interval of <11, 12> minutes.

Our goal is to find the dependence between the time of the active work with the learning object and the student's knowledge which he acquired. At what time of the reading the user starts to acquire the knowledge and at what time he acquires the maximum value of knowledge? We assume that the user acquires the highest value of knowledge in time close to the other times of reading the similar learning objects on the basis of the average time of reading by other students. In order to start the whole process of estimation of the student's knowledge, the minimal time of the active work with the learning object has to run out. It is therefore that student might have displayed the learning object accidentally or knowingly however somebody or something disturbed him and he interrupted the work. Based on these we estimate the level of the student knowledge to the learning object, as is in the following formula:

$$\boldsymbol{V_n = V_o + \alpha \times V_i} \tag{3}$$

where $\boldsymbol{V_n}$ is the final knowledge, $\boldsymbol{V_o}$ is the original knowledge and $\boldsymbol{\alpha} \times \boldsymbol{V_i}$ is the increment of the knowledge measured during the last visitation of the given learning object. The result value is normalized to a value from the enclosed interval of <0, 1>.

We are aware of the limitations of our method like a number of visitations of the learning objects and an impact of the characteristics of the learning objects on the estimation the student's knowledge for just visited learning object. We are trying to find out whether the characteristics of the learning object affect the resultant graph and if so, we try to find how. We also investigate what effect does the number of visits of the learning object have on the overall knowledge.

For the concept contained in the given learning object we compute the level of the student's knowledge, such as:

$$\boldsymbol{V_c} = \boldsymbol{V_{c_s}} + \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{V_{i_c}} \tag{4}$$

where we add to the original level of the knowledge the sum of the knowledge acquired in the learning objects, which are connected with the given concept.

4 Evaluation and conclusion

We performed an experiment to discover interesting user behavior during study. We focused on the observation of 8 users studying in the adaptive educational system Alef. According to the degree of the study we could divide these users into several groups: 2 students of the bachelor degree study program, 4 students of the master's degree program, 1 student of the doctorate degree study program and 1 student, who has finished the doctorate degree study program. The students were male users and two of them had no knowledge about the ongoing Functional and logical programming course. One of the students was studying it and others had completed the course with the different time interval and while answering the questions they had only been repeating the materials. Each participant was tasked to do exercises in this course. If the users did not know how to answer the questions from the exercises they could also study the related study materials included in the course. We chose the questions according to the user skills in the given domain, while trying to maximize the studying of the study materials.

During the whole experiment we were using the Eye-Tracker technology. We observed the gaze of the users, the movements of the cursor, and the presses of the keys on the keyboard. We recorded the behavior of each user using the web camera with additional information recorded on paper – opening of new tabs in the browser, interesting cursor movements, and other behavior characteristic for the given user.

The aim of this experiment was besides just addressing the limitations of the modeling of knowledge only on the basis of the implicit feedback, collected via the server side, to reveal the interesting patterns of the student behavior during the study. We analyzed the data using the Tobii Studio software, with additional data coming from the Alef system database. We found out that only one of the students used for solving the exercises and studying more tabs in browser. While the students were working, there were an unexpected circumstances, which were not captured by the Alef system or by the Tobii studio. They were only captured by the web camera and this was mainly the action when the user's mobile phone rang, or the user was disturbed by someone after which they spent some time talking. The time, which the students spent on activities other than studying, was added to the overall time of the active work with the learning objects and influenced the computation of the student's knowledge.

We also designed the second experiment, which goal is to find the dependence between the time of the active work with the learning object and the student's knowledge. We plan to achieve this by comparison of the evaluation of the acquired knowledge based on the student results from the pretest and posttest from the Functional and logical programming course with the time of the student's active work spent on the learning objects. As first, we plan to ask students to solve the pretest from the course. Subsequently, the students will have the task to study the materials in

the course with the goal to prepare themselves for the other test in the course. During the process we will monitor them by our browser extension, which captures the mouse movements, scrolls and in cooperation with the web camera detects the frontal face. Finally, we will ask them to solve the posttest from the course. We will focus on the observation of 20 users of the first year of the bachelor degree study program, who have no knowledge about the given course.

The user model is an important part of the adaptive learning system. It influences the quality of the content adaptation to the student. Despite the existing implicit methods of evaluation for various signals of user activity, which aim to explore its characteristics, there is still room for improvement.

In addition to existing approaches we present a method for the estimation of student's knowledge based only on the active time of working with the learning object. Furthermore, we bring students characteristics and the learning objects characteristics which haven't been used in previous approaches. For example the average time of work with the learning object for all users or a single user, the readability and the understandability of the text contained in the learning object. We use the browser extension which captures the mouse movements, scrolling and detects the face in front of the display with web camera. In the near future we plan to realize this experiment and based on the results we plan to improve our method.

Acknowledgement: This work was partially supported by the Cultural and Educational Grant Agency of the Slovak Republic, grant No. 009STU-4/2014.

References

[1] Berendt, B., Spiliopoulou, M.: Analysis of navigation behavior in web sites integrating multiple information systems. The VLDB Journal, (2000), vol. 9, no. 1, pp. 56–75.

[2] Brusilovsky, P., Millán, E.: User models for adaptive hypermedia and adaptive educational systems. The adaptive web, (2007), vol. 4321, pp. 3–53.

[3] Fox, S., Karnawat, K., Mydland, M., Dumais, S., White, T.: Evaluating implicit measures to improve web search. In: *ACM Transactions on Information Systems (TOIS)*, ACM, 2005, pp. 147–168.

[4] Hauger, D., Paramythis, A., Weibelzahl, S.: Using Browser Interaction Data to Determine Page Reading Behavior. In: *UMAP'11 Proc. of the 19th Int. Conference on User modeling, adaption, and personalization*, ACM, 2011, pp.147–158.

[5] Henze, N., Nejdl, W.: Adaptation in open corpus hypermedia. In: *Int. Journal of Artificial Intelligence in Education*, (2001), 12(4), pp. 325–350.

[6] Kelly, D., Belkin, N.J.: Reading time, scrolling and interaction: exploring implicit sources of user preferences for relevance feedback. In: *Proc. of the 24th annual Int. ACM SIGIR Conference on Research and Development in Information Retrieval*, ACM, (2001), pp. 408–409.

[7] Longo, L., Barrett, S.: A context-aware approach based on selforganizing maps to study web-users' tendencies from their behavior. In: *Proc. of the 1st Int. Workshop on Context-Aware Middleware and Services affiliated with the 4th Int. Conference on Communication System Software and Middleware (COMSWARE 2009) – CAMS '09*, ACM, (2009), pp. 12–17.

[8] Unčík, M.: Modelovanie používateľa v doméne webovo-orientovanej podpory vzdelávania. Master's thesis, (2012), (in Slovak).

[9] Yilmaz, E., Verma, M.,Craswell, N., Radlinski, F., Bailey, P.: Relevance and Effort: An Analysis of Document Utility. In: *CIKM '14 Proc. of the 23rd ACM Int. Conference on Information and Knowledge Management*, ACM, (2014), pp. 91–100.

Students' Behaviour Analysis in e-Learning System

Ondrej KAŠŠÁK*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
ondrej.kassak@stuba.sk

Abstract[1]

The ability to predict the user's future behaviour represents an important information, which can be used for personalization of the system the user interacts with. Personalized system is able to adapt its content for purposes of every user individually, which improves the interaction quality.

In this paper we focus on e-learning domain, where we predicted users' behaviour in various learning courses. Our aim was to answer the question: "Will the user leave the system in his/her next action or will he/she continue browsing the learning objects (web pages) in the system?" In this task, we have to deal with the high unbalance between the predicted classes (leave or stay), because user normally browses through more than ten learning objects before he/she leave. In the e-leaning, there is also a large heterogeneity in the behaviour among the users and also they behave differently in various course parts (e.g. at begin of the course and at night before the exam). The data from come in a continuous stream, so they should be processed in one iteration.

To reduce mentioned problems we proposed the polynomial classifier using stochastic gradient descent method to learn the classification attributes' importance. These importance is in addition considered individually for every user. The attributes used by classifier describe, for every observation (user's interaction with learning object), the properties of learning objects, characteristic traits of interaction and its context. One part of attributes is directly logged by learning system, other ones were derived by us. The individual attributes importance eased the problem with the heterogeneous behaviour.

Usage of stochastic gradient descent allowed us to deal with high amount of data and also of their diversity in time (various users' behaviour in time). We reduced the unbalance of the multiplicity of the observations belonging to different classes by the assigning the different weights to them based on the class they belong to (leave or stay) with favour to rarer class.

To evaluate our approach we applied the method to the data from real e-learning system and also on the bank telemarketing domain data. In both domains, we reached promising results.

Acknowledgement: This work was partially supported by the Cultural and Educational Grant Agency of the Slovak Republic, grant No. 009STU-4/2014.

---

* Doctoral degree study programme in field: Software Engineering
Supervisor: Professor Mária Bieliková, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

[1] Full paper available in printed proceedings, pages 187–195.

Considering Navigational Value of Keywords in the Process of Navigational Leads Selection

Róbert MÓRO*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
robert.moro@stuba.sk

Abstract[1]

For the type of searches that are open-ended, start with ill-defined information needs which can change over time as the users get more acquainted with the domain and the problem at hand and which require the use of different search strategies, the term *exploratory search* was coined. A prime example of this type of search that is often conducted by the novice researchers is an exploration of a new domain and search for the state-of-the-art.

Considering the novice researcher scenario, we have proposed a method of supporting the exploratory navigation using the navigation leads in the abstract or a summary of a document that help the novice researchers to filter the information space. For the term to be selected as a navigation lead, it should be relevant for the document in which it is identified as well as reflect the information subspace that is covered by the lead. We denote the latter as *navigational value* of the term. In order to compute the navigational value of a term (a lead candidate), we need to identify the subspace behind the term. We employ clustering Latent Dirichlet Allocation; the subspace is thus represented by a cluster that the document, for which we identify the lead, belongs to.

In order to evaluate our approach we have conducted an experiment on data from a web-based bookmarking system Annota. We have compared the resulting clusters when using different sources of keywords; for the evaluation of the clusters we have used a measure of perplexity. The best performance was achieved by the keywords extracted from the whole content of a document followed by the user-added tags. This suggests that the users use descriptive tags of a high quality and also that the abstracts themselves (which were also compared) are not enough to find good clusters of the documents. The worst performance was achieved by the author-added keywords.

The main contribution of the proposed approach is in consideration of the navigational value of the term when computing its relevancy. In addition, it combines the depth-first navigation with wider exploration, because it considers also less relevant topics (clusters) that are associated with the documents. As future work, we plan to conduct a user study that would evaluate the proposed method of exploratory navigation and the leads selection during real search sessions.

Acknowledgement: This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG 1/0646/15.

---

* Doctoral degree study programme in field: Software Engineering
Supervisor: Professor Mária Bieliková, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

[1] Full paper available in printed proceedings, pages 195–202.

Towards Preservation of Sustainability in Community Question Answering

Ivan SRBA*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
ivan.srba@stuba.sk

Abstract[1]

With increasing popularity of online communities gathered in knowledge sharing systems (e.g. Wikipedia, forums, and mailing lists), new forms of systems based on collective intelligence constantly emerge. One of the most popular among them is a concept of Community Question Answering (CQA) sites, such as Yahoo! Answers or Quora. Currently, the existing CQA systems are perceived mainly as a successful example of open community-based knowledge management systems due to their high popularity, millions of answered questions, very fast question answering process as well as availability to anyone.

In spite of that, some of the biggest and the most popular CQA systems are recently not such successful as they used to be. More specifically, we can witness new emerging problems in some CQA systems – an increasing failure (a proportion of unanswered questions) and churn (a proportion of users who leave the community) rate. In order to investigate this trend in more details, we conducted a case study focused on Stack Overflow. At first, we analyzed emerging problems by means of a community perception. Community proposed that the emerging problems are highly related to the growing amount of undesired groups of users that produce a great amount of low quality content: help vampires, noobs and reputation wh*res. Afterwards, we confirmed the community perception by quantitative analyzes on data from the question answering process itself.

With respect to the findings obtained in the case study, we found out that the majority of state-of-the-art to collaboration support in CQA systems does reflect the emerging problem, moreover some approaches, which can be also characterized as asker-oriented, even indirectly support and motivate the undesired groups of users in posting of additional low-quality content. Therefore, we propose novel answerer-oriented approaches to collaboration support that, in addition, involve a whole community in a question answering process. These approaches represent an eminent attitude change in the existing question-answering support methods with the aim to preserve a long-term sustainability of CQA ecosystems.

Acknowledgement: This work was partially supported by the Cultural and Educational Grant Agency of the Slovak Republic, grant No. 009STU-4/2014.

---

* Doctoral degree study programme in field: Software Engineering
Supervisor: Professor Mária Bieliková, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

[1] Full paper available in printed proceedings, pages 203–210.

Computer Graphics, Multimedia and Computer Vision

Fast gSLIC for Superpixel Segmentation

Róbert BIRKUS*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
rbirkus@gmail.com

Abstract[1]

The SLIC algorithm proposed by Achanta et al. [1] is a superpixel segmentation algorithm based on k-means clustering, which efficiently generates superpixels. It seems to be a good trade-off between the time consumption and robustness. Important advancement towards the real time applications using superpixels has been proposed by the authors of the gSLIC – a modified SLIC implementation on the GPU (Graphics Processing Unit) [2].

We have presented an important modification in the clustering part of the gSLIC algorithm. We implemented a parallel reduction through which we achieved significant acceleration. The results can be seen in the Table 1. The modification does not have impact to the quality of the segmentation. We have also presented a segmentation quality comparison between SLIC and gSLIC. The results of the comparison are different, depending on the evaluation method.

Table 1. Time evaluation on the NVIDIA GTX 770.

Image Size	gSLIC	Fast gSLIC	Speed-up
320x240	5.326 [ms]	0.426 [ms]	12.5x
640x480	8.0 [ms]	1.539 [ms]	5.20x
1280x960	19 [ms]	6 [ms]	3.17x

Acknowledgement: This research has been supported by a grant VEGA 1/0625/14.

References

[1] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk. Slic superpixels. Technical report, Ecole Polytechnique Fedrale de Lausanne, Report No. EPFL-REPORT-149300, 2010.

[2] Carl Yuheng Ren and Ian Reid. gSLIC: a real-time implementation of SLIC superpixel segmentation. Technical report, University of Oxford, Department of Engineering, 2011.

* Bachelor study programme in field: Informatics
Supervisor: Dr. Vanda Benešová, Institute of Applied Informatics, Faculty of Informatics and Information Technologies STU in Bratislava

[1] Full paper available in printed proceedings, pages 213–218.

Sitting Posture Quality Evaluation Using Depth Camera

Jakub MAČINA, Jozef STAŇO, Matej LEŠKO
Jozef GÁBORÍK*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
imaginecup2015@googlegroups.com

Abstract. On a long term basis, sitting in a wrong posture causes health problems like back pain and headaches. It can leads to decrease of a productivity, loss of confidence or even serious disorders. This paper further analyzes evaluation of a user's sitting posture quality. We propose several methods for detecting posture of a computer user. We extract features from depth and RGB camera streams using image processing and computer vision algorithms. These features are an input for neural networks. We are also detecting points at different parts of user body and comparing depths with user's calibrated image. Experimental results show over 93 % accuracy for the best method of the detection.

1 Introduction

Human posture estimation from images is challenging task due to viewpoints variations, noise and lack of a third dimension. This paper focus on evaluation of sitting posture quality, which is one special case of a human posture. Because it turns out, that nowadays people are sitting for a huge part of the day. And for most of that time, they are sitting incorrectly. Sitting in a wrong posture on a long term basis causes health problems – ranging from pain in back, neck, shoulders to more serious disorders [3]. Impact is huge, according to WHO paper [2], the most common health problem is low back pain with the lifetime prevalence estimated 60–70 % in industrialized countries.

We propose several methods for evaluation of sitting posture quality. We are using streams from depth and RGB camera to evaluate user's sitting posture. *Depth images* provide us many advantages over traditional RGB images. They are illumination invariant and more easier to segment. To make our solution works correctly it is necessary to set up a camera to capture an upper part of the user's body. Firstly, we are using *Histogram of oriented gradients* (HOG) feature descriptor for depth and color images. These feature descriptors are then processed by separated feedforward neural networks. Next method detects *mesh of depth points* at different parts of user's body and compare them with user's calibrated image. The last method computes custom depth features from an image

* Bachelor study programme in field: Informatics
Supervisor: Dr. Jakub Šimko, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

region of interest (ROI), which is the user's body, and these features are then fed into *neural network* for posture classification.

2 Related work

In a field of a real-time posture tracking, there is a software called *PostureMinder*[1]. It is necessary for users to calibrate application by taking photo of themselves in a correct sitting posture and cropping location of their face. However, this application only uses an algorithm that compares area of an actual face location with a calibrated image. As a result, PostureMinder recognizes only three posture states in a comparison with a calibrated image – leaning forward, correct posture and slouching.

In the a hardware field, there are wearable devices such as *Lumoback*[2] and *iPosture*[3]. These devices are worn around your lower back or clipped at clothes. They can monitor user posture for an entire day, but it can lead to discomfort wearing them on the body. User has to calibrate the device before using it by doing several casual activities. But concerning the back posture, they only monitor spinal bending. While consulting with a specialist, we realized that it is not sufficient. Wearable devices uses accelerometer to measure differences in three axis. Alerts are done by vibrations. There is also a cushion that monitors user's sitting posture and his vitals[4].

There are some studies that used skeleton data from *Kinect* for posture detection [4]. However, skeleton data are not working properly in a near camera range.

3 Concept overview

The concept of our solution is aimed at monitoring user's posture while working with a computer. User's posture detection process is characterized by following phases:

1. Acquisition of a current frame from camera every few seconds while running in the background.

2. Classification of user's sitting posture by several methods. Method total result is based on a result and certainty of that result. All methods are then combined into evaluation of sitting posture. Each method contributes to the evaluation by equal weight.

3. Provide an informative feedback for user at the right time. When user is sitting wrong for a longer period of time, it checks user activity like typing, scrolling and windows switching. Notifications are showed at the right time according to these features or when the predefined time limit is reached.

Advantage of our solution, compared to related works, is usage of a depth camera, so it is more precise at detecting a user's posture and capable of recognizing more sitting posture states.

4 Features extraction

A preprocess is required in order to extract features from the image. Firstly we find all *contours* in this image. Contour is a curve along boundary of pixels with the same intensity or depth. Assuming correct setup of camera pointed at a user and capturing only depths in a close range from camera, we

[1] `http://www.postureminder.co.uk/`
[2] `http://www.lumobodytech.com/lumoback/`
[3] `http://www.iposture.com/`
[4] `http://darma.co/`

filter and process only contour with maximum area. It is our region of interest as it is very likely to be a user. In the next step, statistical properties called *moments* are calculated. These moments are very useful for describing objects represented by contours. From these image moments the center of the contour known as centroid is calculated.

4.1 Depth comparison

The following method detects *depths at different parts of a user's body*. Then we calculate differences from the user calibrated image. To find a points on a user's body which are used for comparison, we use similar technique like raycasting in a computer graphics. Given a started point and an angle, algorithm is moving in a specific angle and finds boundary of an actual region. *Boundary of a region* is set when difference from previous pixel depth exceeds a specified threshold. Some of the points are shown in Figure 1.

This method can recognize several posture states comparing differences with the user calibrated image using specified threshold. It checks difference between average head depth and upper body depth. Along with checking head average depth, if it's proximity to the screen is not too close, these are important indicators for detection of a leaning forward. Leaning to the sides is checked by comparing head position x coordinates with the centroid. Detection of a slouching is done by computing absolute differences between horizontal upper body points. Twisting body to the side is detected by an average left depth in a comparison with an average right depth.

Figure 1. Visualization of Depth comparison (left) and ROI features (right).

4.2 Region of interest features

This method extracts other interesting features from ROI of depth image. All values are depth and region size independent. We create an 11 dimensional vector of features that is than fed into feedforward neural network. Features vector is composed of:

- Bounding box extent – Scalar that specifies ratio of ROI area compared to area of its bounding box. Area is defined by a number of pixels in a region.

- Convex hull extent – Similar to previous, but with comparison to convex hull area. Convex hull is the smallest convex polygon which can cover the region.

- ROI orientation – Angle between the horizontal axis and the major axis of the ellipse fitting the region. Degrees are from range 0 to 180 and increasing from right to left.

- Centroid – X and y centroid coordinates as a ratio to bounding box width and height.
- Eccentricity – Scalar that specifies elongation of the ellipse fitting the region. The value is in range between 0 and 1, where 0 represents circle and 1 represents a line.
- Mean depth – Region mean depth value. In order to make it depth invariant, the value is in a relation to the range of all image depth values. Depth values are noisy, we filter only depths with significant counts.
- Depth with maximum count – Similar to previous, but calculated for the depth with the maximum count in a region of interest.
- Ratio of left to right average depth – Computes average depths in a region on the left and the right side from centroid.
- Head to centroid depths difference – Find a neck as the most narrow part of the region. Calculate average head depth from the area above neck boundary. Than difference between head and centroid average depths is calculated.
- Top depths with maximum counts – Calculate a histogram for depths in a region, filter only top 10 depths and get a sum of the absolute differences between first and second top depth, second and third and so on.

4.3 Histogram of oriented gradients

Histogram of oriented gradients [1] is a feature descriptor for an object classification commonly used in a computer vision. It produces the same feature vector for the same object viewed under different lighting conditions. HOG feature descriptor uses a *sliding window* which is moved across the image and at each position, the gradient vector for each pixel is computed. Then it takes all feature vectors and inserts them in the histogram with several bins. Normalization of values is done in blocks to suppress the effects of a change in an image contrast.

We are extracting HOG features for RGB images in a full size. For bounding box of user's upper body we use the same approach too. The upper body bounding box is found by *haarcascade classifier*. Moreover, we try this approach for depth images. For depth images, we can also crop *region of interest* from image and scale it to a constant image size of 320x240 pixels. It leads to a smaller and constant feature vector and invariant to other contours in image. In the both cases we trained the *deep neural network* for a posture classification.

5 Posture classification

Our goal is to perform *binary classification* of an actual image into a correct or a wrong sitting posture state. We implement a *feedforward neural network* with resilient backpropagation learning, which takes into account only a sign of a derivative for each weight and therefore adaptively update each weight. Hidden layers of neural networks use sigmoid activation function. Because it represents a classification problem, output layer uses *softmax activation function*, so the sum of outputs from all neurons in this layer equals one. This gives us a probability for a particular posture state. Lastly, according to these two states, threshold value is used to specify final single value result.

We are using full batch learning and provide training, validation and testing data set. We prevent network from overfitting by setting maximum number of successive iterations that error on the validation set fails to decrease. Other stopping conditions are number of epochs and minimal error on training data set.

For neural network classification with larger number of inputs e.g. HOG, the *stacked autoencoder deep neural network* was implemented. Autoencoder tries to learn compressed representation of its

input. Stacked autoencoder neural network consists of multiple autoencoders, learned one layer at a time. This pretrain deep network. After all, we fine-tune it for a classification.

6 Evaluation

In our experiment, we used RGB and depth photos of four women and six men while sitting in front of a computer. These photos are from variety of places, camera placements and participants are captured in a different posture. Seven people (four men, three women) was selected into the set A and three (two men, one woman) for the set B. Set A was further randomly divided into training and validation set with size ratio approximately about 80–20 %. Set B, that contains different people from set A, was used as a testing set. The distribution of photos are shown in table 1.

Table 1. Correct and wrong sitting samples distribution in the sets.

Set	Correct	Wrong
Training	373	613
Validation	89	145
Test	122	160
Total	584	918

Table 2. Posture classification using depth comparison.

Person	Success rate[%]	TPR	TNR
Woman 1	88.888	0.929	0.864
Woman 2	91.304	0.909	0.916
Woman 3	94.444	0.907	0.969
Woman 4	93.750	1.000	0.875
Man 1	95.477	0.917	0.982
Man 2	88.021	0.795	0.945
Man 3	100.000	1.000	1.000
Man 4	97.441	0.978	0.973
Man 5	90.141	0.802	0.983
Man 6	89.894	0.789	1.000
Total	93.656	0.883	0.971

Depth comparison and neural networks classification results are shown in the table 2 and 3. Depth comparison method was tested on a whole data set. *True positive rate* (TPR) represents proportion of correctly detected correct sitting posture and *true negative rate* (TNR) represents proportion of correctly detected wrong sitting posture.

7 Conclusion

In this paper we propose a solution for a real-time sitting posture quality evaluation. We use several techniques to analyze depth and RGB camera streams. To classify user's sitting posture, we detect different parts of a user's body and compare their average depths with a calibrated image. Next method uses computer vision algorithm Histogram of oriented gradients not only for a color image, but also for depth images. Last method computes a region of interest custom image features.

Experiments show that depth comparison method gained the best results. Using calibration image, it is the most invariant method to body proportions and camera placements. Some of the

Table 3. Posture classification using neural networks.

Net name with layers counts	Threshold[%]	Success rate[%]	TPR	TNR
ROI 12-8-2	50	83.688	0.910	0.781
ROI 12-8-2	80	81.915	0.770	0.856
ROI 12-8-2	30	82.624	0.984	0.706
ROI 12-6-2	50	86.879	0.992	0.775
ROI 12-6-2	85	87.943	0.885	0.875
HOG 2304-900-200-2	50	76.241	0.819	0.718
HOG 2304-1200-400-40-2	50	92.553	0.910	0.938
HOG 1728-800-200-20-2	50	81.915	0.844	0.800
HOG-RGB 2268-900-200-2	70	73.338	0.459	0.956
HOG-RGB 2268-1200-300-80-20-2	50	68.905	0.303	0.987
HOG-RGB 2268-1200-300-2	50	73.243	0.713	0.737
HOG-RGB 5680-900-200-2	70	51.739	0.359	0.587
HOG-RGB 1260-900-200-2	70	53.043	0.621	0.373

methods that use neural networks are good, but they are depended on a quality of training samples. In the future, we would like to add other features in an image extraction, e.g. for detecting position of arms. We would like to validate our solution on bigger data set containing even more people with various body types and camera placements.

Acknowledgement: The authors extend their gratitude to Mgr. Lucia Merceková for sharing her professional knowledge of common bad sitting habits and spine related health problems. This contribution was created with the support of the Ministry of Education, Science, Research and Sport of the Slovak Republic within the Research and Development Operational Programme for the project University Science Park of STU Bratislava, ITMS 26240220084, co-funded by the ERDF.

References

[1] Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*. Volume 1., IEEE, 2005, pp. 886–893.

[2] Duthey, B.: Background Paper 6.24 Low back pain, 2013.

[3] Melrose, A.S., Graveling, R.A., Cowie, H., Ritchie, P., Hutchison, P., Mulholland, R.M.: Better Display ScreenEquipment (DSE) work-related ill health data, 2007.

[4] Paliyawan, P., Nukoolkit, C., Mongkolnam, P.: Prolonged sitting detection for office workers syndrome prevention using kinect. In: *2014 11th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, IEEE, 2014, pp. 1–6.

Vehicle Recognition in Image Sequences Using Mobile Devices

Jakub MROCEK*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
mrocek.jakub@gmial.com

Abstract. This paper deals with vehicle recognition in the sequence of images captured using built-in camera of mobile device in real time. This creates specific conditions for solutions. Systems have to be effective as the computing resources of mobile devices are limiting. We proposed and implemented system for Android based on the OpenCV library using cascade classifier to detect vehicles in image. The cascade classifier was created using AdaBoost learning algorithm. Our implemented system runs on mobile device in real time and according to our testing correctly detects more than 93 % of vehicles in real life conditions.

1 Introduction

Image recognition and computer vision in general has many ways of utilization in broad sphere of fields. Therefore, there was huge progress in research and development in this area during past decades. Increasing computing power of mobile devices and new, efficient, methods for solving partial problems allowed us to implement systems capable of detecting vehicles (or any other objects) in real time.

The main sources of motivation for research of vehicle detection are the efforts to create an autonomous car, creating systems assisting drivers, development of parking systems, increasing safety on the road by accident prediction or measuring traffic data such as number of vehicles or their speed.

2 Related Work

A huge variability of vehicles on the market causes difficulties while identifying the vehicle in the image. There are many models of vehicles from different manufacturers what produces a wide variety of shapes and colors. However, there are several key features that can help us detect the vehicle in the picture, as they are typical for all vehicles. Various authors have described different

* Bachelor study programme in field: Informatics
Supervisor: Dr. Peter Lacko, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

methods by which the vehicle can be identified in the image in their studies. However, it may be impossible to achieve accurate results using only a single indicator. Their combinations are therefore often used what results into satisfying results under different conditions.

The basic method used for object detection in an image consists of the two following phases [6]:

Hypothesis Generation In the first phase, the image is scanned quickly (but less accurately), while we try to find areas where searched objects could be possibly located. For these purposes can be used reduced-resolution images, which contain less details, but allow faster processing of the image, while focusing only on the key signs of the presence of the searched object in the image. This phase results into finding such areas in image that could hypothetically contain the searched object (see Figure 1).

Figure 1. During the first phase (hypothesis generation) are found areas, which could hypothetically contain the searched object.

Hypothesis Verification In the second phase are examined in more detail only areas found in the first phase. For every area we are trying to confirm or deny the presence of the searched object in the image. There are several approaches that can be used to confirm the presence of an object using classifiers. Each has advantages and disadvantages and is suitable for different conditions. Therefore, a combination of classifiers is often used to achieve better results. This phase resluts into correctly detected object (see Figure 2).

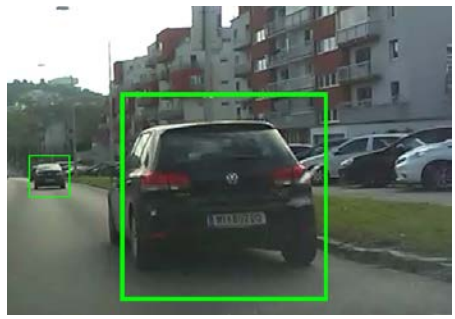


Figure 2. During the second phase (hypothesis verification) is confirmed or denied presence of searched object in evry area found in the first phase what resluts into correctly detected object.

The advantage of this approach is that only areas for which it makes sense are searched in more detail, which brings great savings of computational costs.

2.1 Classifiers Used for Vehicle Detection

Symmetry There is great diversity in vehicle shapes that need to be definitely distinguished from surrounding. But almost in all cases, if seen from behind, the right and left half of vehicle looks identical, they are just mirror-flipped. This fact helps us in many cases to reliably say that object in the image is a vehicle. Drawback of this classifier, however, is high false rate, as any symmetrical (or homogeneous) area is claimed to be a vehicle. Therefore should be this classifier used in combination with other classifiers to distinguish those falsely indicated objects.

Hsieh et al. [2] used symmetry as the main prerequisite for detecting vehicles in the image. In order to eliminate false detection have been used shadows under the vehicle and the horizontal edges of the bumper and the bonnet and the windscreen. The advantage of detecting vehicles based on the symmetry properties is the high success rate even at night scenes, where the tail lights are dominant as they are symmetrical.

It is appropriate, prior to the analysis, to convert the image to gray scale what reduces the amount of processed data. This helps achieve a higher number of frames processed per unit of time.

Shadow One of the other characteristics of the vehicle in the picture is its shadow. In most lighting conditions, there is shaded area near to or under the vehicle. The only exception are situations with very specific light or dark road surface where the shadow can be hardly identified. It would be difficult to achieve reliable results using the shadow as main classifier, as the shadow can be thrown by any other object and any dark area in the picture can be wrongly considered to be a shadow. Therefore can this classifier be better used to confirm vague results gained using some different classifier.

The fact that not only vehicles, but all objects cast a shadow on the road was used to detect obstacles on the road, more specifically other vehicles and pedestrians [4]. However, even in this case this was accompanied by another identifier – a characteristic movement of the object.

Corners and Edges Corners and edges are characteristic for the vehicle, especially from rear view, either horizontal (top, chassis, rear window, bumper, license plate) or vertical (sides of the vehicle, the rear window). Using corners and edges to find vehicle in image is effective mainly for vehicle closer to camera, as those are more significant. Detection of edges and corners can be a good basis for determining the speed of vehicle from a sequence of images, since the vehicle is localized with high precision [3].

2.2 Vehicle Detection Using Cascade Classifier

Face recognition is very similar problem to vehicle recognition, despite the differences in characteristics between the face and the vehicle. Ability to detect faces developed in past years in real time brings many opportunities in many areas and it has received much attention. The basis of this method is in works "Robust real-time face detection" [9] and "Rapid object detection using a boosted cascade of simple features" [8], where the authors Viola and Jones were able to use 3 revolutionary ideas to develop a method for detecting faces in an image in real time (15 frames per second at 700 MHz CPU). The three main ideas that made this possible were: integral image, creating weak classifiers using only a small amount of characteristics with the learning algorithm AdaBoost and the combination of weak classifiers into a complex cascade structure. The method classifies images based on image features and not on the basis of the values of individual pixels. This makes it easier to learn the characteristics of the searched object from the limited set of samples. a key feature of this approach is the high search speed, as the system based on finding features is significantly faster than a system based on finding pixels.

Haar Wavelet Representation First one to use Haar wavelet representation to represent image feature was Papageorgiou et al. in 1998 [5]. It collects statistical data about the object obtained by

analysis of multiple instances of the object. This data can be later used to detect the object in the image. The advantage of such representation is that for the correct detection there is not necessary any knowledge of the specific features or characteristics of the object. This leads to the creation of a general framework for object recognition in the image. Authors of this study successfully used this method to detect faces even in very disturbing scenes. Haar wavelets form a set of basic functions that detect differences in average intensity between regions. This helps to determine which parts of the image are specific and thus typical of the object and is used for learning the features of an object from its instances. Learning of the attribute of the object consists of two phases: a) identification of a small amount of basic functions describing the structure of the class, b) derivation of the precise properties of the object using the classifier. a) We obtain coefficients for every part of the image using different (horizontal, vertical and diagonal) wavelet basis functions. Those coefficients are subsequently normalized. We can determine whether a characteristic is significant or not using the values of the coefficients. It is interesting that each of the functions captures different kind of properties what helps to create a complex picture of the features of the object. b)We derive a small amount of coefficients from the calculated values that statistically describe large amount of original coefficients. These derived coefficients are significant and broadly describe studied object.

Integral Image We can describe features of rectangular areas of the image using Haar Wavelet representation in relation to the whole image (1). Rectangular features can be computed quickly if the image is represented using median values. Such representation is called integral image. Integral image of the point x, y is the sum of the values of previous pixels that lie at the top left of the point. Calculating the value of the integral image of the point x, y:

$$ii(x, y) = \sum_{x'<x, y'<y} i(x', y') \tag{1}$$

Where ii is the integral image and i is the original image.

AdaBoost AdaBoost is an algorithm developed to enhance the performance of other learning algorithms. Its essence is the combination of weak (low specific) classifiers into a strong one (very specific). A typical example of its usage are situations where we have a lot of rules that describe an object. These rules are not very specific and we can not determine with high probability that an object is described precisely by these rules using only a few of them. The more weak classifiers we have, the higher percentage of correct determinations. When we have a sufficiently large number of rules describing the object, the determining success rate is sufficient for practical usage [1]. The great advantage is that the weak classifiers does not need to be accurate at all, generally only error rate below 50 % of cases is needed. The weight of individual rules depends on their accuracy, respectively on their error rate. AdaBoost algorithm is fast, stable and easy to program.

2.2.1 Creating Classifier from Image

We need a lot of weak classifiers to create the final strong classifier, which we will be able to use to determine accurately whether there is a vehicle in the image. Those can be obtained by splitting image into many smaller ones. Each part of the image becomes classifier. Afterwards it is necessary to test each sector test find out how accurate it is to determine whether it is a vehicle or not. We count errors for each area – classifier. Consequently, according to the error rate is weight assigned to individual classifier in order to increase the impact of those achieving lower error rate. This is followed by further testing using revised classifiers and reallocation of weights according to a new error rates. This process is repeated several times to produce accurate final classifier.

2.2.2 Combination of Classifiers into Cascade

The creation of a single strong classifier might bring impressive results, nevertheless, even these may not be sufficient for real life. The need for greater accuracy increases computational complexity heavily. But it is unnecessary to crawl with great accuracy through the parts of the image in which there is only a very low probability of occurrence of the searched object. To increase the efficiency it is better to search in bigger rectangles using less accurate classifier first. This allows us to quickly determine whether there is a real chance that the area contains the search object. If so, this area is left for further, more precise search. If not, there is no need to consume more computing resources to accurately scan the area. In this case, the area is simply excluded from other, more specific searches. The same principle is then used for areas which stayed. This results in a significant speed up the whole process. To create such a cascade classifier we need to achieve following: a) to be able to create fast, but not necessarily too accurate classifier for the first level of the cascade. b) to be able to create an accurate classifier that is capable of deciding about the presence of the searched object in the image with very high probability, even at the expense of greater computational complexity. This will be used for the last level of the cascade. In all levels of cascade is accuracy as well as complexity of the calculation increased gradually. To create such, gradually changing, classifiers can be well used AdaBoost classifier described above with changing threshold.

3 Our Approach

Our approach was largely inspired by the face detection method as it was designed by authors Viola and Jones. We decided to train our own cascade classifier, which would be capable of detecting vehicles in the image using the above described method.

We trained cascade classifier using approximately 1800 positive and 1000 negative images. Some images were used from computer vision datasets, but to increase performance and accuracy we included images taken by mobile device as well. These were taken on our public roads and show cars commonly used in our region. The created cascade classifier was then used in Android application we implemented using the OpenCV library.

4 Experimental Results

The implemented system running on a mobile device equipped with 1 Ghz processor was capable of processing 9–13 frames per second according to complexity of the scene. There are some enhancements which should increase the performance, but we consider achieved frame rate to be a good result for first prototype of the proposed system.

We tested our prototype in real-life conditions. Android device with installed system was placed in a vehicle moving at the speed range from 15 to 90 km/h. While using the built-in camera, our prototype was detecting other vehicles in front of it for approximately 15 minutes. During the testing we counted all the vehicles which appeared in the field of view as well as those that were correctly detected. There were often multiple vehicles in the frame and all were considered. Detection percentage reached 93,47 %.

We noted false positives in some frames. In most cases these were caused by visually similar objects to vehicles as road signs or advertisements. The false positive rate was 4,81 %.

5 Conclusion and Future Works

We consider the results achieved with prototype to be satisfying, but still plan to improve it's performance and accuracy.

The next big step will be determining speed of detected vehicle. We suggest to use inverse perspective mapping [7] as it seems to be capable or real time running even on mobile devices and

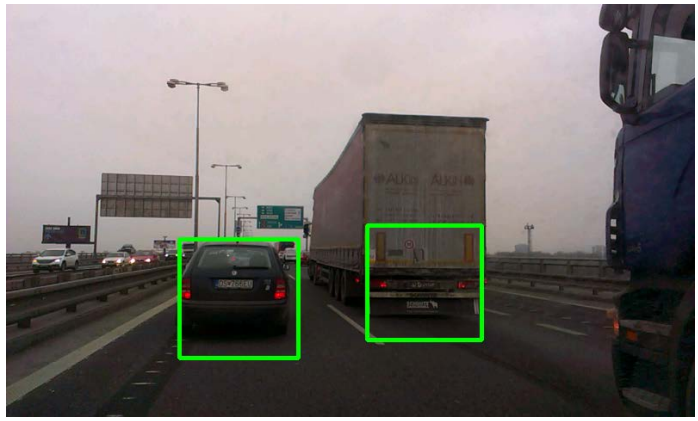


Figure 3. Screenshot from our application running on a mobile device in real life conditions. Application is capable of detecting different types of vehicles.

is perfectly suited for our conditions, where the mobile device with camera is placed in vehicle monitoring scene in front of it.

Acknowledgement: This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG 1/0752/14.

References

[1] Freund, Y., Schapire, R.E.: A Short Introduction to Boosting. In: *In Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence*, Morgan Kaufmann, 1999, pp. 1401–1406.

[2] Hsieh, J.W., Chen, L.C., Chen, D.Y., Cheng, S.C.: Vehicle make and model recognition using symmetrical SURF. In: *Advanced Video and Signal Based Surveillance (AVSS), 2013 10th IEEE International Conference on*, 2013, pp. 472–477.

[3] Kumar, K., Chandrakant, P., Kumar, S., Kushal, K.: Vehicle Speed Detection Using Corner Detection. In: *Signal and Image Processing (ICSIP), 2014 Fifth International Conference on*, 2014, pp. 253–258.

[4] Mori, H., Charkari, N.: Shadow and rhythm as sign patterns of obstacle detection. In: *Industrial Electronics, 1993. Conference Proceedings, ISIE'93 – Budapest., IEEE International Symposium on*, 1993, pp. 271–277.

[5] Papageorgiou, C., Oren, M., Poggio, T.: A general framework for object detection. In: *Computer Vision, 1998. Sixth International Conference on*, 1998, pp. 555–562.

[6] Sun, Z., Bebis, G., Miller, R.: On-road vehicle detection: a review. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 2006, vol. 28, no. 5, pp. 694–711.

[7] Tuohy, S., O'Cualain, D., Jones, E., Glavin, M.: Distance determination for an automobile environment using Inverse Perspective Mapping in OpenCV. In: *Signals and Systems Conference (ISSC 2010), IET Irish*, 2010, pp. 100–105.

[8] Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*. Volume 1., 2001, pp. I–511–I–518 vol.1.

[9] Viola, P., Jones, M.: Robust real-time face detection. In: *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*. Volume 2., 2001, pp. 747–747.

3D Object Recognition Based on Local Descriptors

Marek JAKAB*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
marko.jakab@gmail.com

Abstract[1]

In our research, we propose an enhanced method of 3D object description and recognition based on the method of local descriptors using RGB image along with depth information (D) acquired by depth sensors. Currently, several local descriptors include depth information to increase the robustness of the descriptor as an extension of the single descriptor vector (for example CSHOT or BRAND). Overall, the use of depth information proved to have increased effect in accuracy. However, the task of object recognition becomes harder with the increased dataset as the recognition time will increase.

Hence, we propose a novel local depth descriptor (DD) that will enter the part of the decision-making process as a standalone descriptor in combination with the SIFT descriptor. Proposed depth descriptor is responsible for the pre-selection of the objects in the dataset which will pass to the second, accurate recognition. Because of the 4 dimensional vector the depth descriptor is currently formed of, compared to the 128 dimensional vector in the SIFT descriptor, it can significantly reduce the recognition time and increase the number of objects we are able to recognize in real time.

While the depth descriptor takes first part in the object recognition process, we need to maintain its robustness and invariance at least to the level of all consecutive processes. Therefore all the values in the descriptor vector are based on the statistical data unrelated to key point position. This will achieve the rotation and small perspective transformation invariance to the descriptor. Scale invariance is given by changing the size of the descriptor pattern related to the distance of actual key points.

The pattern used in the descriptor is formed from the 4 triangles in a star pattern. Three edge points from each triangle are used with their depth information to get normal vector as they define the plane. Overall 4 normal vectors are computed which are later used in the process of creating the descriptor vector.

The results of the paper have been published at SPIE – Electronic & Imaging conference held on February 8–12, 2015 in San Francisco, California.

Acknowledgement: This research has been supported by a grant VEGA 1/0625/14.

---

* Master degree study programme in field: Applied Informatics
Supervisor: Dr. Vanda Benešová, Institute of Applied Informatics, Faculty of Informatics and Information Technologies STU in Bratislava

[1] Full paper available in printed proceedings, pages 231–236.

Hierarchical Superpixel-Based Saliency Model

Patrik POLATSEK*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
patrik.polatsek@gmail.com

Abstract. Prediction of human visual attention is more and more frequently applicable in computer graphics, image processing, human-computer interaction and computer vision. Saliency models implement bottom-up factors of visual attention and represent the conspicuousness of a given environment using a saliency map. Standard hierarchical saliency methods do not respect the shape of objects and model the saliency as the pixel-by-pixel difference between the centre and its surround. The aim of our work is to improve the saliency prediction using a superpixel-based approach whose regions should correspond to object's borders. In this paper we propose a novel saliency method that combines a hierarchical processing of visual features and a superpixel-based segmentation. The proposed method is compared with existing saliency models and evaluated on a publicly available dataset.

1 Introduction

Our environment contains many objects which provide us a huge amount of visual information. The human brain has limited computational capacities, due to which it cannot process all incoming visual data. Thus *attention* provides mechanisms of reducing and selecting important information. Visual attention helps us to decide where to move our eyes and which parts of a scene should be deeper processed [2].

There are various factors that influence our attention. We can divide them into two main categories [5]:

1. stimulus-driven *bottom-up* factors such as colour, contrast, orientation, texture and movement,

2. and goal-driven *top-down* factors involving prior knowledge, experiences, expectations, tasks or goals.

Bottom-up attention is based on visual characteristics of a scene which automatically draw the attention. It is related to the term *saliency*. Saliency is the vividness of a stimulus which stands out relatively from its neighbour.

---

* Master study programme in field: Information Systems
Supervisor: Dr. Vanda Benešová, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava.

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 237–242.

In recent decades scientists have studied mechanisms of human attention to determine regions of interest. Visual attention modelling has a wide range of applications such as computer graphics, image processing, human-computer interaction, psychology, neurophysiology and computer vision.

2 Related Work

According to the type of processing, we divide attention models on bottom-up, top-down and those that combine both processes. The majority of them models bottom-up attention. The result of such models is a *saliency map* which is a topographic representation of visual saliency of a scene [2].

Hierarchical (cognitive) models are biologically inspired models based on hierarchical decomposition of visual features inspired by *Feature Integration Theory* (FIT) [11]. According to the theory, in early and parallel *pre-attentive* processing a scene is analysed to identify individual features. Within the second, *focused* attention phase various features are combined and integrated to perceive whole objects.

One of the most known bottom-up hierarchical model is presented in [7]. It extracts three visual features – colour, intensity and orientation. The model uses the Opponent-Process Theory of Colour Vision based on two opponent-colour mechanisms: red-green and blue-yellow. The characteristics of texture and local orientation are obtained using the Gabor kernel. This model creates *Gaussian pyramids* for red, green, blue, yellow, intensity channel and local orientations. The structure of retinal ganglion receptive fields is characterised by *center-surround* organisation. This model achieves center-surround operations as the point-to-point difference between finer scales of Gaussian pyramid representing the center and coarser scales representing the surroundings that leads to multiple *feature maps*. Normalised feature maps are combined into three *conspicuous maps* for intensity, colour and orientation and finally into a single saliency map.

Bayesian models [13] are probabilistic frameworks which combine a bottom-up saliency with the effects of prior visual experience.

Decision-theoretic models [4] are based on the theory known as *discriminant saliency* selecting optimal attributes that most distinguish a visual class of interest from the other classes.

Information-theoretic models [3] are based on the theory which assumes that saliency results in the maximum information sampled from a given environment.

Graphical models use graph-based computations to create a saliency map. Nodes of a graph present a set of variables and edges their probabilistic dependencies. an eye movement sequence treated as a time-ordered sequence is modelled in [10] using a *Markov model*.

Spectral analysis models process images in the frequency domain instead of the spatial domain. The model mentioned in [6] is based on the *spectral residual* obtained as the difference between the original and smoothed version of the log spectrum.

Pattern classification [9] models use supervised machine learning algorithms to learn the visual attention from eye-tracking data or labelled salient regions.

Reinforcement learning models [8] predict the saliency using the reinforcement learning algorithm.

3 Proposed Algorithm

We introduce a novel method called *Hierarchical Superpixel-Based Saliency Model* for the detection of bottom-up saliency.

In order to at least partially cover the focused phase of attention that includes the integration of visual features to objects, we implement a superpixel-based saliency in our model instead of a simple pixel-by-pixel-based difference of Gaussian pyramid layers proposed in [7]. Our model is also a hierarchical saliency model based on FIT that integrates intensity, colour and orientation.

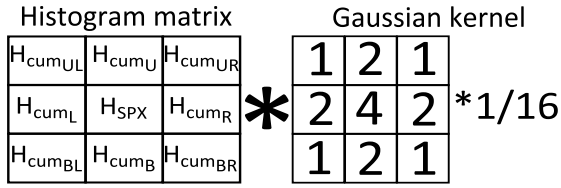


Figure 1. Superpixel Gaussian convolution.

A *superpixel* represents a visually coherent region which can better correspond to object contours than a rigid structure of pixels. The usage of superpixels is the primary difference between our model and the standard hierarchical model in [7].

The proposed solution segments input images using *Simple linear iterative clustering* (SLIC) [1] algorithm. SLIC is performed twice in the model with two different region sizes of 15 and 30.

3.1 Superpixel Gaussian Pyramid

Due to the usage of superpixels, the standard algorithm for Gaussian pyramid is replaced by our superpixel version. Each pyramid layer consists of a superpixel map representing the locations of all superpixels and a set of superpixel histograms.

Within the first pyramid layer 1D superpixel histograms are constructed using one of three visual features. In order to create the next layers, we have to downsample the superpixel map to the half size.

Then we search neighbours to all superpixels in this resized superpixel map. The neighbour assignment procedure processes superpixel borders per pixel. Each border pixel is classified into one of the following categories based on its location to the analysed superpixel: left (L), right (R), upper (U), bottom (B), upper-left (UL), bottom-left (BL), upper-right (UR) and bottom-right (BR). Within each category, neighbours are characterised by a weight which depends on the boundary length with the superpixel. Longer the mutual boundary is, higher weight is assigned to the neighbour.

After the processing of the whole superpixel neighbourhood, we can build a histogram matrix of size 3×3. The center element corresponds to the analysed superpixel histogram H_{SPX}.

All 8 location categories are presented with a cumulated histogram H_{cum_i} defined as the weighted sum of all neighbour histograms connected to the category.

The rest of the histogram matrix is build using the 8 cumulated histograms of all location categories. Each cumulated histogram is assigned to the position in the matrix depending on the category name, for example the upper-left cumulated histogram takes place in the first (upper-left) position of the matrix. The histogram matrix is finally convolved with the discrete Gaussian kernel (see Figure 1). In case of convolution at image borders where empty location categories without any neighbours may occur, the cumulated histogram of such categories equals to the histogram of analysed superpixel H_{SPX}.

In order to create a pyramid layer this procedure is repeated for all superpixels.

To produce the rest of Gaussian pyramid layers, the whole process is iteratively performed with the half size of an input superpixel map.

3.2 Superpixel Feature Processing

After the segmentation of an input image into superpixels using *SLIC* algorithm, our saliency model processes subsequently all features. For each feature it represents individual superpixels by a histogram. The superpixel map and the histogram set enter in the iterative process of the *superpixel Gaussian pyramid* as the first layer.

After the creation of the whole Gaussian pyramid, our model compares center and surround layers of the pyramid per pixel. In order to achieve the center-surround differences, we find

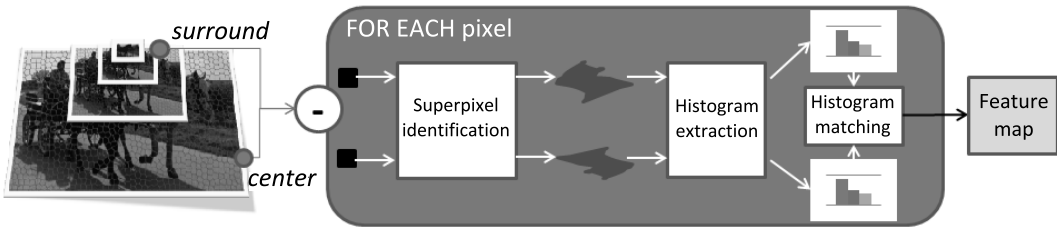


Figure 2. Difference between center and surround scales of superpixel Gaussian pyramid.

superpixels which contain compared pixels at the center and surround scale. Then we measure the similarity between the histograms belonging to these selected superpixels using a *histogram matching algorithm*. Such a difference between the layers produces a *feature map* (see Figure 2).

3.2.1 Intensity

In order to analyse the intensity feature, an image is simply converted to *grayscale*. Superpixels are described by 1D histograms with 128 bins. a value of each pixel in the resulting feature map is computed using the *correlation* method as the follows:

$$FM_I(x,y) = 1 - abs(d_{correl}(H_c(x,y), H_s(x,y)), \tag{1}$$

where d_{correl} is a correlation coefficient, $H_i(x,y)$ is a histogram of a superpixel at the scale i which contains a pixel at position $[x,y]$ and subscripts c and s denote a center and a surround scales.

3.2.2 Colour

An input image is at first converted to 4-channel RGBY colour space for colour feature processing. Each colour channel is defined by its 1D histogram with 128 bins. The process implements the *Opponent-Process Theory*. The center-surround is expressed as the difference between *mean colour values* of compared superpixels for both opponent pairs:

$$FM_{RG}(x,y) = norm(abs(\max(H_{diff_R}) - \max(H_{diff_G}))), \tag{2}$$

$$FM_{BY}(x,y) = norm(abs(\max(H_{diff_B}) - \max(H_{diff_Y}))), \tag{3}$$

where $norm$ normalises values within the interval $\langle 0,1 \rangle$, $\max$ is the most frequently colour and $H_{diff_{COL}}$ is a difference of histograms of colour channel COL at a center c and a surround s scale defined as $H_{diff_{COL}}(x,y) = abs(H_{COL_c}(x,y) - H_{COL_s}(x,y))$.

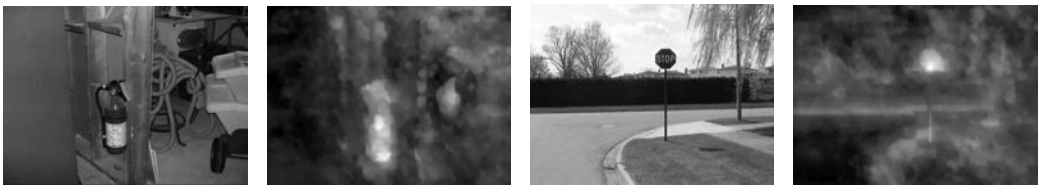
3.2.3 Orientation

The processing of orientation feature starts with the image conversion to *grayscale* colour space. Each superpixel is characterised with a *histogram of oriented gradients* with 9 bins. Orientation differences are computed with the same *correlation*-based method as in intensity feature maps:

$$FM_O(x,y) = 1 - abs(d_{correl}(H_c(x,y), H_s(x,y)). \tag{4}$$

3.3 Saliency Map

Extracted feature maps are combined into 3 *conspicuous maps* of intensity, colour and orientation for each used region size in SLIC algorithm. After their normalisation, they are linearly combined to create a single *saliency map* (see Figure 3).



(a) Source image. (b) Saliency map. (c) Source image. (d) Saliency map.

Figure 3. Hierarchical superpixel-based saliency map.

4 Evaluation

In order to evaluate the novel method we have implemented our superpixel-based model as well as a standard hierarchical model inspired by [7] in C++ using OpenCV library. Both models were tested on a publicly available Toronto dataset [12] (120 images, 20 subjects).

In order to measure the accuracy we use our own method – *maxima distance*. It measures the distance between the most salient location on a saliency map and the location with the highest density on a fixation map. In order to compute the distance, we take iteratively into account together with the first maximum value of the fixation map also the next highest value as follows: $d_{MAX_i} = \min_i \left(d(\max(SM), \max_i(FM))\right)$, where SM is a saliency map, FM is a fixation map and $\max_i$ is the i^{th} maximum value. The results are visualised in Figure 4.

If the highest values of saliency and fixation maps are only considered, the total Euclidean distance produced by our model is 18250 px what is very similar to the results of a standard hierarchical model – the total distance of 18261 px.

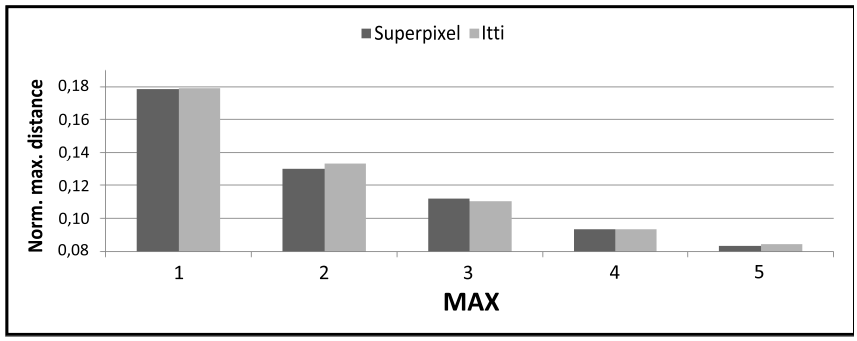


Figure 4. Normalised maxima distance.

To compare the models, we have also used standard methods such as the *similarity metric* defined as $S = \sum_x \min(SM(x), FM(x))$ and the *Kullback-Leibler (KL) divergence* computed as $KL_{div} = \sum_x FM(x) * \log\left(\frac{FM(x)}{SM(x)+\epsilon} + \epsilon\right)$, where ϵ is a small constant.

The results of all evaluation methods produced by our model are very close to the results of a standard hierarchical model as shown in Table 1.

Wrong classification of the most salient location is often the result of the absenting top-down attention in this hierarchical model. The top-down part of our saliency can affect the conspicuousness of objects more significantly in such complex scenes than in simple ones. It is focused in most cases on objects such as texts, human faces or traffic signs that may not be salient in terms of our model.

Despite of that, there are scenes where our model outperforms the standard one. The main reason is the superpixel segmentation of an input which can correspond to object edges.

Table 1. Comparison of a standard hierarchical model based on [7] and a novel model on Toronto [12].

Model	Max.dist.	Similarity		KL-div.	
	MAX_1	*Avg.*	*Med.*	*Avg.*	*Med.*
Standard	18261 px	0.3969	0.3992	1.1649	1.1484
Superpixel	18250 px	0.3939	0.3979	1.1870	1.1334

5 Conclusion and Future Work

In this paper we have presented a novel saliency model that integrates a hierarchical and a superpixel-based approach. The main benefit of superpixels in our model is respecting the shape of objects in the visual attention processing. Our model achieves on a publicly available dataset very similar results to a standard hierarchical model based on [7].

We will further focus on dynamic stimuli such as a motion contrast that significantly influence our attention. Combining static and dynamic attentional factors we create a complex spatiotemporal model that may better predict human attention.

References

[1] Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Susstrunk, S.: SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2012, vol. 34, no. 11, pp. 2274–2282.

[2] Borji, A., Itti, L.: State-of-the-Art in Visual Attention Modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, vol. 35, no. 1, pp. 185–207.

[3] Bruce, N.D.B.: *Saliency, Attention and Visual Search: An Information Theoretic Approach*. PhD thesis, Canada, 2008, AAINR45988.

[4] Gao, D., Vasconcelos, N.: Discriminant Saliency for Visual Recognition from Cluttered Scenes. In: *NIPS*, 2004.

[5] Goldstein, E.: *The Blackwell Handbook of Sensation and Perception*. Blackwell Handbooks of Experimental Psychology. Wiley, 2008.

[6] Hou, X., Zhang, L.: Saliency Detection: A Spectral Residual Approach. In: *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on*, 2007, pp. 1–8.

[7] Itti, L., Koch, C., Niebur, E.: A model of saliency-based visual attention for rapid scene analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 1998, vol. 20, no. 11, pp. 1254–1259.

[8] Jodogne, S., Piater, J.H.: Closed-loop Learning of Visual Control Policies. *J. Artif. Int. Res.*, 2007, vol. 28, no. 1, pp. 349–391.

[9] Kienzle, W., Franz, M.O., Schölkopf, B., Wichmann, F.A.: Center-surround patterns emerge as optimal predictors for human saccade targets. *Journal of Vision*, 2009, vol. 9, no. 5, p. 7.

[10] Salah, A., Alpaydin, E., Akarun, L.: A selective attention-based method for visual pattern recognition with application to handwritten digit recognition and face recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 2002, vol. 24, no. 3, pp. 420–425.

[11] Treisman, A.M., Gelade, G.: A Feature-Integration Theory of Attention. *Cognitive Psychology*, 1980, vol. 12, pp. 97–136.

[12] Tsotsos, J.K., Bruce, N.D.B.: Saliency based on information maximization. In: *Advances in Neural Information Processing Systems 18*, MIT Press, MIT Press, 2006, pp. 155–162.

[13] Zhang, L., Tong, M.H., Marks, T.K., Shan, H., Cottrell, G.W.: Sun: A Bayesian framework for saliency using natural statistics. *Journal of Vision*, 2008.

Software Engineering

Transformation of UML Combined Fragments from 2D to 3D

Richard BELAN*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
xbelanr1@is.stuba.sk

Abstract. Software modeling and visualization are a large part of software's life cycle. Despite the large improvement in computer graphics and numerous attempts the three dimensional software modeling and visualization were not yet adopted. This paper deals with the way of using the third dimension with purpose to improve the work experience with large, complex diagrams which are hard to fully understand and are hard for managing, editing and refactoring. The main focus of this paper is on combined fragments in sequence diagrams. We propose a way how to transform combined fragment from 2D to 3D with the use of layers and animating it's operands in third dimension.

1 Introduction

Two dimensional UML is industry standard for visualizing and modeling software. Although it is very popular and common in the process of software development it has limitations with large, complex diagram models of huge enterprise application and also with their comprehension. Complex diagrams are serious limitations for user comprehension and in case of huge applications it's hard to find scheme in the diagrams. One of the major problems is that even a large monitor can show only a small portion of the diagram as seen in Figure 1A, where the black box shows what can a FullHD monitor display. But even the part which is displayed on FullHD monitor as in Figure 1B is still very complex and missing important informations, connections to fully understand the displayed part of the diagram. Also managing, repairing, editing and refactoring such large diagrams is very hard. This paper propose a way how to increase comprehension of complex diagrams with additional dimension and method of layering informations, focusing on informations encapsulated in combined fragments in three dimensional UML diagrams. It also shows how fragments and nested fragments can help user comprehension.

* Bachelor study programme in field: Informatics
Supervisor: Dr. Ivan Polášek, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 245–250.

Figure 1. Example of a large complex diagram. Black box showing what can be seen on a FullHD monitor.

2 Related Work

In the past 25 years there were numerous attempts to visualize and model software in three dimensions. These trials can be divided into three groups (*mental models*, *three dimensional objects* and *layering*) by the approach they chose to use the third dimension.

Idea of using mental models was first proposed by H. Koike in 1993 [6]. The user creates different mental models from each diagram, then he reconstructs them to one which satisfies all constraints [6]. Having one diagram on front side of a cube, second diagram on rear side of the cube and in space between them showing connections and flow between both diagram creates one 3D diagram. The advantage of this approach is that if the final diagram is easy for comprehension, then the user cognitive load is reduced twice due to necessity of creating two times less mental models. Determinate how to join two diagrams and make information easily readable and not confusable at two complex diagrams are the biggest disadvantages of modeling using mental models.

Using 3D objects in UML was first proposed by M. Gogolla et al. in 1999 [4]. They introduced 3D UML notation, for Class and Sequence diagrams and claimed that 3D objects can help user comprehension and also that sorting nodes with connections by some criteria and pushing the uninteresting ones in background is possible only in 3D [4]. Improving the comprehension with 3D objects was also presented in a paper by P. Irani in 2000 [5]. He not only given the UML notation a 3D shape, but also changed the representation of UML notation with geons. Geons are small 3D geometrical primitives, introduced by I. Biederman in 1987 [1]. Biederman claimed that objects made of geons are much easier for human comprehension [1]. Irani evaluated geon diagrams resulting with 11 person from 15 identifying faster geon diagrams than UML diagrams and 26 perso nfrom 35 recalling more geon than UML diagrams [5].

The last approach first introduced by S. Kent et al. in 1998, proposed separating information by some criteria or state in time and showing them on layers in 3D space [3]. The idea of layers was expanded by one of the last contributors to 3D software modeling P. McIntosh and J. Pilgrim et al [7,8]. McIntosh focused on UML State diagram and also rejected the previous ideas of combining more diagrams into one rather showing large complex diagrams of one type spread on layers in space. The same idea of using layers for modeling complex diagrams, but using own rather then UML notation was proposed by Pilgrim et al. As mentioned in introduction the large complex diagrams can be divided into layers and this way made easier for comprehension. Although all these approaches have their pros and on some particular diagrams they might be even better as similar diagram in two dimensions, they were not accepted by the industry so far.

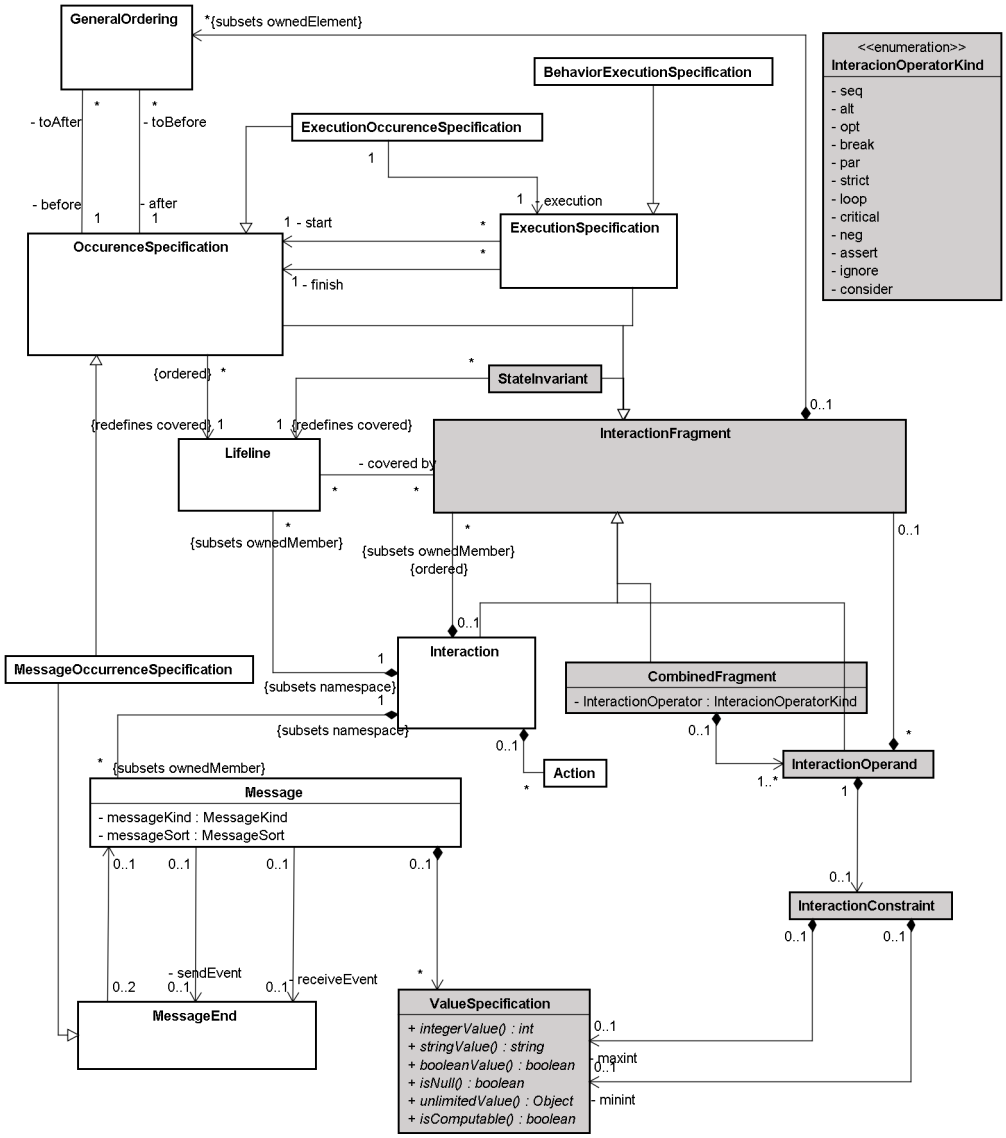


Figure 2. Metamodel we created and used for storing data of 3D sequence diagram and 3D combined fragment.

3 3D Combined Fragment in UML Sequence Diagram

3.1 Evolution of 3D UML Sequence Diagram

MacIntosh mentions 6 major reasons why three dimensional software modeling have not been adopted so far: *Capabilities of computer graphics, Evolution, Open Standards, Perceived Return on Investment, Software Changes, Desktop Computer Limitations* [7]. In our prototype we tried to be consistent with all these reasons. Because of improvement in graphics and desktop computers and their capability of advanced three dimensional modeling, we do not need to worry about hardware limitations. Bearing in mind open standards and evolution we tried to improve a working and

accepted environment rather than providing a truly new one as it might not be accepted by users because of their choice to use the traditional one, with which they are used to work. Traditional modeling in 2D is drawing diagrams on one layer. By multiplying the layers in three dimensions we can divide informations, make complex diagrams easier to user comprehension and still keep the traditional concept in the process of evolution. We chose UML as our standard since UML 2.4.1 already supports three dimensional modeling [2] and also because UML gained a wide acceptance and is now an industry standard.

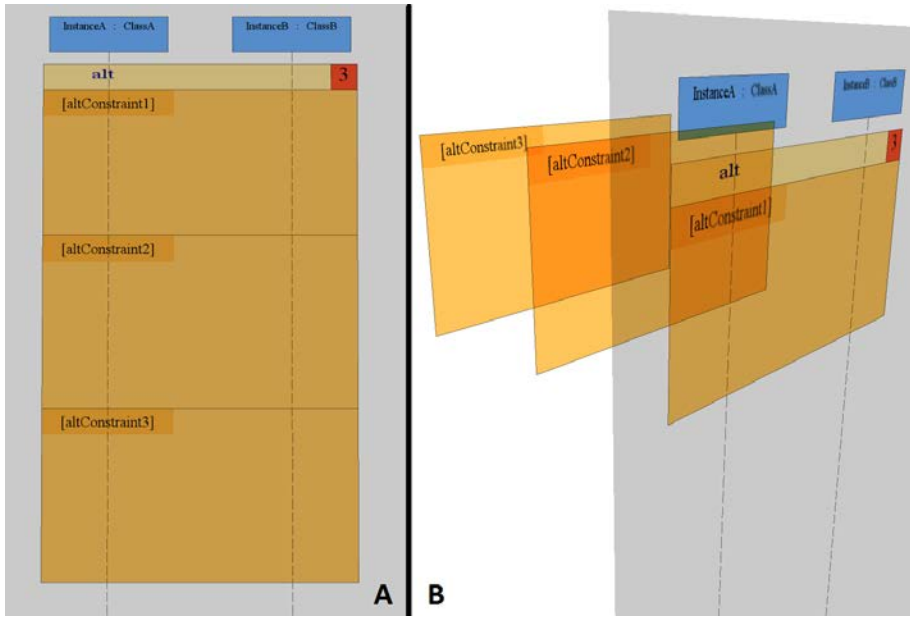


Figure 3. 2D combined fragment on the left and 3D transformed equivalent on the right.

3.2 Transformation of Combined Fragments from 2D to 3D

For creating the fragment in three dimensions we used OGRE graphic engine and C++. Our fragment contains information about it's graphics, position in the world, number of operands (in the dimension of interactions on y axis), number of lifelines (in the dimension of lifelines on x axis), condition, constraint and many other important data. To store the data for the 3D combined fragment and also for the whole 3D sequence diagram, we created a metamodel. This metamodel in Figure 2 we created consistent with UML notation, semantics and model elements described in OMG UML Superstructure and Infrastructure 2.4.1 [2] The metamodel we divided into two parts, in Figure 2 differentiated by color. White color is used in part which we created to store general diagram data, layers and also 3D sequence diagram data. Grey colored part in Figure 2 we created to represent data of 3D combined fragment and it's operands, constraints, type and other data directly relevant to fragment. This way our representation of 3D combined fragment is consistent with UML open standard and also stores everything important to reuse, store and recreate the fragment.

In the process of creating a fragment, we first choose the layers where the fragment will be present. Then by mouse *onClick* and *onRelease* we define the position of the fragment and also the size of it's operand. Next step is defining the fragment's type, constraint and number of operands. There are two kinds of fragments. The simple fragments like loop, opt, etc with only one operand and the more complicated ones with at least one operand like alt and parallel. Last step before drawing the fragment is checking whether the fragment is nested in another fragment or not. Every nested fragment is shown few pixels above the fragment they are nested in. Fragments are nested in three

dimensional space using a Composite design pattern. Our main effort is on fragments like alt and parallel with usually more than one operand. These types of fragments we show like in 2D on same level each next operand under the previous one as in Figure 3A with possibility to animate fragment's operands each behind the previous in three dimensions like in Figure 3B By this transformation of fragment's operands we can display alternatives like nobody had displayed before and we do not know the effect and the further possibilities of our approach yet.

3.3 Transformation algorithm

Each fragment has in it's right corner a red bookmark button as in Figure 3A in right corner. Bookmark button contains a number which indicates how many operands the corresponding fragment has. Only the fragments with more than one operand are transformed into three dimensions. Transformations begins with activating the bookmark button. All Operands corresponding to the fragment are collected with every information they contain, except the first operand which is not being transformed. Also every layer is collected and sent to the function. Then the operands are transformed into three dimensions and layers shifted back to make space for operands. This is made by algorithm in Listing 1. With deactivating the Bookmark button, the animation roll backs from the 3D in Figure 3B back to the original 2D position in Figure 3A.

Listing 1. Transfomation of fragment's operands from 2D to 3D.

```
transform(InteractionOperand[] operands, Layer[] layers, int direction)
 //transform fragments
 int N = 50*direction; //N is a~value of the offset of repositioning
 for each operands do
  SceneNode node = operands[i]->getGraphicsData->getSceneNode();
  Vector3 start = node->getPosition();
  Vector3 middle = Vector3(start.x, start.y, start.z+i*N);
  Vector3 end = Vector3(start.x, start.y-i*operandHeight, start.z+i*N);
  Animation animation = new splineInterpolationn(start, middle, end);
  animation->animate();
 //transform layers
 for each layers do
  SceneNode node = layers[i]->getGraphicsData->getSceneNode();
  Vector3 start = node->getPosition();
  Vector3 middle = Vector3(start.x, start.y, start.z+fragments->size*N);
  Vector3 end = Vector3(start.x, start.y, start.z+fragments->size*N);
  Animation animation = new splineInterpolation(start, middle, end);
  animation->animate();
```

4 Evaluation

To verify, test and evaluate the prototype we created a set of test cases. These test cases defined some basic diagrams and diagram structures which we tried to represent in our 3D sequence diagram. During the simulation of these test cases in our prototype we focused not only on verifying the correctness of the notation, but also whether the prototype responds correctly and either displaying the 3D transformed combined fragment in sequence diagram was proper. Some minor problems were detected, but after identifying them we fixed them immediately.

We have performed some preliminary and initial tests given the functionality of displaying, but our prototype is not yet completely tested. In future we prepare to do further tests and evaluations, but because there's nothing similar in this field, we have nothing to compare with our prototype and evaluate the pros and cons of each approach. To evaluate our prototype despite having nothing to compare it with we will need to use relevance judgment technique. First we will create some basic metrics for the technique and then evaluate them in the context of clarity and simplification.

5 Conclusion and Future Work

In this paper we discussed the problem which occurs with large complex diagram models for huge enterprise applications like in Figure 1. We proposed a way how to improve complex diagrams by layering them with the main focus on *combined fragments* in sequence diagrams. Having in mind the 6 reasons of adopting three dimensional modeling by MacIntosh we transformed fragments with more than one operand from 2D to 3D with availability to transform it back anytime.

In future we want to insert the 3D fragment into other UML diagrams. Then the next step will be using these diagrams for source code generation and also vice versa to reverse engineering, generating diagrams from source code. With these two things done we can achieve round-trip engineering by synchronizing source code, actual model and documentation.

Acknowledgement: This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG1/1221/12.

References

[1] Biederman, I.: Recognition-by-components: A theory of human image understanding. *Psychological Review*, 1987, vol. 94, pp. 115–147.

[2] Booch, G., Jacobson, I., Rumbaugh, J., eds.: *OMG Unified Modeling Language Specification Version 2.4.1*. 2.4.1 edn., 2011.

[3] Gil, J., Kent, S.: Three dimensional software modelling. In: *ICSE '98 Proceedings of the 20th international conference on Software engineering*. ICSE Proceedings, 1998.

[4] Gogolla, M., Radfelder, O., Richters, M.: Towards three-dimensional representation and animation of UML diagrams. In: *UML'99 Proceedings of the 2nd international conference on The unified modeling language: beyond the standard*. UML Proceedings, 1999.

[5] Irani, P., Ware, C.: Diagrams based on structural object perception. In: *AVI '00 Proceedings of the working conference on Advanced visual interfaces*. AVI Proceedings, ACM New York, 2000, pp. 61–67.

[6] Koike, H.: The role of another spatial dimension in software visualization. *ACM Transactions on Information Systems (TOIS)*, 1993, vol. 11.

[7] McIntosh, P., Hamilton, M., Schyndel, R.v.: X3D-UML: enabling advanced UML visualisation through X3D. In: *Web3D '05 Proceedings of the tenth international conference on 3D Web technology*. Web3D Proceedings, ACM New York, 2005, pp. 135–142.

[8] Pilgrim, J.v., Duske, K.: Gef3D: a framework for two-, two-and-a-half-, and three-dimensional graphical editors. In: *SoftVis '08 Proceedings of the 4th ACM symposium on Software visualization*. SoftVis Proceedings, ACM New York, 2008, pp. 95–104.

Source Code Review Recommendation

Matej CHLEBANA*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava 4, Slovakia
matej.chlebana@gmail.com

Abstract. In this paper we introduce method, which supports utilities and ways for analysing developers' activities during the development of software products. We monitor developing processes, build process models and collect annotations for these processes. We compare existing annotated processes with the newly created processes. According to annotations of a process, we determine its character. We consider the processes with high similarity as the processes with the same character. If a new process has error character, we create a recommendation for the code review for the code which was written in this process. This way brings more effective source code reviews.

1 Introduction and related work

Systems for managing and processing information systems create metadata that help them in analyzing the managed content. One of the types of metadata are the information tags that contain structured information associated with a particular piece of content [1]. Information tags can be used to store information about the activities of developers and the properties of source code artifacts in the domain of information systems development. Information tags allow abstracting from computationally complex direct processing code. Their use can thus help to solve the problems connected with the development of information systems such as optimization of the development process by supporting the assessment of source code.

There are several approaches for monitoring and analysing developers' activity. Some approaches are engaged in measuring of developers' activity which analyse mutual collaboration among developers in team [2]. Many from existing metrics take into consideration only data which have been obtained in the discrete timestamps from changes in versions. According to these metrics we are not able to get information about common problem-solving process. Therefore Meneely created annotations to be used in ticketing systems: author of solution and approver of solution. Meneely moreover analysed discussions for tickets. These refinement of metric provides connection between results from project measurement and team problem solving process [4].

Corral [3] created his own system, which solves implicit obtaining of information, too. This system logs information from environment around developers. The system monitors distance of collaborating developers in team projects through wireless technology *Bluetooth*. In this way it is also recognized that the developed source code occurs alone in collaborating with a number of

* Master degree study programme in field: Information Systems
Supervisor: Karol Rástočný, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

developers. His aim was to obtain these data and their visualization by using jQuery and Google Visualization API[1] for the purpose of the following analyses. It provides means for division of developers into teams based on their frequent collaboration.

2 Our method

Current approaches to the source code review require manual inspect of all code changes by code reviewers. In this approach there are often overloaded reviewers which directly affect the degradation of the quality of source code. The main objective of this work is to facilitate the time-consuming work of assessors of the source code and thereby contribute to the quality of the developed software. We analyze developers' activity during the development of software product by using the process mining technique, watching developing processes which are divided by commits. We collect annotations for individual processes and according to these annotations we determine character of the process. We build process models and compare annotated process models with new process models that have not been already annotated. We consider the processes with high similarity as the processes with the same character and in this way we are able to find out processes without any annotation. If new process has error character, we create recommendation for code review for the code which was written in this process.

In this paper we focus on processes and the technique which is called process mining [5]. The process is simple structured quantifiable set of ordered activities which execute transformation from inputs to required outputs. Process is an ordered set of steps which lead to the desired object. In our case, the steps are activities obtained from system for monitoring of a user's activity. Every process can be represented by a model – oriented graph [2]. One of the possible approaches how to analyze processes is to create models of processes. These models are analyzed using a technique called process mining. Process mining technique deals with methods of creating process models. Many processes can be generalized, eventually inherited similar properties (steps). Similarity of processes allows us to their comparison and searching for their mutual properties. According to their mutual properties we can predict their following behavior, optionally create recommendations.

Process mining is relatively young research discipline. Its main idea is to discover, monitor and improve real processes (not assumed processes) by extracting knowledge from event logs readily available in today's information system. Process mining involves process discovery (building process models from event logs), conformance (monitoring deviations between model and event log), extension of model, model correction, prediction cases and creating recommendation on historical base [5]. The starting point for the process of discovery is a record of events. All techniques of process mining assume that individual records events relate to activity linked to a particular case. Event logs may include various additional information which the techniques use for the detailed description of the process. Most solutions are focused on the mining processes to analyze data in offline environment as well as in our case are applied to the complete cases of event logs.

Our method extracts useful information from event logs of developers' activities, creates processes from these activities, obtains annotations and thus determines the character of processes. From these processes we build process models that we compare with newly formed process models (with no annotations). This way enables us to know a character of new process and thus in case of wrong character of process create recommendation.

We have implemented a prototype application which uses our method. The prototype consists of activity analysing, process creating and evaluation processes based on the obtained annotations. The application is called ProcessMiner.

[1] https://developers.google.com/chart/interactive/docs/reference

2.1 Architecture

ProcessMiner has a database which contains processed data in process form. ProcessMiner loads data from four systems:

- UACA[2] – The User Activity is client-server application designed to acquisition of user's work-related activities. Client application UACA (User Activity Central Application) for collection and management of user activities is running in developers computers. Activities we work with are web browser and IDE activities. User activity is monitored by web browser plug-ins and IDE plug-ins. UACA creates events from monitored activities. Several events are created for a short time interval.
- CORD – The CORD is service that allows access to GIT repositories, history of project and commits from API.
- CodeReview[3] – The CodeReview is a system which provides an access to the user data and allows insertion and management information tags in the source code. Users are identified in UACA, CORD and ITM different, but CodeReview contains user aliases from these systems.
- ITM – The ITM is a database of information tag management system. From this database we get annotations for processes in the form of information tags.

2.2 Taxonomy

We have created our own event taxonomy that defines the type of events (see Figure 2). This taxonomy is used to model processes and according to this type we are able to filter and analyze the influence of granularity of events to the resulting process. When we model processes we examine the impact on abstraction of events to the similarity of processes.

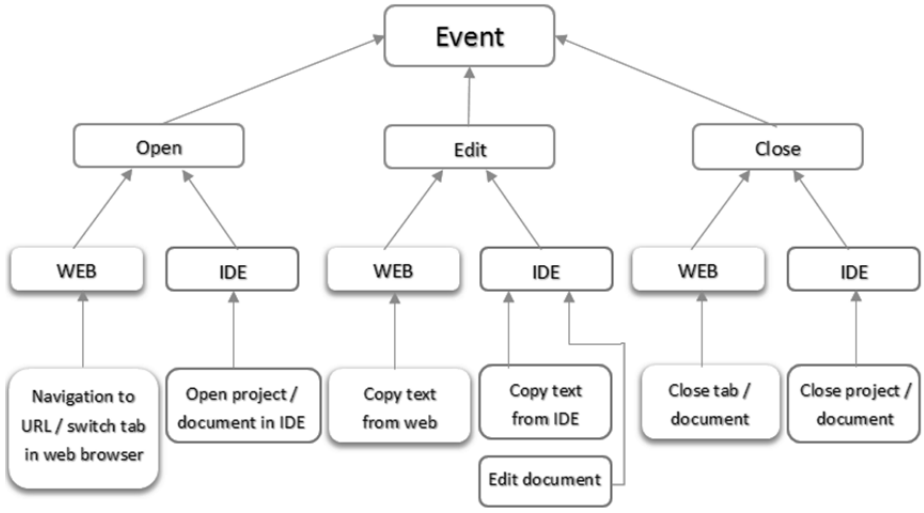


Figure 1. The event type taxonomy.

Inheritance ensures that the type which is lower in hierarchy is type of higher defined type. All of the types inherit from the type *Event* which is on the top of hierarchy. The first division of event is split into three types – actions (each action has a defined environment where the event comes from). The lowest layer defines an action processed in this environment.

[2] http://perconik.fiit.stuba.sk/UserActivity
[3] http://perconik.fiit.stuba.sk/CodeReview

2.3 Identification of process interval

The data we work with are only sequence of activities sorted by timestamp of creation. Whereas the process is defined as ordered set of steps which lead to the desired objective. During the development of software, programmers solve tasks which can be decomposed to the sequence of events. The objective that we monitor is to accomplish the task which ends by creating a new version into some Revision Control System. Events recorded during the development of one task we consider in our solution as a process. Event of creation of a new version is used as a delimiter of this data sequence as shown in the Figure 3.

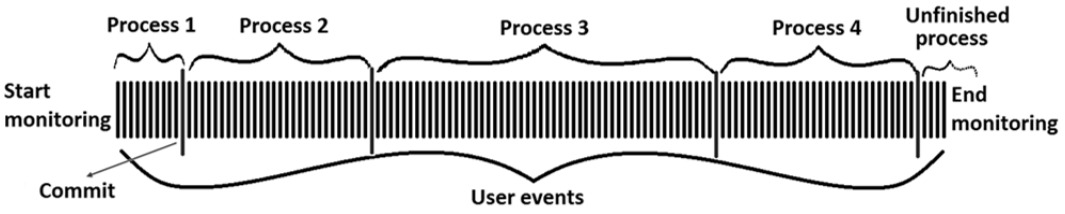


Figure 2. Splitting of the sequence of activities into processes.

The events are loaded in regular intervals. As we can see in the Figure 3, processes do not have the same number of events. This is due to different difficulty of tasks or data noise [6]. Last process in the Figure 3 is not completed because it is not finished by commit. This process will be completed in the following interval.

2.4 Annotations of processes

To be able to define character of processes we have to classify these processes in some way. For classification of processes we use annotations which are extracted from information tags and commit messages.

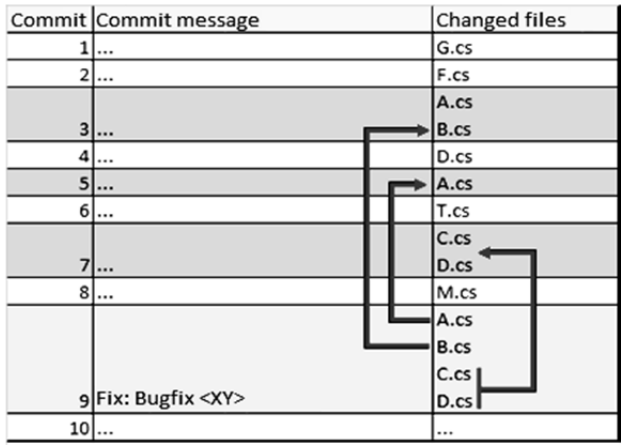


Figure 3. Reverse marking of erroneous commits according to changed files.

Information tags are obtained from the ITM database. These tags are created by reviewers for a specific part of a code. One file in one version can contain a number of information tags. Information tags contain information about tagged part of a code (*bug*, *smell*, etc.). If a file contains some information tag which type is *bug* – we can say that the process by which these files are changed was erroneous.

Commit messages can contain additional information about new committed versions. These comments shortly describe changes that were made in new versions. Good practice when writing comments is using keywords (*fix*, *add*, *delete*, etc.) that describe character of new versions. If commit message contains keyword *Fix*, we assume that previous change of these files from this commit was erroneous. Problem of this approach is that we do not have access to files' contents (we don't know which part of a file was changed) therefore we mark several the previous commits that changed (or added) these files (see Figure 4).

3 Evaluation

Data acquisition is one of the most difficult but also very important important part of process mining methods. ProcessMiner works with data obtained from the four systems accessible via REST or SOAP services provided by systems developed within the research project PerConIK [1] (*Personalized Conveying of Information and Knowledge*). Data acquisition architecture is shown in Figure 1.

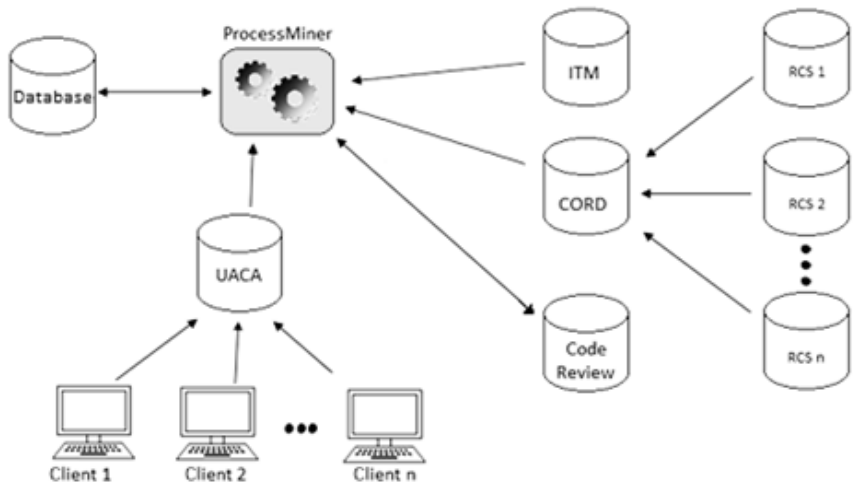


Figure 4. Architecture of data acquisition for ProcessMiner.

3.1 Experiments

In experiments we observed working procedures between academic and corporate environments. Academic environment is represented by data from FIIT STU in Bratislava, where the *UserActivity* is used in the team projects. Corporate environment is represented by data from company Gratex International[4] which collaborates on development of the project PerConIK and actively uses PerConIK in their work. We analyse data obtained from these environments for 36 days from Gratex and 84 days from FIIT (see the results in Table 1).

We have annotated created processes in order to determine the character of processes. So that we can examine the character of processes and find out what proportion of the process must be revived. In the first experiments we classified processes into three categories: *Error, Repairing and Neutral* (see Table 2). In the following experiments we will examine and validate processes in more details.

[4] http://www.gratex.com

Table 1. Measured values from obtained data.

	Gratex International	FIIT STU
Number of active UserActivity users	14	7
Average number of processes per user per day	1,38	0.47
Average number of events per user per day	642	312
Average number of events per process	465	710

Table 2. Process classification.

	Gratex International	FIIT STU
Number of all processes	696	277
Number of error processes	68 (10 %)	44 (16 %)
Number of repairing processes	196 (28 %)	11 (4 %)
Number of neutral processes	432 (62 %)	222 (80 %)

4 Conclusions and feature work

In this paper we introduced the method for classification of software developing processes. We implemented a prototype which obtained data from external services and from these data it built processes. We also introduced the method for annotation of processes from information tags and extract annotations from commit messages. In the experiment we compared data between corporate and academic environment. From the results we observed differences between these environments, especially in number of created processes and the number of events per process.

Our following step is building of process models. We plan do model processes in the program RapidMiner using ProM Framework extension which provides tools for process mining. Built process models will be compared and the character of processes will be detected.

We plan to perform additional experiments in larger projects from company Gratex. We also plan to create more detailed classes for classifying processes.

Acknowledgement: This contribution is the partial result of the Research & Development Operational Programme for the project Research of methods for acquisition, analysis and personalized conveying of information and knowledge, ITMS 26240220039, co-funded by the ERDF.

References

[1] Bieliková, M. et al.: Platform Independent Software Development Monitoring: Design of an Architecture. In*: SOFSEM 2014*, LNCS 8327, Springer, 2014, pp. 126–137.

[2] Cook, J.E., and A.Wolf, Discovering models of software processes from event-based data, *ACM Trans. Softw. Eng. Methodol.* 7, 3 (July 1998), pp. 215–249

[3] Corral, L. et al.: DroidSense: A Mobile Tool to Analyze Software Development Processes by Measuring Team Proximity. In C. A. Furia & S. Nanz (Eds.), *Objects, Models, Components, Patterns* (Vol. 7304, pp. 17–33). Springer Berlin Heidelberg, 2012.

[4] Meneely, A., Corcoran, M., & Williams, L. Improving developer activity metrics with issue tracking annotations. In *WETSoM 2010*, pp. 75–80. New York, USA: ACM Press. 2010.

[5] Van der Aalst, W. M. P. 2012. IEEE Task Force on Process Mining Manifesto. In*: Business Process Management Workshops LNBIP*, Springer, 2012; 99, pp. 169–194.

[6] Thomson C., Gheorghe M.: Using Process Mining Metrics to Measure Noisy Process Fidelity. In *EASE'09 Proceedings of the 13th international conference on Evaluation and Assessment in Software Engineering*, 2009, pp. 132–135.

Collaboration in Development of Active Conceptual Model and its Management

Ondrej JURČÁK*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
jurcakondrej@gmail.com

Abstract. The active conceptual model is a new approach to traditional conceptual modelling. It enriches the traditional conceptual model with new time dimension, which represents the world more dynamically. The dynamic aspects and evolution of a domain scheme force the evolution of the conceptual model itself. Collaboration support and proper model management during the evolution changes allow creating complex problem solutions and track changes in the past. We propose a system which allows several developers to collaborate during evolution of the active conceptual model. The proposed system stores versions of the model to improve its management during the development and maintenance.

1 Introduction

Conceptual modelling captures the real world into abstract patterns. Traditional conceptual modelling is focused on a static view of the world. It gives us a better understanding of the real world and makes it easier to communicate between domain specialists.

Fast changes in the real world are forcing us to change the process of conceptualization, which brings us to a new view of the reality and exploitation of our knowledge about the real world. Changes in the real world need to be captured and we have to understand these changes for future predictions. Active conceptual modelling is a new approach to conceptual modelling. It is a multi-perspective, abstract model of the reality which integrates temporal and spatial entities, temporal relations and events [3]. It describes all aspects of a domain and activities in this domain from different perspectives.

To describe a learning content of some profession, we need to model reactions to events, which can occur. For example, if in a library someone wants loan a book, a librarian has to find this book and register the loan into the system. The loan of the book is an event, which the librarian has to handle, if he wants do his job. We can describe required reactions to events using scenarios. In software design we use the UML Use Case diagrams to model the desired software reactions to events. But instead of software we can also model professions. Modelling with the Use cases adds only a dynamic dimension to the model. To catch the static structure we use Class

* Master degree study programme in field: Software Engineering
Supervisor: Dr. Ján Lang, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

diagram elements. Domain model is representation of the real world, where relation between concepts and concepts itself may be changed. Domain model evolves during the time, as changes in world occur and new requirements come from model usage. We need to manage these evolution changes. They allow us to analyze decisions made in past and find changes and responsibilities for changes, which led to errors in the model.

Each human being has some limits, which make impossible to resolve some issues during the work. Collaboration in teams removes these limits and allows developers to find more efficient solutions. Group of specialists cover bigger space of knowledge and each individual sees domain from different perspective. When developers work on high level of abstraction, form requirements, design software, write source code or create a test, they are slow and they make mistakes. So realization of big projects require specialists to collaborate [6]. Collaboration support in teams is not a trivial task and in software engineering, version systems are used to manage work of multiple developers on one product.

2 Related works

Conceptual model could be represented in many ways. Simple way to represent domain are hierarchical models called “Mind maps”. Main principle of this model is that concepts have better semantic value together with their perceptual representation (image, sound, color, etc.) and human mind can better associate concepts with the real world [1].

The Unified Modelling Language (UML) is based on main principles of E-R model. UML is set of graphical languages and diagrams used to describe a software product. Most used UML diagram for describing static view of software is Class diagram [3]. Active conceptual model can be represented by a model based on two UML diagrams. First Class diagram, which holds static view and Use Case diagram which represents dynamic dimension [5]. Such active conceptual model can holds both static structure of domain and dynamic domain reaction on events.

Effective collaboration support in software engineering is one of the critical task during development process of large projects. Version management system supports collaboration and also provides management tool for software product during its evolution. Modern and the most used version system, Git or older SVN, are designed for source code versioning. We can use collaboration principles used in file based versioning system and use it in model versioning system [4]. UML is commonly used to describe a software design. Oliviera et al. proposed versioning system Odyssey-VCS to support management of UML models exported in XMI format [4]. From one perspective, we may see source code as model, too. There are relations between files and they are structured hierarchically. Principles of such version system may be used in version management system for an active conceptual model.

3 Collaboration and management support during evolution of active conceptual model

We design versioning system for an active conceptual model to support collaboration and management. Active conceptual model notation is based on combination of UML Class diagram and Use Case diagram. We analyzed existing versioning system, as Git or SVN, and chose main features and solutions which support collaboration. We researched history of main version systems in software engineering to better understand evolution of collaboration features in such a system in past. During the time, designers of such a system faced numbers of problems. Main problem, however, is how to support several people’s collaboration on the same project artefacts.

When the same project artefacts are modified in the same time, some changes may be override by another. For example, if one developer adds a method to class in a file, but another in the same time deletes the property, we don’t want to lose any of these changes. In our system, we used an optimistic strategy to resolve this problem. Main principle is that each developer have

local version of a model on the computer and makes his own changes with model artefacts. These changes are persisted locally and to publish these changes to remote repository, developer has to merge the local version with the remote one. When developer pushes their changes to remote, two main scenario may occur:

1. No changes had been pushed to repository by any other user – in this case, all changes can be stored in the remote repository immediately.
2. New changes had been pushed by second developer – in this scenario, user has to merge his changes with version from remote repository and resolve detected conflicts.

On diagram below (see Figure 1) we can see the second case. The second developer pushes changes to the remote repository after the first developer. Client 2 detects that it's outdated and they fetch an actual version from the server. The second developer merges the remote version with his own version and pushes changes to the server.

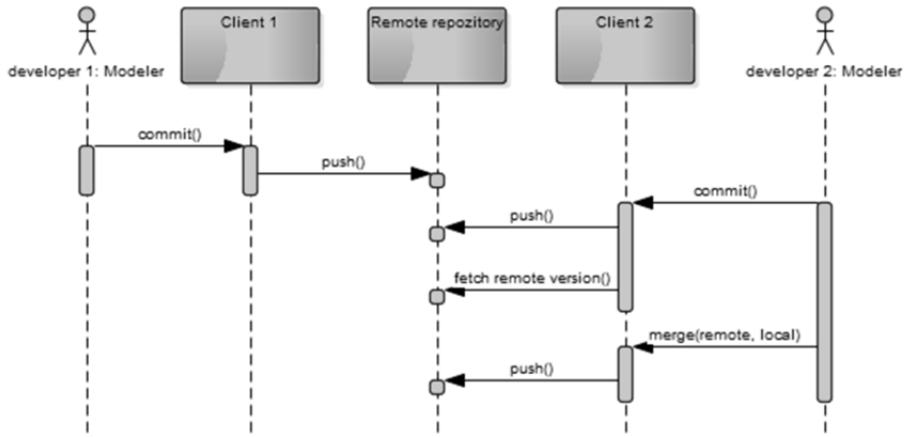


Figure 1. Sequence diagram of an optimistic collaboration between two developers with the proposed system.

Main difference from a pessimistic collaboration used in an older system is that it allows more than one developer to work on the same artefacts in the same time. One of the main lack of the pessimistic collaboration strategy was that, if developer wants to store some changes, he needs connection to the server. Local version of the model allows developers to make changes on the model without the internet connection. All changes are stored locally on developer's machine. Connection is required only if developer wants push changes on the remote repository. Main requirements to make the optimistic strategy to work in practice, is implementation of branching, solid version merging mechanism, conflict detection and resolution algorithm.

During model evolution, a new requirement may call for a new variation of model. When we add new features, we make changes which are not desired to be in our stable version before finishing the work. We use branching to resolve this problem. Branches are known from a versioning systems for source code. Our system allows to manage different versions of a model in separate branches. Developer can create their own branch from a selected version and there he can store his own changes. Changes in one branch don't affects another branch. In other hand, splitting versions into branches also requires its merging.

When we merge two different branches, we need to deal with conflicts. They occurs when the same model elements were changed in both branches. Any changes loss is undesirable. For this purpose, we design our system to detect conflicts and to resolve them. System conflict detection and conflict resolution algorithm are based on the element of comparison (proposed by Oliviera, Murta and Werner [4]). In our system, elements of comparison are EObjects. They are composed from attributes and references to another EObjects. Such an object can be stored in a used

repository. If any of object's attributes or references have been changed in both branches, conflict is detected. System using two types of conflict resolution:

1. Automatic – different properties or relations of an object had been changed.
2. Manual – the same properties or relation types had been changed. For example, developer changes the object's attribute name to "Bob" in one branch and another developer changes name to "Ann" in second branch.

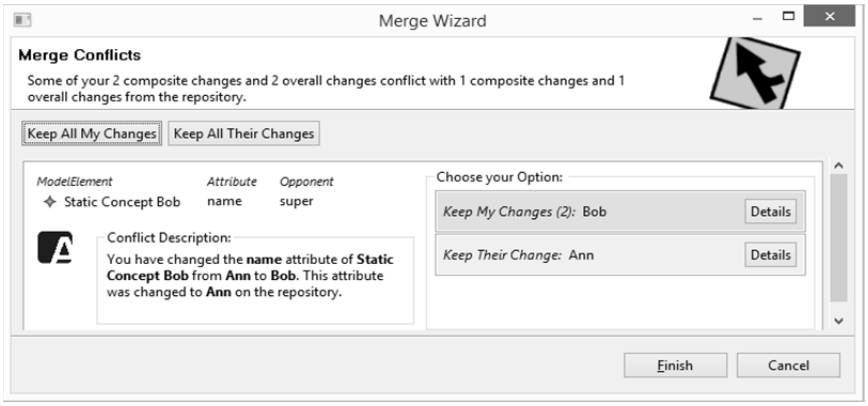


Figure 2. Manual conflict resolution. System list changes and user has to choose which to apply.

In the first case, we can automatically resolve conflicts without losing some operations. But in the second, we don't know which value of attribute name to choose from. In this cases we use developer's assistance to resolve the conflict. On the image above, (see Figure 2) we see a wizard for user to decide which changes we should apply to model element. All changes in the versioning system are represented by the atomic operation like deleting, renaming, adding, etc.

Some changes require two or more atomic operations to perform. For example, moving the attribute from one concept to another requires two operations. The first is deleting an attribute from one concept and the second is adding the attribute to another concept. When we track changes in the past and we see this operation alone, we may lost the original purpose of these changes. The solution is to composite operations. We can put the atomic operation together and make the composite operation (for example: "move property"). This composite operation has better semantic value for user and allows user to better understand changes of the model.

In the section below, we describe the system for versioning the active conceptual model of learning. To evaluate collaboration features of such system, we developed the prototype of the versioning system.

4 Evaluation

We implemented the solution described below in the prototype of the versioning system for the active conceptual model. We used Eclipse Modelling Framework to create a model and we stored such model using EMF Store. To simulate collaboration and model management we created Library domain that was modelled using the active conceptual model of learning.

On the image below (see Figure 3), we can see a simple example of the active conceptual model of learning. On the left, we see the initial version of the librarian model. To demonstrate optimistic collaboration, the second developer created a new branch "develop". In this branch, the second developer added a new concept "Section" and he created associations. After finishing work, developer merged "develop" branch into "trunk". During merging, process prototype detected conflicts and the developer had to manually resolve conflicts.

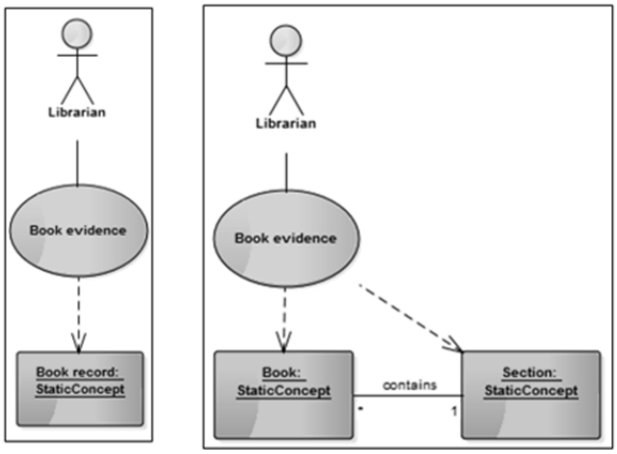


Figure 3. Example of the active conceptual model. On left we see the initial version of the diagram representation, on the right we see the final version of the diagram representation after merging.

To simulate conflicts, we used Book concept (see Figure 3). On develop branch "Version 4" (see Figure 4) contains of "Book record" concept to "Publication" and on branch trunk in "Version 5" name was changed to "Book". When develop branch is merged to trunk, system detects conflict on "Book record" elements. System needs user assistance to resolve this conflict, so it proposes two options to user. The first is to change the name to "Book" or the second is to "Publication". Developer then selects the desired option and in merged version ("Version 6") new name of the concept is set properly.

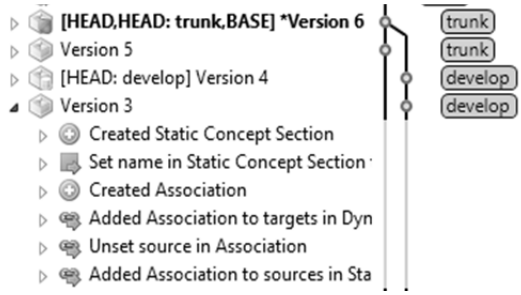


Figure 4. History view of library repository.

There are two branches, trunk and develop, in the history view above (see Figure 4). Each version contains bunch of operations made with the model. In the history view, we can also see version's date and developer's name. This information leads to a better model management. We can track changes in the past and find responsibilities of these changes.

In the history view, we can see branching. Developers can create branches explicitly from selected version, but to support the optimistic strategy, the system automatically creates a local branch for each developer. He saves their changes into this branch and this branch can be merged with a remote branch.

In our work we used principles of systems used in software engineering to support collaboration and to help managing model in content engineering. Our active conceptual model is based on main principles of UML Class and Use Case diagram. Existing versioning system for model, Odyssey-VCS, are used for managing UML models for software design. Main difference from our system is that we manage active conceptual model. We use this model to describe

learning content. In our work we adapt main principles of existing versioning tools (Odyssey VCS for model and Git or SVN for source codes) to allow many developers to collaborate during development of learning content. This existing versioning system are being used on real software projects by many developers, so principles used in such tools are based on real requirements of software developers.

5 Conclusion

We used active conceptual model of learning based on UML Class and Use case diagram to model Library domain. We modelled librarian profession using active conceptual model. Domain is space of concepts, roles and roles reaction on domain events. We used Class and Use case diagram elements to model domain. Static structure is modelled using Class diagram and we adapt Use cases to model behaviour in model, which are reaction on events.

Main focus of our work is collaboration on active conceptual model. Software projects use version management systems to manage work of multiple developers. Version systems for source code files are widely used on many projects. Such system helps developers collaborate and maintain different versions. Changes can be track down in past and determine developers, whose made such changes. We design version system for active conceptual model of learning, which adapt some main features of collaboration tools used in software projects.

Proposed system use optimistic strategy to support collaboration during model evolution. Branching, conflict detection and conflict resolution during merging process are requirement for this cooperation strategy to work. We used two types of branching. Implicit branching for local version of model to allow developer make his own changes and work without connection to server. Second explicit branching for management of multiple model variation. Side effect of branches merge is conflicts when different changes are made on same model artefacts. We focused on detection and resolution of this conflicts to avoid loses of information during merge.

Acknowledgement: This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG1/1221/12.

References

[1] Buzan, T.: Mind Mapping. *The Mind Map Book: How to Use Radiant Thinking to Maximize Your Brain's Untapped Potential*, BBC Books, London (1993), 83–91.

[2] Devedzić, V.: Understanding ontological engineering. *Communications of the ACM*, (2002), vol.45 n.4, pp. 136–144.

[3] Gogolla, M.: UML and OCL in Conceptual Modeling. In: *Handbook of Conceptual Modeling*, (2011), pp. 85–122.

[4] Oliveira, H., Murta, L., Werner, C., Odyssey-VCS: a flexible version control system for UML model elements. In: *Proc. of the 12th international workshop on Software configuration management (SCM '05). ACM*, New York, USA, (2005), pp. 1–16.

[5] Urban, O., Lang, J.: A Learning Content Active Conceptual Modeling. In: *Proc. of 2012 15th International Conference on Interactive Collaborative Learning (ICL)*, Villach, (2012), pp. 26–28.

[6] Whitehead, J.: Collaboration in Software Engineering: A Roadmap. In: *Proc. International Conference on Software Engineering (FOSE '07)*, pp. 214–225.

Irrlicht Based Three-Dimensional Complex Conceptual Models Visualization

Tomáš KEPIČ*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
tomas.kepic@hotmail.com

Abstract. Conceptual models represent some realities or ideas. Although models are mostly simple, sometimes they can be very complex. Increasing complexity is caused by large number of elements – nodes, relationships. Using three-dimensional space can be considered as one of the methods given for their simplification. There are many related works and realizations of this idea. In this paper, we visualize conceptual models in modelling language UML using graphical engine Irrlicht. Processing XMI file exported from Enterprise Architect and division the complex model to layers by different types of elements appears to be promising approach in creating three-dimensional models.

1 Introduction

When we try to express some realities or ideas, we create models, which represent them. When the models are created from some concepts, we talk about conceptual models. Sometimes there is also activity, which we want express in model. Models, which include activity are called active conceptual models. Introduction of activity bring some more complexity to simple models. On other hand, they also can include large number of elements. The models become more complex.

In our work, we are finding method how to simplify complex conceptual models. One of the examined method is using of three-dimensional space. There exist some works which examined three-dimensional models. In our paper, we describe creating three-dimensional models with graphical engine Irrlicht[1]. We choose UML[2] as language for modelling, because it is the most used modelling language in software engineering. It is also possible to use it in other domains. Our three-dimensional UML models are based on two-dimensional models. These models we create in CASE tool Enterprise Architect[3]. We use XMI format for saving information about models. We expect that this approach bring simplification to the modelling of large complex models.

---

* Master degree study programme in field: Software Engineering
Supervisor: Dr. Ján Lang, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

1 http://www.irrlicht.sourceforge.net/

2 http://www.uml.org

3 http://www.sparxsystems.com/products/ea/index.html

2 Related works

In the introduction we describe how complexity in conceptual models is increasing. For simplification we propose to use three-dimensional space. First, we examined existing solution which introduce third dimension to UML modelling. There exist many other works. One of the first work is called On better understanding UML diagrams through interactive three-dimensional visualization and animation [4]. Authors describe possibilities how to use third dimension in UML modelling. One of the most developed work is related to three-dimensional state diagram [3]. Authors divided state diagram to layers, which express nesting in some state. Solution was developed and tested. Professionals said that it is good solution and they do not need to switch screen for gathering information from substate. On the other hand, three-dimensional orientation bring new view to model and they have to accustom for using this orientation.

In our faculty, there are also some works which deal with three-dimensional models [6]. One author created a three-dimensional class and activity diagram. It is also possible to extend this solution to other diagrams, which are described. For example other group of students created three-dimensional sequence diagram [1]. Their software prototype allow them to add and remove elements in three-dimensional space.

3 Creating and reading model in XMI format

In our work, we do not create CASE tool for adding elements in three-dimensional space. We created tool for visualize two-dimensional models in three-dimensional space. So we use other existing CASE tool for creating models and we convert them.

For saving UML models we choose XMI format exported from Enterprise Architect. This decision was made because Enterprise Architect save lots of adding information about created model. With them it is easier to convert two-dimensional model to three-dimensional. There are also existing some student research bachelor thesis about examining XMI format and third dimension [2, 5]. As we mentioned in introduction, we use UML language for modelling. It is the most used language in software engineering, but it is also possible to use it in content engineering. Because we try to catch both modelling domains, we create tool which visualise active conceptual model. One of these models was described in diploma thesis Urban, O.: Využitie aktívneho konceptuálneho modelovania na tvorbu obsahu [7]. The model use static and dynamic elements of language UML. It includes elements from class diagram – classes and corresponding relationships for modelling static parts and use case diagram with actors and corresponding relationships for modelling activity. These two types of diagrams are joined. We also want to join these two types of diagrams, but we introduce collaboration to connect these diagrams.

To summarize, our conceptual model involves four types of elements (nodes) and many types of relationships between them. For modelling this model we use collaboration model from Enterprise Architect with some additional elements from class and use case diagram. This model we save in format XMI, which is based on XML format and it is simple to create application to read it.

The XMI file from Enterprise Architect consists of three parts:

- Header
- Basic information about elements and relationships
- Further information
  - Information about elements
  - Information about relationships
  - Information about diagram

We are interested in second and third part. We do not need information from header. Further information are information stored by Enterprise Architect, which describe basic information for better processing of elements and relationships. Very important for our work are information about diagram. There are stored coordinates where are elements placed in the model. Relationships do not store their coordinates, they are calculated.

For compare, CASE tool IBM RSA[4] includes only basic information about elements and relationships. We can read them, but we do not know where the elements are placed. It is not possible to create same visual model by importing XMI file from IBM RSA. That is reason why we chose Enterprise Architect XMI file and we get information from it.

The following fragment of the XMI file represents the element type of Actor:

```
<packagedElement xmi:type="uml:Actor" xmi:id="EAID_91E9AC8E"
    name="Citatel" visibility="public"/>
...
<Association xmi:id="EAID_3E18C4F0" start="EAID_91E9AC8E"
    end="EAID_4203A906"/>
<Association xmi:id="EAID_9F642B37" start="EAID_91E9AC8E"
    end="EAID_CBF04F1F"/>
<Association xmi:id="EAID_D540C8CA" start="EAID_91E9AC8E"
    end="EAID_5AA33E96"/>
...
<element geometry="Left=1482;Top=509;Right=1527;Bottom=599;"
    subject="EAID_91E9AC8E" seqno="3" style="DUID=E8CCECA9;"/>
```

We read this XMI in the following sequence:

1. From basic information we get ID, type and name of item
2. We divided items on elements and relationships by their type
3. From further information we get all information about relationships. They consist of type as name and they have attributes, which are ID, starting and ending element
4. We mapped information from third point to information from first point by relationship's ID
5. From further information about diagram we get information about location of elements

After getting information from XMI file, we create graph from elements and relationships. For every element we add list of his neighbours. Using this algorithm we can read information about model which we visualize.

4 Three-dimensional model visualization

For creating three-dimensional space, we add third coordinate to element's coordinate in model which are captured in third part of XMI file. In actual solution, we add value of this coordinate by element type. In the result, it means that each type of element is located on other layer. We supposed that these layers provide better orientation in model. Recognition of elements will be also simple for users, because it will not depend only on element shape, but it also depends on layer in space where it will be placed. Our layers consist from these elements in this order:

- Use case actors
- Use cases
- Collaborations
- Classes

[4] http://www.ibm.com/developerworks/downloads/r/architect/

These four layers create complex conceptual model. Between these elements are also relationships, which connect them between layers or on the same layer. Used elements and relationships are also formally described in UML metamodel[5]. Due to third dimension, we need to add space and layers to metamodel. For example, in the work ŠKODA, M.: Trojdimenzionálne zobrazenie UML diagramov [6], an author describe adding space and layers into UML metamodel.

For visualization of these models in three-dimensional space we create software prototype. It consists from importing XMI file and visualization of elements and corresponding relationships in three-dimensional space.

We choose external software to create XMI file with UML model information, but we need software to visualize model in three-dimensional space. Although there exist lots of other solutions, we decided to create brand new visualization solution. After implementation we can compare our chosen technology with others. In our work we use graphical engine Irrlicht. It is three-dimensional graphical engine. Primary programming language is C++. There also exist some wrappers for other languages. We decided to create our solution in programming language C#. Because of this decision, we use graphical library Irrlicht Lime[6]. It is C# wrapper for classical C++ Irrlicht library. Besides standard programming, we can use some other C# and .NET technologies. For example we can create application with standard graphical GUI based on Windows Forms and we can insert Irrlicht screen into this window GUI or call separate window from GUI. We do not need to use other library or framework for standard Windows GUI, where we can make some changes to the read XMI model. This functionality can be very useful in further work.

Other reason for choosing Irrlicht is community of users. On the website of Irrlicht, there exists lots of tutorials, which can be used and they are very helpful during the work. Other support with programming is provided by forum, where we can find answers for many questions and problems. Thanks to this community, work with graphical engine Irrlicht is interesting and it saves lots of time.

5 Implementation in graphical engine Irrlicht

In graphical engine Irrlicht are three-dimensional object defined as scene nodes. They have certain shape and they are placed somewhere in the space. They have also other features as possibility to assign material, application of texture on surface and so on. We mentioned community which created lots of tutorials. One of them describes creating our own scene node. This node is created by inheritance and adjusting basic properties from base library node. This is very important feature which we use in our implementation for creating model's nodes. We created several nodes which represents items with same properties as shape or texture.

First of our nodes is representing elements. Element can be class, use case, actor or collaboration. These elements are placed on same layer by type. All of them can be created as two-dimensional squares with different texture on surface, which give them specific visible shape.

Second type of node is relationship. It is not same node as element, because relationship can be placed between different layers. This node need more coordinates, which specific starting and ending layer. As element, each type of relationship is differed by its texture.

These graphical scene nodes are specific for each type. But we need one more node for text on elements. It is different for each element. We can not use same approach as was used for elements or relationships, which are using two types of general nodes and certain textures. We decided to show text next to element node. We examined possible visualization of text in graphical engine Irrlicht. There are some possibilities to visualize custom text, but they have one problem. We can define coordinates where the text is placed, but we cannot define direction. It means that text is always readable from camera and after rotate it looks same. One option can change scale of

[5] UML Superstructure Specification, `http://www.omg.org/spec/UML/2.4.1/`

[6] `http://www.irrlichtlime.sourceforge.net`

text with increasing camera distance, but direction was always same and we cannot see text from different angles. We were looking for solution, where the text should be placed in space and turned always in same direction. After many attempts to implement we found solution on forum. One user created his own node which consisted from square and texture with custom text. It was possible to change text and square automatically adapts to text size. After testing, it looks like ideally solution for our problem. It was created in language C++, which is primary programming language for Irrlicht, so we convert it to C# language and we use this solution.

To sum up, in our visualization application we designed three types of custom scene nodes – node for element, relationship and text. We visualize complex conceptual model using our nodes.

As we mentioned, we used textures to differ elements and relationships in our solution. It was important to create proper texture mapping to surface. After many tests and settings we get appropriate solution. We also had to familiarize with Irrlicht coordinates system due to proper placement of nodes. Overview of coordinates in space (x, y, z) and coordinates for mapping (u, v) can be seen in Figure 1. In model visualization, textual node need only basic x, y, z coordinates so we do not show him in Figure 1.

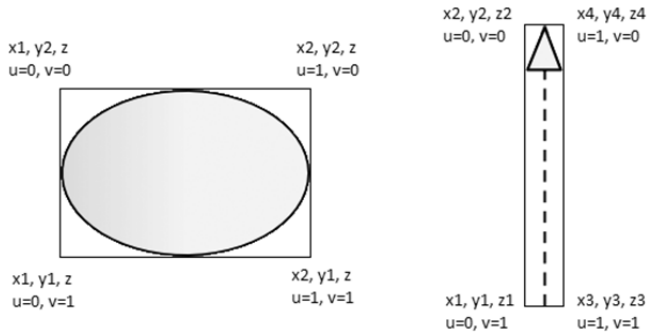


Figure 1. Coordinates for location in space and texture mapping of elements and relationships.

Another important task was setting transparency of texture on node's surface, because elements are not only squares. For example use case element is a circle and corners must be completely transparent. After many tests and experiments of graphical engine Irrlicht we chose option, where transparent part of texture are removed and colour parts are preserved without any transparency. We tested also partially transparent colour parts, but when we rotate camera or move in space, there reveal problem as missing texture, so we decided to do not use this method of transparency.

After solving all problems we have discovered during implementation, we created software prototype which can visualize model from imported XMI file.

6 Conclusions and further work

As an example of three-dimensional visualization we use active conceptual model of the diploma thesis Urban, O.: Využitie aktívneho konceptuálneho modelovania na tvorbu obsahu [7]. We created this model in Enterprise Architect. We added collaborations into model between use cases and classes. Then we saved created model into XMI file, which we can visualize in three-dimensional space using our software prototype. For visualization we use graphical engine Irrlicht and library Irrlicht Lime for programming language C#. The result of visualization can be seen in Figure 2. We can see four types of elements which are divided into layers.

Graphical library Irrlicht Lime can be also used in Windows Forms application. This is one of the ways of further work. We can implement some functionality to modify three-dimensional models and immediately apply changes. For example, it can be used for filtering complex models.

We suppose that three-dimensional space can be helpful for recognition of elements in large and complex models. Recognition is mainly supported by layers, which are created by same type of elements. As verification of this assumption we can create questionnaire or interview with different groups of people, such as students or professionals. We suppose, that creating three-dimensional models from Enterprise Architect XMI file can be interesting for professionals who everyday create large models in this CASE tool and it brings new possibilities to their work.

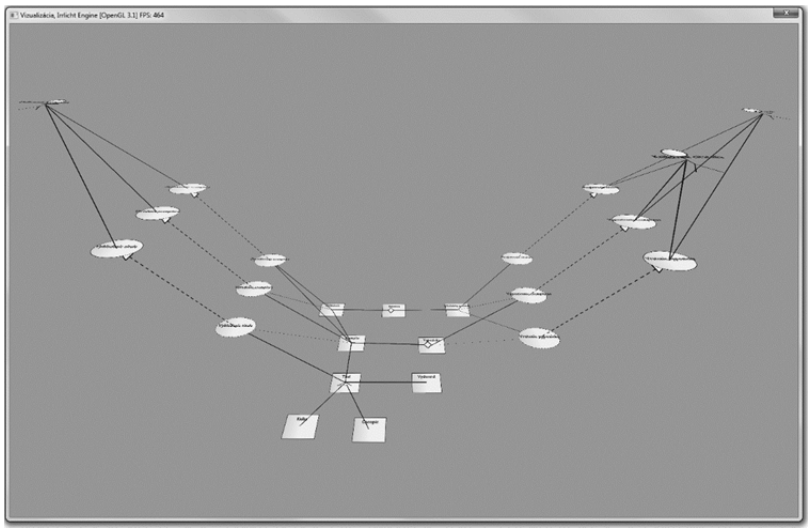


Figure 2. Visualization of complex conceptual model in software prototype.

Acknowledgement: This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG1/1221/12.

References

[1] Brndiarová, G., Martoš, I., Štajer, A., Štetiar, M., Šuta, E., Valko, A.: Three-dimensional Visualization of UML Diagrams. In: *Student Research Conference 2014, Vol. 2: 10th Student Research Conference in Informatics and Information Technologies Bratislava, April 29, 2014 Proceedings.* pp. 607–608. ISBN 978-80-227-4152-1.

[2] Kostoláni, M.: *Viacrozmerné zobrazenie modelu v jazyku UML*. Bachelor thesis. FIIT STU, 2014. (in Slovak)

[3] Mcintosh, P., Hamilton, M., Van Schyndel, R.: X3D-UML: 3D UML State Machine Diagrams. In: *11th international conference on Model Driven Engineering Languages and Systems* (MoDELS '08). Springer-Verlag Berlin, Heidelberg, 2008, pp. 264–279.

[4] Radfelder, O., Gogolla, M.: On better understanding UML diagrams through interactive three-dimensional visualization and animation. In: *Proceedings of the working conference on Advanced visual interfaces*, Palermo, Italy. ACM, pp. 292–295.

[5] Sýkora, T.: *Viacrozmerné zobrazenie modelu v jazyku UML*. Bachelor thesis. FIIT STU, 2014. (in Slovak)

[6] Škoda, M.: *Trojdimenzionálne zobrazenie UML diagramov*. Diploma thesis. Bratislava: FIIT STU, 2014. (in Slovak)

[7] Urban, O.: *Využitie aktívneho konceptuálneho modelovania na tvorbu obsahu*. Diploma thesis. FIIT STU, 2012. 83 p. (in Slovak)

Use Case Based Approach in Conceptual Modeling

Peter KYSEĽ*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
xkyselp@fiit.stuba.sk

Abstract. Common problem in Conceptual modeling is in accuracy of names and structure in Class Diagrams. Various approaches offer different degrees of reality approximation. This is also related to the quality of the result itself. In this paper we present our approach in conceptual modeling based on use cases treatment. We have performed finding candidates for classes even their attributes automatically by use cases analysis. We focus on extraction nouns from detailed description of selected use cases. Finally we have compared such set of nouns by the set obtained from the class diagram created in unified process methodology.

1 Introduction

Conceptual Modeling is research technique, the essence of which is the compensation system examined the original of modeling the system model, in order to get through experiments. With information on the original model also allows for reduced complexity and abstract from details of the real system [5]. To include a verification of the various influences on the system without risk to real systems, or monitoring changes in the system in a much shorter time. Modeling is often used in software engineering. One of a kind of modeling is a conceptual modeling.

Use cases are primarily descriptions – a textual requirements specification. They used to capture the software requirements. They depict how a user interacts with a system to achieve a goal.

In terms of in-depth view of the matching concept seems the sign of a formal system a bit exaggerated. Although just concept conceptual schema in the database world is essential, use only a limited number of concepts. Limiting case of conceptual model is natural language [4].

Outcome of those activities is the use case model. This model contains four types of components: actors, use cases, relationships, system boundary. Use Case models provide a major source of objects and classes and they are early input in class modelling [3].

Structure of this document contains following parts: Introduction, Related Works, Quality Metrics, Concept of Experiment, Discussion of Results, Conclusions and References.

* Master degree study programme in field: Software Engineering
Supervisor: Dr. Ján Lang, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 269–274.

2 Related Work

Derivation Class diagram directly from Use Case diagram is one approach (Booch et al., 1999; Jacobson et al., 1992). Use Case diagram not contains all necessary classes, attributes or methods for Class diagram but is used for validating and improving Class diagram. [1]

Bente Anda and Dag I. K. Sjoberg made a experiment where they use nouns from Use Case diagrams followed with common class pattern, responsibility-driven approach and case-driven approach. [1]

They identified two alternative approaches which derivate Class diagrams from Use Case diagrams. Directly from the use cases (called the derivation technique) and grammatical analysis of requirements documents (called the validation technique) [1].

3 Quality Metrics

Conceptual models are validated by quality metrics. One of most accurate standards for software validation is IEEE standard. In validation we will use metrics defined by IEEE methodology of software quality metrics: integrity, reliability, correctness, verifiability, survivability, maintainability, expandability, flexibility, reusability, efficiency, survivability, interoperability.[2]

In our experiment we will extract nouns from use case diagram and confront results with quality metrics. We will measure their impact on results from tokenisation.

4 Concept of Experiment

Concept of our experimental approach is described in Figure 1. We have Use Case diagrams with detailed use case description – flow of events. This events flow we will put into application which cooperates with English National Language Corpus[1]. It will choose all literary nouns from specific language.

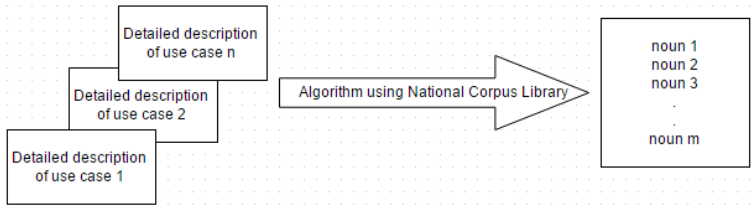


Figure 1. Experimental noun extraction using National Corpus library.

In Figure 2 is described approach used by Bente Anda and Dag I. K. Sjoberg and commonly used with RUP technique. Use Case diagrams will be transformed into sequence, communication, activity diagrams and from them will be constructed Class diagram. From Class diagram we can extract manually nouns from names, attributes and methods.

After Noun extraction from our or Anda and Sjoberg approaches we will compare extracted sets. Nouns set from our approach we will denote as M1 and other M2. As a result we can face these four cases:

$$\mathbf{M1 = M2}$$
$$\mathbf{M1 \subset M2}$$
$$\mathbf{M2 \subset M1}$$
$$\mathbf{M1 \neq M2}$$

[1] Available on: http://www.cis.upenn.edu/~xtag

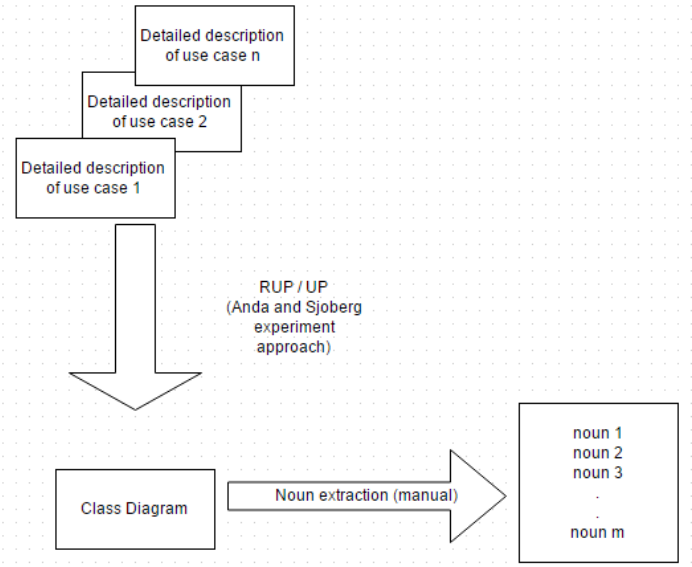


Figure 2. Extracting nouns from class diagram.

In next part of this paper we will demonstrate this concept on real diagrams and discuss about impact on quality.

5 Discussion of Results

As we can see in Figure 3 in our approach after tokenization we have extracted all nouns from detailed use case descriptions. After distinct selection and A-Z order we get them in normalized form.

In Figure 4 is Anda, Sjoberg experimental approach. As we can see they choose to build sequence diagrams and then from them class diagrams. Manual extraction nouns from Class diagram we can get list of nouns.

Let split nouns from Figure 5 to two groups. $M1$ will be left side which is result of our approach. $M2$ will be right side as a result from Anda, Sjoberg experimental method. As we can see its case where $M1 \subset M2$. But if we look closer words are different. If we hide plural words situation will change. After this step we can see that $M2 = M1$. Obviously one or more words are different.

Now we will try to compare how much are M1 and M2 similar. We will use simple arithmetic operation. Equal nouns will be nouns which are same in M1 and M2. MAX (All nouns) represents the largest group of nouns from this two. In our case it is M2. 100 is used for percentual representation of result.

$$Similarity = Equal\ nouns / MAX\ (All\ nouns) * 100\%$$

After insertion into formula we get result:

$$Similarity\ with\ plurals = 11 / 14 * 100 = 78{,}57\ \%$$

Now we will compare similarity after hide plural words. We will use same form.

$$Similarity\ without\ plurals = 11 / 13 * 100 = 84{,}61\ \%$$

As we can see similarity of both approaches is changed.

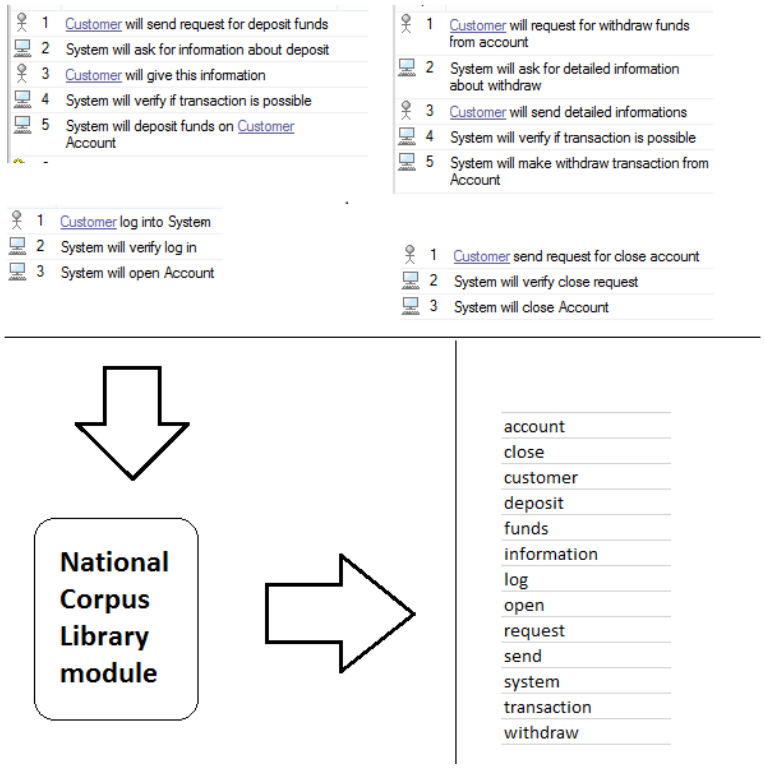


Figure 3. Experimental sample no.1.

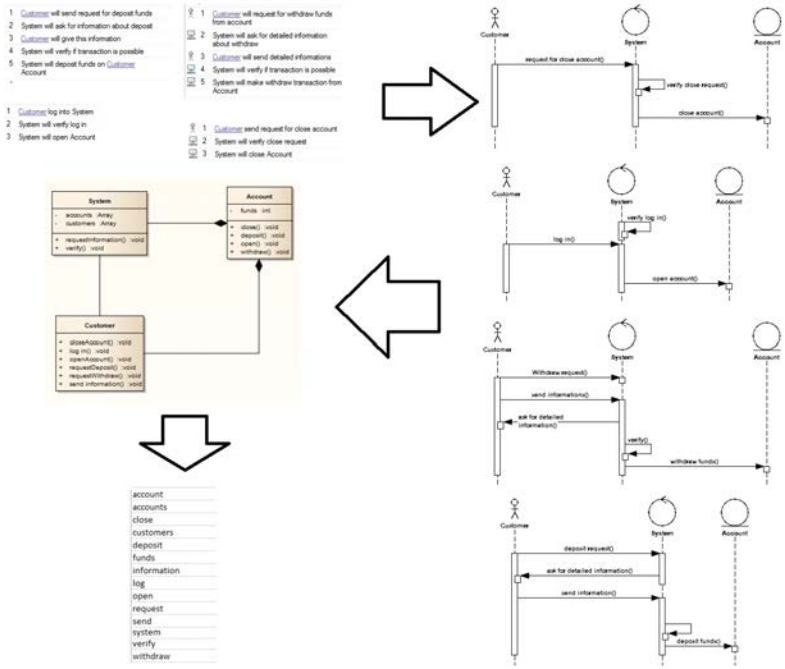


Figure 4. Experimental sample no. 2.

account	account
close	accounts
customer	close
deposit	customers
funds	deposit
information	funds
log	information
open	log
request	open
send	request
system	send
transaction	system
withdraw	verify
	withdraw

Figure 5. Sample no. 1 and Sample no. 2 nouns.

Now we will confront results with quality metrics. We are identified several quality metrics which are affected by nouns:

- Correctness: nouns extracted from detailed use case descriptions in both approaches shows us main points, base words from software specification.
- Verifiability: extracted nouns provides basic verification if model contains specific part of software specification.
- In some aspects it can affect flexibility and reusability. In identifying reusable model parts or in changing software missions, functions.

6 Conclusions

Use Cases representations of a user interaction with a system. Use Case detailed description (event flow) provides a major source for classes identification. They are early input in class modeling. Problem in Class Diagram creating is to determine correct Classes.

Derivation from Use Case diagram directly to the Class diagram is one of Jackobson and Booch approaches. On this approach experimented Anda and Sjoberg. They identified two alternative approaches which derivate Class diagrams from Use Case diagrams. Our solution continues in their experiment and expanding in grammatical analysis.

As validation technique we used quality IEEE standard of quality metrics in software engineering. Grammatical analysis is provided through Slovak National Corpus library which extracted nouns from detailed use case descriptions.

In a part Discussion of Results we compared results from noun extraction and used simply formula for comparison. In validation part we confronted results with quality metrics. We are identified several metrics which are affected by our approach.

Acknowledgement: This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG1/1221/12.

References

[1] Bente Anda and Dag I. K. Sjoberg, Investigating the role of use cases in the construction of Class Diagrams, Empirical Software Engineering, Volume 10 Issue 3, pages 285-309, ACM DL, 2005 Kluwer Academic Publishers Hingham, MA, USA, ISSN: 1382-3256

[2] IEEE Standard for a Software Quality Metrics Methodology, Institute of Electrical and Electronics Engineers Inc., 345 East 47th Street, New York, NY 10017,USA. Class e Engineering, 10:285-309, 2005.

[3] Jim Arlow, Ila Neustadt, UML a unifikovaný process vývoje aplikací (2003), Brno, ISBN 80-7226-947-X, pp. 55.

[4] Karol Matiaško, Databázové systémy (2008,) Žilina, ISBN 978-80-8070-821-4

[5] P.Márton, N.Adamko, Praktický úvod do modelovania a simulácie, ISBN 978-80-554-0387-8, pp. 9–10.

Exploring Possibilities for Symmetric Implementation of Aspect-Oriented Design Patterns in Scala

Pavol PIDANIČ*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
ppidanic@gmail.com

Abstract. In aspect-oriented programming, it may be distinguished between asymmetric and symmetric implementation. Aspect-oriented design patterns are known in their AspectJ-like, asymmetric implementation. One of their categorization is by dominating element of aspect-oriented programming – pointcut patterns, advice patterns and inter-type declaration patterns. Scala is a programming language which joins object-oriented and functional paradigm. Despite Scala is not fully aspect-oriented, it contains language mechanisms which can implement aspect-oriented feature – advice in symmetric way. This paper presents three aspect-oriented design patterns. We have demonstrated that Cuckoo's Egg and Director described in Coplien's form can be implemented in Scala. We have also described Worker Object Creation pattern in similar manner and created an implementation in Scala.

1 Introduction

Scala[1] is a multiplatform, strong static typed language that runs on Java Virtual Machine. Scala joins object-oriented and functional programming paradigm together. It is possible to import existing Java classes and libraries into Scala code.

In aspect-oriented programming, it may be distinguished between asymmetric and symmetric implementation. In asymmetric approach we distinct between the base and aspects that affects this base and base should not be aware of the aspects. AspectJ is known as asymmetric and has influenced use of aspect-oriented programming. In symmetric approach we consider how to merge things, we treat all elements equally. Elements are composed without explicitly denoting as aspects and base [2]. It has its importance in modeling. Scala contains some aspect-oriented features.

In Section 2 we describe aspect-oriented features in Scala. In Sections 3, 4 and 5 we present three aspect-oriented design patterns and their symmetric implementation in Scala. In section 6 we

* Master study programme in field: Software Engineering
Supervisor: Assoc. Professor Valentino Vranić, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

[1] http://www.scala-lang.org/

discuss asymmetric implementation of presented design patterns. In Section 7 we describe related work and conclusion of this work is in Section 8.

2 Scala and Aspect-oriented Programming

Despite Scala is not designed as aspect-oriented, it contains language mechanisms which exhibit symmetric aspect-oriented features [2]. These features partly reproduce all features of aspect-oriented programming as known in AspectJ.

The language mechanism is *trait*. Traits can be used as *mixin(s)*. With mixins we can not only define, but also provide behaviour for descendants. Traits can also define state and be combined together.

AspectJ-like advice *before*, *after*, *around* can be created with idiom *stackable modification* [9]. Advice is implemented with special construct `abstract override` in extending class.

Generally, one can handle all advice on method joinpoints. However, there is no quantification over joinpoints and pointcuts as in aspect-oriented programming [8].

Aspect-oriented design patterns are known almost exclusively in their asymmetric, AspectJ-like implementation. Design patterns should be described in general way independent of implementation details. One of the way is a Coplien's form [4] – set of definitions that describe patterns in general manner – as presented by Bálik and Vranić [1]. They implemented some aspect-oriented design patterns in HyperJ programming language.

One of the categorization of aspect-oriented design pattern is by dominating element of the asymmetric aspect-oriented programming – pointcut patterns, advice patterns and inter-type declaration patterns [6].

As described above, Scala is not fully aspect-oriented. Hence we focus only on inter-type declaration and advice patterns – Cuckoo's Egg, Director, Worker Object Creation.

3 Cuckoo's Egg Pattern

Main use of the Cuckoo's Egg is to capture a constructor call and instead create or provide an object of an another type.

Bálik and Vranić [1] created a description in the Coplien's form as follows:

Problem: Instead of an object of the original type, under certain conditions, an object of some other type is needed.
Context: The original type may be used in various contexts. The need for the object of another type can be determined before the instantiation takes place.
Forces: An object of some other type is needed, but the type that is going to be instantiated may not be altered.
Solution: Make the other type subtype of the original type and provide its instance instead of the original type instance at the moment of instantiation if the conditions for this are fulfilled.
Resulting Context: The original type remains unchanged, while it appears to give instances of the other type under certain conditions. There may be several such types chosen for instantiation according to the conditions.
Rationale: The other type has to be a subtype of the original type.

We used this description to a demonstration that the pattern can be implemented in Scala programming language.

Instead of direct instantiation of an object one could use the object-oriented design pattern Factory Method [7]. Hence we were able to implement the Cuckoo's Egg pattern in Scala as follows where we substitute instance of `Bird` with instance of `Cuckoo`:

```
trait Egg {
  def hatch: Unit
}

trait Nest {
  var eggs: List[Egg] = Nil

  def newEgg: Unit = {
    var egg = layEgg
    eggs = egg :: eggs
    egg.hatch
  }

  def layEgg: Egg
}

class BirdEgg extends Egg {
  def hatch: Unit = {
    println ("bird")
  }
}

class BirdNest extends Nest {
  def layEgg: Egg = {
    new BirdEgg
  }
}

class CuckooEgg extends Egg {
  def hatch: Unit = {
    println ("cuckoo")
  }
}

trait Cuckoo extends Nest {
  abstract override def layEgg = {
    new CuckooEgg
  }
}
```

4 Director Pattern

The Director pattern defines additional roles that are enforced onto the existing types without changing an existing implementation.

A general description of the Director pattern in the Coplien's form by Bálik and Vranić [1]:

Problem: Additional roles have to be defined in application.
Context: A type hierarchy that defines the roles.
Forces: The application has to be extended with additional roles, but the original class hierarchy in the source code has to remain free from these roles.
Solution: Introduce the additional roles as types and enforce their implementation by the corresponding types externally.
Resulting context: The type hierarchy preserved in the source code, but extended with new roles in

execution.
Rationale: Director provides two main benefits: the application behavior can be easily changed by replacing a particular concern and the core functionality is less complicated.

We used the same approach as with the Cuckoo's Egg pattern in Section 3 and demonstrated a Scala implementation of the Director pattern. The Observer object-oriented design pattern presented by Skeel [10] exhibits the Director pattern features at the same time.

Trait `SuperSensor` extends a behaviour of the class `Sensor` with a new functionality of notifying attached observers, when a value on sensor changes:

```
trait Subject[T] {
  self : T =>
    private var observers: List[T => Unit] = Nil

    def subscribe(obs: T => Unit):Unit =
        observers = obs :: observers

    def unsubscribe(obs: T => Unit):Unit =
        observers = observers filter (_ != obs)

    def publish = for(obs <- observers ) obs(this)
}

class Sensor(val label: String){
  var value:Double = _

  def changeValue(v: Double) =
    value = v
}

class Display(val label: String) {
  def notify(s: Sensor) =
    println(label + "␣" + s.label + "␣" + s.value )
}

trait SuperSensor extends Sensor with Subject[Sensor] {
  override def changeValue(v: Double) = {
    super.changeValue(v)
    publish
  }
}
```

5 Worker Object Creation Pattern

The Worker Object Creation pattern delays executing (mostly time-consuming) methods with a worker object without interfering main thread of an application.

We created a description in the Coplien's form as follows:

Problem: Delaying execution of some method out of main application thread.
Context: Order and time of an execution of some methods is not important.
Forces: Delayed execution of a method runs later.
Solution: Create an worker object and encapsulate a method and delay its execution later.
Resulting context: Methods calls are preserved, but executed later.

Rationale: Particular method calls must be out of main application thread. Worker objects can be passed to an other context.

```
class Later {
  def doLater: Unit =
    println("method execution")
}


import java.awt.EventQueue

trait WorkerObject extends Later {
  abstract override def doLater = {
     def proceed = super.doLater _

     val worker =  new Runnable() {
                def run() {
                  Thread.sleep(5000);
                  println("late execution")
                  proceed()
                }
         }
     EventQueue.invokeLater(worker)
  }
}
```

6 Asymmetric Implementation of Design Patterns

As mentioned above in asymmetric approach we distinct between base and aspects that affects base functionality and base should not be aware of the aspects. In AspectJ poincuts are used.

In Scala there is no general quantification over joinpoints, specially pointcuts definition. Spiewak and Zhao [11] introduced an experimental framework that supports pointcuts declaration. This framework uses its own Domain Specific Languages and Proxy Methods. Scala AOP framework uses non-transparent mechanism for method interception and every captured class method requires a call to a singleton delegate. Delegate function accepts a functional which represents original method body. Poincuts can capture types and names of methods and they have access to a context of instance variables and match `execution` a `within` in AspectJ programming language. Before and after-like advice functionality is supported.

As consequence that the framework lacks around-like advice we are not able to implement two of presented design pattern because both Cuckoo's Eg [7] and Worker Object Creation [5] require replacing original method execution in their original AspectJ-like implementation.

7 Related Work

In the section 3 we described the implementation of the Cuckoo's Egg pattern that uses the Factory Method pattern and in the section 4 Scala implementation of the object-oriented Observer pattern exhibits Director aspect-oriented design pattern. Bača and Vranić [3] show that there is a correlation between object-oriented and aspect-oriented design patterns. The Director pattern can substitute many patterns, indicating Prototype pattern. The Worker Object Creation substitutes Proxy pattern and the Cuckoo's Egg replaces Abstract Factory, Singleton and Flyweight.

8 Conclusion and Future Work

Scala is a modern programming language which joins object-oriented and functional paradigm that become more popular in business. The same holds for the aspect-oriented programming with main language AspectJ. Aspect-oriented design patterns are known almost exclusively in their asymmetric, AspectJ-like implementation. Design patterns should be described in the general way.

This paper presented three aspect-oriented design patterns – Cuckoo's Egg, Director and Worker Object Creation. The Cuckoo's Egg and the Director are described in the Coplien's form as presented by Bálik and Vranić [1]. We used the same description without any modification and sustained with the implementation in Scala language. We also described and provided the implementation of the Worker Object Creation pattern in the similar manner.

Although our description of the Worker Object Creation has not been sustained, the provided implementation of two design patterns advances the Coplien's form as the general description of the design patterns independent of the symmetric or the asymmetric view of the aspect-oriented programming.

A future work could focus on a description of another aspect-oriented design patterns and an implementation in Scala also with the framework. an extension of the framework is another interesting study for the future work.

Acknowledgement: The work reported here was supported by the Scientific Grant Agency of Slovak Republic (VEGA) under the grant No. VG 1/1221/12.

References

[1] Bálik, J., Vranić, V.: Sustaining Composability of Aspect-Oriented Design Patterns in Their Symmetric Implementation. In: *2nd International Workshop on Empirical Evaluation of Software Composition Techniques*. ESCOT 2011, at ECOOP 2011, 2011.

[2] Bálik, J., Vranić, V.: Symmetric Aspect-orientation: Some Practical Consequences. In: *Proceedings of the 2012 Workshop on Next Generation Modularity Approaches for Requirements and Architecture*. NEMARA '12, New York, NY, USA, ACM, 2012, pp. 7–12.

[3] Bača, P., Vranić, V.: Replacing Object-Oriented Design Patterns with Intrinsic Aspect-Oriented Design Patterns. In: *Engineering of Computer Based Systems (ECBS-EERC), 2011 2nd Eastern European Regional Conference on the*, 2011, pp. 19–26.

[4] Coplien, J.: Design pattern definition, `http://www.hillside.net/component/content/article/50-patterns/222-design-pattern-definition`.

[5] Laddad, L.: AspectJ in Action: Practical Aspect-Oriented Programming, 2003, Manning Publications Co., Greenwich, CT, USA.

[6] Menkya, R.: Aspect-Oriented Design Patterns, 2007, Bachelor Thesis, Faculty of informatics and information technologies, Slovak university of technology in Bratislava.

[7] Miles, R.: AspectJ Cookbook, 2004, O'Reilly Media, Inc.

[8] Odersky, M.: The Scala Experiment – Can We Provide Better Language Support for Component Systems? In Chin, W.N., ed.: *Programming Languages and Systems*. Volume 3302 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2004, pp. 364–365.

[9] Odersky, M., Spoon, L., Venners, B.: Programming in Scala: A Comprehensive Step-by-step Guide, 2008, Artima Incorporation.

[10] Skeel Løkke, F.: Scala & Design Patterns (Exploring Language Expressivity). Masters thesis, University of Aarhus, Department of Computer Science, 2009, `www.scala-lang.org/old/sites/default/files/FrederikThesis.pdf`.

[11] Spiewak, D., Zhao, T.: Method Proxy-Based AOP in Scala. In: *Journal of Object Technology, vol. 8, no. 7*, 2009, pp. 149–169.

Interactive Visualization of Developer's Actions

Lukáš SEKERÁK*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova2, 842 16 Bratislava, Slovakia
sekerak.luk@gmail.com

Abstract. Software product development provokes numerous unanswered questions which may not be easily resolved. Many visualization techniques do not offer the option to view the development process from the perspective of particular steps of the developer. Step-by-step approach enables visualisation of various actions taken during the pro-cess – such as utilization of a browser or communication/information channels for trouble-shooting etc. Another significant indicator to be mindful of is the transcription of own or externally acquired code. These actions are parts of software product development and their visualization may give answer to questions.

1 Introduction

In the development of software products there are many unanswered questions, which cannot always be easily answered. Many visualization techniques do not offer the opportunity to view the development of software product in terms of developer's behaviour. Which is based on collected developer actions. Through the visualization of these actions, we could identify different patterns of developer's behaviour at his work. We could identify interesting events or anomalies that have detrimental impact on development.

Actions that we may be interested in would include, for example, use of a browser, use a of communication tools for problems solving. Further development environment and etc. Because of the fact that, the developer uses many tools in his work, we will focus only on some certain actions. We will be interested in actions such as, when the developer activate the browser. Did he open it just, when he had implemented the functionality into the system, or did he copy something from a website, or did he just read the documentation from portal?

The sequence of actions consists of developer activity. As result of the analysis of work activities of the user (developer), we obtain user model as a set of user characteristics and his behaviour over the time and in the working environment. Work activity is a set of user actions for a certain period, in the context of data he is working with, the working environment, respectively tool equipment (hardware and software) computer. The activity can be understood as a demarcation of the two major events, in which the work activity acquires logical meaning.

---

* Master degree study programme in field: Information Systems
Supervisor: Dr. Alena Kovárová, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 281–286.

2 Related work

In this section we will have a look at the different views of software development. These views watch / monitor certain features of the software project. These features are then measured and assessed.

1. Tracking code structure
2. Tracking source code metrics
3. Tracking tasks (Gantt diagram) [8]

We analysed several tools in visualization of development area. For example: Source Miner Evolution (SME) [9], YARN [7], Forest Metaphor [3], ClonEvol [6], Gevol [2], EPOSee [1]. Wefound, that these tools are focused only on the code structure or metrics. Often these views were combined. However, no tool does not focuses on developer's actions or his behaviour during the development. The measured information are visualized in animation using the time line.

The analysis of work "Using developer activity data to enhance awareness during collaborative software development" by Ferguson et al. provides a good overview of 12 tools [10], for different purposes and dedicated to the developer's activities. For example, Team Tracks, the visualization tool, that displays the current activity of the developer. It display which files are currently edited and who altered them. Accessory to the IDE environment monitors which files were most frequently visited and assumes that these files can be problematic. The tool is designed for teams and management of collaborative work. None of the mentioned tools dealt with visualization of software development by a view similar to ours. Article indicates the importance of capturing data for tracking developer's activities and of course the visualization of this data.

3 Software development visualization method

Our method is based on developer's actions. These actions can be tracked by several tools, we have chosen PerConIK project. Shortly, we will describe this project, types of events and how we will build activity from these events.

3.1 PerConIK data source

PerConIK (*Personalized Conveying of Information and Knowledge, per-conik.fiit.stuba.sk* [4]). Part of the project is also a solution that records and analyses user work activity, in order to analyse behaviour of the user in the context of used software tools, that are used, but also in the context of data (documents, source code, Web...), he works with. The solution collects data on each keystroke, mouse click, browser opening and many others. Indeed, these data are linked to version control system. [5]

3.2 Activity tracker

It is an application that monitors the activities of the developer, as a user of the operating system. It is the central application in the operating system, which collects data also from other environments with additional components. These components follow document editing, web browser etc. In general, it monitors the currently running processes, open windows and the users' interaction. Tracker monitors activities also in the integrated developer environment (monitor the operations over source code).

3.3 Building activity from actions

When the developer's activity is created, it has two major events, such as the opening and the closing of tab in the browser. During this time period, the developer could look at the website, or maybe not have to. If the developer opened tab and closed it in a short period of time, it is highly

probable that the developer was on the web site. This time period will be defined by *threshold* and this will determine the creation of activities from events (actions). If an event of the same type occurs within a given time interval, we can say that the developer has worked on the activity, that is declared by the type of the event.

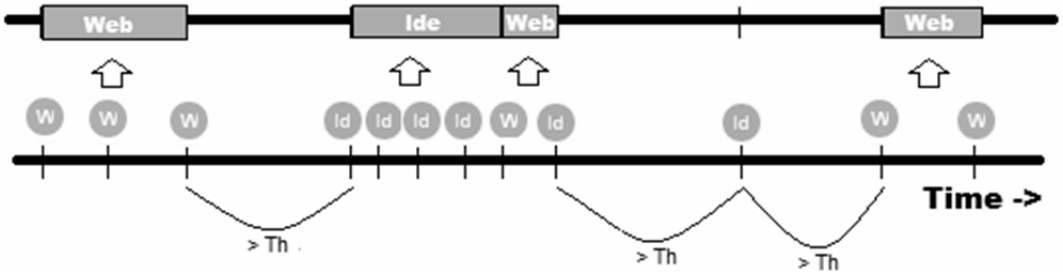


Figure 1. Building of activities from events. First line showing builded activites, second line events at time line. We use threshold value between events to determine, when event is part of activity and when it is not.

3.4 Privacy policy of developer

Captured developer's actions affect all processes of the developer's workstation. Including the web browser. Program also traces the web addresses visited by developer. However, these processes may interfere with the privacy of the developer.

4 Evaluation

4.1 IVDA tool

We have designed and implemented a tool to analyse features of activities and to visualize them. Individual features are visualized with the aid of specific graphs, focusing on data type. The user creates graphs and visually monitors the information. Tool support several types of graphs:

1. Duration of the time spent on web portals
2. Graph of source code changes (specifies when changes occur and how big they were)
3. Graph of picked features for activities (specify when activities occur, how much time did they take and shows metric determined by the type of activity)
4. Graph showing distribution of events (like a histogram)
5. Graph showing distribution of activities (focuses only on duration of the activity)
6. Graph showing changes over files and projects
7. Count of visited domains
8. Detail of run and used processes at developer's workstation
9. Graph showing grouped information for activities per some period
10. Timeline visualization showing detail of events and measured metrics

Each graph is generated as the user has requested with the help of interactive toolbar. Graphs are generated for specified developer and time period. User can interact with graph too, can set another period or see some details about events and activities.

4.2 Graph of activities and their metrics

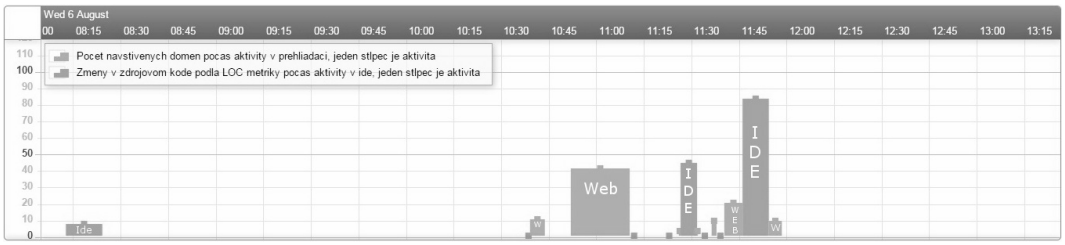


Figure 2. Graph showing developer activities at 6. August. Activities are tracked at Web and IDE enviroment.

Figure 2 shows the developer's activities within the browser and in the IDE. We have chosen the developer called Mouse, and the time period from 8:00 to 13:20, 6 August. From the graph, modified histogram, we can see what time the activities occurred. By looking at the width of the columns, we can see how long they lasted. Furthermore, we can see the characteristic of activity, according to the height of the column. For example, for the activity in the browser, we count the number of domains visited. For IDE activity, we count how many lines has changed within the activity.

Using this modified histogram, we can deduce that at 8:15 he devoted 10 minutes at Ide. Where he made the changes in the source code, and has changed 8 lines. Between 8:25 and 10:30, developer worked with something other than the Web or IDE, or has been off-line.

The developer visited the most of the domains between 10:45 and 11:10, when he visited about 40 domains. It is possible for the developer to visit such high number of domains in such a short time, for example, when dealing with a problem. At the time of 11:45 developer made a change or many changes, when he changed about 80 lines of code in 10 minutes. These changes may have been made, for example, by commenting 80 lines of source code or by refactoring. Next graphs bring next details about this change. After 12:00, the developer did not made any activity. Histogram mainly refers to the duration of activities and their order.

4.3 Hypothesis result

We defined our main aim, hypothesis. "If a developer uses a browser (a particular portal) frequently during code-writing, how often is the code rewritten? If so, does this apply to every other developer in the same manner?" One of our graphs attempts to answer to this hypothesis. We believe that this graph will have bigger signification with more data.

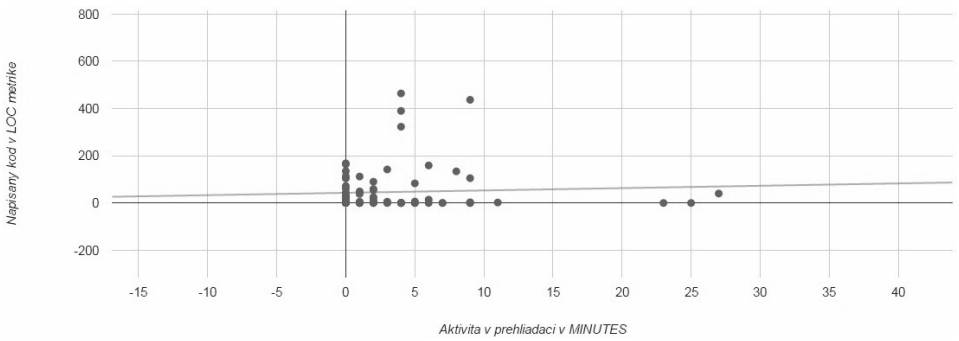


Figure 3. Scatter graph. Every point represent ratio between measured time of the Web activity (in minutes) and changes at source code (by LOC metric) in the next corresponded IDE activity. Corresponded IDE activity is activity, which happen immediately after the previous (Web) activity.

Actually usercan generate another interested graphs to insight into developer's work.

Figure 4. Timeline view of source code adjustment by specific developer (in LOC metric) on year scale.

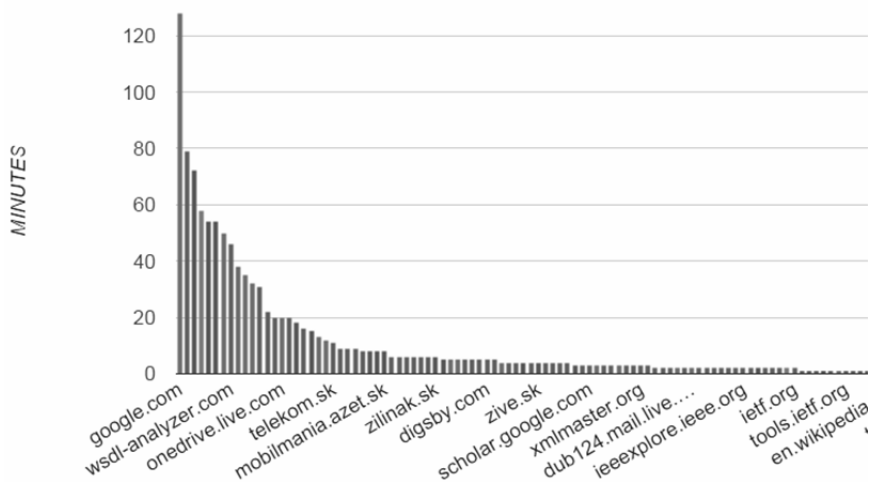


Figure 5. Visited domains and sum of time spent on them.
Data were tracked in one month of developer's work.

5 Conclusions

At the beginning of this paper we pointed out that a lot of visualization tools do not use view of developer's behaviour to visualize the development of the software. We think that this view has big potential and can be used to find answers to various interested questions. To support this thought, we have created generic visualization, where user can generate graphs. Visualization show information about activates and developer's actions and their metrics. In this paper we have showed only two graphs, although many graphs can be generated. By some changes we could create any graphs which will answer any software development questions based on developer's actions.

Our solution is based on restful service and visualization. Service is reusable for any *PerConIK* project. Visualization is web-based and can be easily part of another tool, such as a tool for managing project. Any manager may use this tool to identify anomaly and trends of their developers. For example, he can find out if using some particular portal has a big impact on developer work. Also this tool can help any scientist to answer hypothesis in software development area.

We have shown that this new view of software development is important and has the potential to be developed further. The aim was to create a tool, visualization, which would respond to the hypothesis which the other instruments could not answer. The scatter graph shows no direct answer whether the hypothesis is true or not. This may be due to lack of data. The next graphs dealt with another hypothesis, they provide better quality of view at the developer's activities.

Acknowledgement: This Software is the partial result of the Research and Development Operational Programme for the project Research of methods for acquisition, analysis and personalized conveying of information and knowledge, ITMS 26240220039, co-funded by the ERDF.

References

[1] Burch, M. et al. EPOSee: A Tool For Visualizing Software Evolution. In: *3rd IEEE International Workshop on Visualizing Software for Understanding and Analysis*, (2005), pp. 1–2.

[2] Collberg, C. et al. A system for graph-based visualization of the evolution of software. In: *Proceedings of the 2003 ACM symposium on Software visualization – SoftVis '03*. New York, New York, USA: ACM Press, (2003). p. 77.

[3] Erra, U. et al. Visualizing the Evolution of Software Systems Using the Forest Metaphor. In: *2012 16th International Conference on Information Visualisation*. 2012. s. 87–92.

[4] GRATEX PerConIK. [Online, accessed December 12, 2014]. Available at: <http://perconik.fiit.stuba.sk/>.

[5] GRATEX PerConIK : Používateľská príručka. [Online, accessed December 12, 2014]. Available at: <http://perconik.fiit.stuba.sk/UserActivity/Default/UserManual>.

[6] Hanjalic, A. ClonEvol: Visualizing software evolution with code clones. In *Software Visualization (VISSOFT), 2013 First IEEE* 2013. pp. 1–4.

[7] Hindle, A. et al. YARN: Animating Software Evolution. In: *2007 4th IEEE International Workshop on Visualizing Software for Understanding and Analysis*, (2007), pp. 129–136.

[8] Maletic, J.I. et al. A task oriented view of software visualization. In: *Proceedings First International Workshop on Visualizing Software for Understanding and Analysis*, (2002).

[9] Novais, R., Lima, C. An interactive differential and temporal approach to visually analyze software evolution. In: *Visualizing Software for Understanding and Analysis*, (2011), pp. 1–4.

[10] Roehm, T., Maalej, W. Automatically detecting developer activities and problems in software development work. In: *Proceedings – International Conference on Software Engineering*, (2012), pp. 1261–1264.

Software Services Recommendation Using Context

Ivan MARTOŠ*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
martos.ivan@gmail.com

Extended Abstract

Nowadays Service-oriented architecture (SOA) is one of the most popular techniques for creating Enterprise applications. SOA is based on using independent units of functionality called services [1]. These services are composed together so they can propose a solution for specific problem. One of the biggest advantages of SOA is that existing external services can be used for solving various problems, therefore it is not mandatory to create new service when specific functionality is desired [2].

Nowadays many different software services are available for usage. These services can be composed to create service compositions that can fulfil needs of a user. However, with a great number of services, proper selection of software service is very difficult and non-trivial task.

In this paper we focus on recommending software services. By recommendation, we want to achieve selection of software service that fulfils needs of the user as much as possible. In this work we focus on web services (WS) described by SOAP and WSDL.

Currently there are many tools that can be used for recommendation of WS, but they are often facing Cold-Start problem and recommendation is mostly created for a group of similar users, rather than a specific user. Also, most of the recommendation approaches failed in using contextual information for recommendation [3, 4].

In this paper we focus on creating a new method for recommendation of services using context. We define specific context for each service and each user separately. Information stored inside given context describes each subject individually. In this work both contexts extend base abstract context, resulting in few common attributes. Each context has also specific attributes. Using contextual information we are able to compare contexts of users and services and therefore recommend those services to user, which we consider for best suited.

Approach proposed by this paper also uses standard techniques for recommendation – Collaborative filtering and Content-based approach, resulting in Hybrid approach for services recommendation. Using Hybrid approach and contextual information, we propose a sorted list of services (sorted by supposed likelihood) as a result of users request for a service. Equation used for computation of likelihood is shown in Figure 1. If user submits request for a service, likelihood is calculated for each service separately. Likelihood of a service is calculated by arithmetical average of partial likelihoods:

---

* Master degree study programme in field: Software Engineering
Supervisor: Assoc. Professor Viera Rozinajová, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 287–288.

- Semantic – calculated using Content-based approach for recommendation. It represents similarity between functional request from user and semantic description of a service,
- Context – calculated as similarity of contexts of a user and context of a service,
- Collaborative – calculated using Collaborative filtering. It represents similarity between user and users that were satisfied with the service in the past.

Each partial likelihood is weighted and users can specify weights so each result of recommendation is tailored to their exact specifications.

$$likelihood_{total} = \frac{\begin{array}{l}(weight_{\text{semantic}} * likelihood_{\text{semantic}}) + \\ (weight_{\text{context}} * likelihood_{\text{contextual}}) + \\ (weight_{\text{collaborative}} * likelihood_{\text{collaborative}})\end{array}}{3}$$

Figure 1. Equation used for computation of service likelihood.

Advantages of our approach are:

- Elimination of Cold-start problem – since contextual information is available from the start, there is no problem with missing initial data on which recommendation can be performed,
- Recommendation focused on specific user – contextual information is specific for each user (and service) individually, thus recommendation can be performed for given specific user and not for a group of similar users.

Proposed recommendation algorithm and system are implemented. System has more than 1000 different semantic services annotated by SA-WSDL available for recommendation, obtained from [5]. System is implemented using standard and modern technologies. Application logic is implemented using EJB and database access is provided by JPA.

Main contribution of this work is the recommendation system that is based on modern and still evolving approach of recommendation using context. With Hybrid approach for recommendation we can provide accurate results, while with usage of weights in calculation of service likelihood, each recommendation can be tailored to exact specification of the user.

Acknowledgement: This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG1/1221/12.

References

[1] Erl, T.: *Service-Oriented Architecture: Concepts, Technology, and Design*, Prentice Hall Professional Technical Reference, (2005).

[2] Erl, T., Utschig, C., Maier, B., Normann, H., Trops, B.: *Next Generation SOA: A Real-World Guide to Modern Service-Oriented Computing*. Prentice Hall Computer, (2014).

[3] Lie, L., Lecue, F., Mehandjiev, N.: Semantic content-based recommendation of software services using context. *ACM Transactions on the Web*, (2013), vol. 12, no. 3, pp. 1–20.

[4] Na, Y., Jian, L., Yinliang, Z.: A Context-based Framework for Web Services Discovery. In: *2011 3rd Int. Conference on Advanced Computer Control*, IEEE, (2011), pp. 238–241.

[5] Khalid, M.A., Fries, B., Vasileski, M., Kapahnke, P., Klusch, M.: *SAWSDL TC3 WSDL11*. [Online; accessed March 20, 2014]. Available at: http://projects.semwebcentral.org/projects/sawsdl-tc/

Organizational Unit Structure and Provisioning in Federated Identity Management

Erik ŠUTA*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
suta.erik@gmail.com

Extended Abstract

Identity management is a part of software engineering that aims to provide concepts, mechanisms and technologies to work with digital identities in cyberspace in efficient, secure and privacy-respected manner. *"Identity management is commonly considered as the practice of managing features that can identify an individual uniquely within a given context."* [1] Digital identity is an electronic representation of physical persons. Digital identity is a compound of personal information about a physical person in form of attributes, credentials and identifiers. Digital identities are used to interact with software systems, whether it is in government or enterprise environment or even social networking. In general, operations related to digital identities can be described as identity lifecycle [2] that is composed of operations related to identity creation, usage, update and revocation.

While classic identity management use cases and scenarios are focused on managing identities in a single organization, the need of cooperation between such organizations is present as well. Federated identity management contains technologies and standards for such scenarios. However, current technologies and industry standards focus on federated identity management and provisioning from a single identity point of view and the concept of organizational units or organizational unit hierarchies is missing.

Current research is also mostly focused on single identities in the context of identity federations. The concept of *virtual organizations* [3] proposed by F. Vercoulen et al. notices the potential of organizational structure and allows a connection of several resources (relying parties) from various members of identity federation.

Federation space brings several interesting questions in the context of organizational units. Since federation is basically composed of identity providers of separate organizations that with high probability contain their own org. unit hierarchy, each identity provider must be able to set, which of its internal org. units should be local and which should be made available to other organizations in federation (and also select the members of identity federation it will be visible for). Also, an identity provider should be able to determine an origin of an org. unit. This ability implies what can be done with org. unit or org. unit hierarchy on different identity providers. A dynamic mechanism should allow different handling of common operations with org. units depending on the origin of such unit. A delete operation on other than origin org. unit may delete only local copy of org. unit. An identity provider may define different resource allocation for subjects from different organizations grouped in one org. unit. Org. unit hierarchy may be even

* Master degree study programme in field: Software Engineering
Supervisor: Dr. Pavol Mederly, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 289–290.

extended on other identity providers in federation to meet specific needs (see *Figure 1*). The possibilities of such mechanism are vast.

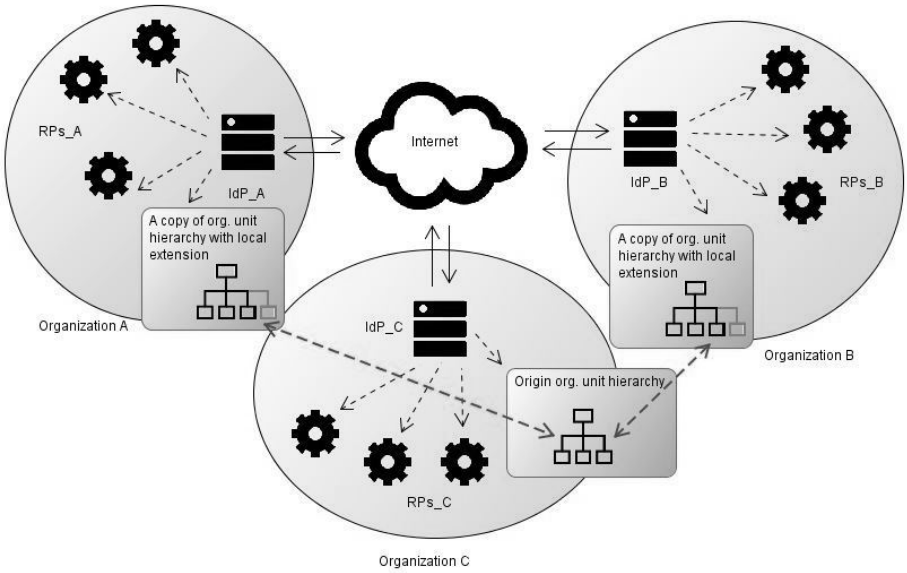


Figure 1. An example of application of org. type origin determination ability of identity providers: org. unit hierarchy in IdP_C is synchronized in federation and local copies of it on other identity providers may extend it with local org. units.

Another interesting area in context of org. units and identity federations is org. unit provisioning. When and how should we distribute local changes to other identity providers? Specific cases may require different approaches. Just-in-time (JIT) provisioning may be enough and change propagation may be distributed and evaluated only after defined action is triggered. Sometimes the distribution of change must be immediate, so pro-active provisioning is also available. In a scenario, where change propagation affects thousands of subjects on several systems, we may want to trigger it in nightly hours to minimise risk of system unavailability. An identity provider should be able to define a type for provisioning for org. unit, or even for a specific action with org. unit.

Acknowledgement: This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG1/1221/12.

References

[1] Leskinen, J., Evaluation Criteria for Future Identity Management. In: *IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*. Liverpool: IEEE, 2012, pp. 801-806. ISBN 978-1-4673-2172-3Friedman, A.D., Menon, P.R.: *Theory and Design of Switching Circuits*. Computer Science Press, Inc., (1975).

[2] Bertino, E. Takahashi, L., *Identity Management Concepts, Technologies, and Systems*. Norwood: ARTECH HOUSE, 2011. ISBN 978-1-60807-039-8.

[3] Vercoulen F., Badger M., Monaco L., Fransson T., Farinetti L., Corno F., *Deliverable 5.4 – Virtual Campus Hub technical evaluation report*, Education and Student Service Center, University of Technology, Eindhoven, 2013.

Employing Software Developer's Activity and Gaze Tracking to Describe Source Code

Martin KONÔPKA*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
martin_konopka@stuba.sk

Abstract[1]

Software developers use integrated development environments (IDE) to increase their efficiency in navigation and comprehension of source code. Monitoring their interactions in an IDE is nowadays used not only for modelling developer's knowledge or experience, but also to describe source code.

The most intuitive way to comprehend source code is to use source code dependencies. Source code dependencies are traditionally identified using syntactic source code analysis. Developers follow dependencies when navigating or navigate with their own intent, even at random. This inspires us to uncover *potential dependencies* between components solely from interactions of developer's navigation, or identify which dependencies are more or less important for development and maintenance tasks. As an example, we distinguish three kinds of potential dependencies based on the sourced interaction – *navigation, copy-pasting, commit*.

The most important drawback of identifying potential dependencies from interactions is their high level of granularity. As a solution, and opposed to existing works, we track developer's viewport in source code files to identify real dependent parts of source code components. To further specify components, we propose to use gaze tracking, e.g., with Tobii Eye Tracking technology. This brings us to the granularity of source code statements without a syntactic analysis.

Monitoring a software developer in an IDE provides fine grained interaction data about his work although without reasons why those interactions happened, i.e., tasks. Our intention is to infer developer's activities from interaction data and then automatically identify his tasks in real time, to describe his work on a task. Inferred developer's activities may also be used for weighting potential source code dependencies that we identify from interaction data.

We see possibilities of our ideas for untangling changes in source code even before committing them, as opposed to existing works, and also to provide dependencies for dynamically-typed programming languages or even when multiple languages are used in source code.

Acknowledgement: This contribution is the partial result of the Research & Development Operational Programme for the project Research of methods for acquisition, analysis and personalized conveying of information and knowledge, ITMS 26240220039, co-funded by the ERDF.

---

* Doctoral degree study programme in field: Software Engineering
Supervisor: Professor Pavol Návrat, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

[1] Full paper available in printed proceedings, pages 291–298.

Eventually Consistent Distributed Transactional Memory

Aurel PAULOVIČ*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
aurel.paulovic@stuba.sk

Abstract[1]

Achieving consistency in distributed systems is a notoriously difficult task that instigated the research of various concurrency control mechanisms. One of such mechanisms is the distributed transactional memory which uses transactions as a concurrency primitive that wraps a block of computation, making it execute in respect to other concurrently running transactions as a single indivisible atomic and isolated operation.

In contrast to the more traditional approach of optimising a strong consistency model for distributed systems, we built our transactional memory on the basis of an efficient and inherently scalable relaxed consistency model using eventual consistency. This allows us to better cope with the constraints and properties of an unreliable distributed computing environment with high latency, frequent node failures and network partitions. To overcome the disadvantages of relaxed consistency we gradually make the model more usable and easier to program in by stacking additional concurrency control mechanisms.

To aid the user with resolving conflicts between concurrent updates to data we employ shared replicated data types that can automatically merge divergent data replicas. We use a simple grouping construct of eventually consistent transactions with isolation of transient operations to ease the reasoning about shared data and to help maintain some data invariants. To temporarily increase the strictness of the consistency model we devise a form of optional locking that allows to dynamically enforce update-serialization for selected data and transactions while preserving the scalability and performance of the rest of the system.

Using these concurrency constructs enables us to build efficient fault tolerant systems that achieve high availability even in the presence of network partitions while providing the programmer with and understandable consistency model, which can be used in many real-world distributed systems.

Acknowledgement: This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG1/1221/12.

---

* Doctoral study programme in field: Software Engineering
Supervisor: Professor Pavol Návrat, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

[1] Full paper available in printed proceedings, pages 299–306.

Computer Networks, Computer Systems and Security

Autonomous Mapping Platform

Zoltán RUSNÁK*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
zoltan.rusnak@gmail.com

Abstract. It has been a while since first machines replaced people in some tasks. This trend went on and nowadays we can employ machines almost in every field. However, some machines called robots are much more complicated because of need to sense the environment and make an appropriate decision based on some events. For many robots it is important to map the unknown environment so they can securely move in it, avoid the obstacles and successfully complete the task. In this paper we present hardware design of the autonomous mapping platform which can navigate itself in unknown environment and gather data from sensors in order to construct the map of the surrounding world.

1 Introduction

Environment mapping is a process of creating models of the world around. These models are also called maps and they are needed for many applications such as exploration, transportation, rescue and many other tasks. Mapping process consists of two main parts that must be solved in order to create precise map of the unknown environment. The first part is localization and it is a name for the problem of determining the robot's position on the map. To solve this problem is usually used GPS technology or rotary encoders mounted on the robot's wheels. The second part is the mapping itself and it is the problem of gathering information from sensors as well as the integration and interpretation of them. For the measurement of the distances between robot and obstacles are most commonly used ultrasonic, infrared or laser distance sensors.

It is not so difficult to determine the position of the robot on the map if the map is known as well as to create a map of environment if there is a way to know the exact position of the robot. However, if these two problems need to be solved at the same time and they are dependent on each other it is becoming more difficult to find the solution. Solving these two problems together is referred to as the simultaneous localization and mapping (SLAM) problem [1].

Robotic mapping became very popular area in robotics during last three decades therefore at present there exist several robust methods for mapping environment which can be applied in order to accomplish this project.

---

* Bachelor degree study programme in field: Computer Engineering
Supervisor: František Kudlačák, Institute of Computer Systems and Networks, Faculty of Informatics and Information Technologies STU in Bratislava

2 Related work

Research of robotic mapping was divided into two main parts according to representation of the map. One way of representation of the map is metric and the other one is topological. Both representations have some advantages and disadvantages and the use of one of them depends on what type of the map is needed or sufficient for completion of the concrete task.

2.1 Topological maps

Topological map represents the environment as some group of places which are interesting in some way. These places are visualized as nodes and they are connected together via arcs which correspond to paths between two places. Together these nodes and arcs form graph which can also store some other information like name of places or distances between them [2].

This type of the map is not so accurate in geometric way but it does not matter because it is not its primary role. The primary role is only to provide navigation between certain places. The advantage of topological maps is that they can be used more efficiently than metric maps because to solve the problem of navigation is only needed to find the route from one node to another one. Moreover, they do not need so much information to be stored in memory and also not so much computational power. On the other hand there are also some disadvantages of topological type of the map. Sometimes there can be problem with recognition of the same place so once visited place can be considered as new one which was not visited before. Another disadvantage is that routes between places are not always optimal therefore the robot must always travel longer distances than it should.

2.2 Metric maps

Metric type of the map for the visualization of the environment uses occupancy grid. It is two-dimensional array of cells. These cells have square shape and each of them represents a piece of the real world around the robot. The side of the cell can be of any size therefore it can represent a piece of the environment of any size. However, the smaller the size the better because then the resolution and the accuracy of the map are better too. Each cell is also described by one or more values which determine the probability of that whether the cell is occupied or free. If the value of probability that the cell is occupied is above some threshold that was given the cell is considered occupied, otherwise it is considered free[3]. The graphical visualization of the real environment is created by the colouration of the cells depending on these probabilities. If the cell is occupied it is coloured black, otherwise it is coloured white or grey. Then there are all obstacles clearly visible on these maps therefore they are much more accurate than topological maps. Furthermore, humans can better understand this representation of the map.

Metric maps are built in different way than topological ones. The cells of occupancy grid are updated after each reading from sensors. For the correct update of the cells and fusion of data from sensors there were invented some algorithms. The most popular ones of them are Bayesian, Dempster-Shafer and HIMM[4]. First two of them can be used to process data from any type of the distance sensors but the third one is designed for the processing of data from ultrasonic distance sensors and it is considerably simplified.

As the advantage of metric maps can be considered their high accuracy and consistency therefore their usages is very popular for the mapping of indoor environment. The main disadvantage is need of large amount of memory and computational power which makes it unusable for large outdoor environment.

3 Hardware design

The robot needs to be designed to minimize errors in odometry. The right step to ensure that is usage of the tracked wheels rather than ordinary wheels. Tracks provide bigger area of contact

with surface so the slippage is reduced. Another requirement is that the robot can move autonomously but it can be also controlled by remote control. Hardware components and connections between them can be seen in Figure 1.

Regarding microcontroller there are also some needs that have to be fulfilled. Large number of digital pins is needed because of ultrasonic sensors and motors. All of them need two digital pins to be controlled. However, two motors on each side of the robot can be controlled together by two pins, so the number of pins needed for controlling motors can be reduced to the half. Another pins are needed to connect output from RC receiver and rotary encoders to the microcontroller, so about 16 digital pins are needed. Next requirement is serial interface for communication with wireless transceiver.

For the control of the motors is used a pair of H-bridges and each of them controls two motors. Rotary encoders are mounted on the wheels and provide information about movement of the robot. RC receiver provides the possibility of controlling the robot remotely by measuring the length of the pulse transmitted by RC transmitter.

To measure the distance tothe obstacles four ultrasonic distance sensors are used. They are placed on each side of the robot. The measurement of distance is based on time of flight of the generated sonic wave so the delay is noticeable but the measurement should be very accurate.

By the wireless transceiver can be all data from ultrasonic sensors and rotary encoders transmitted in real time to PC where they are processed.

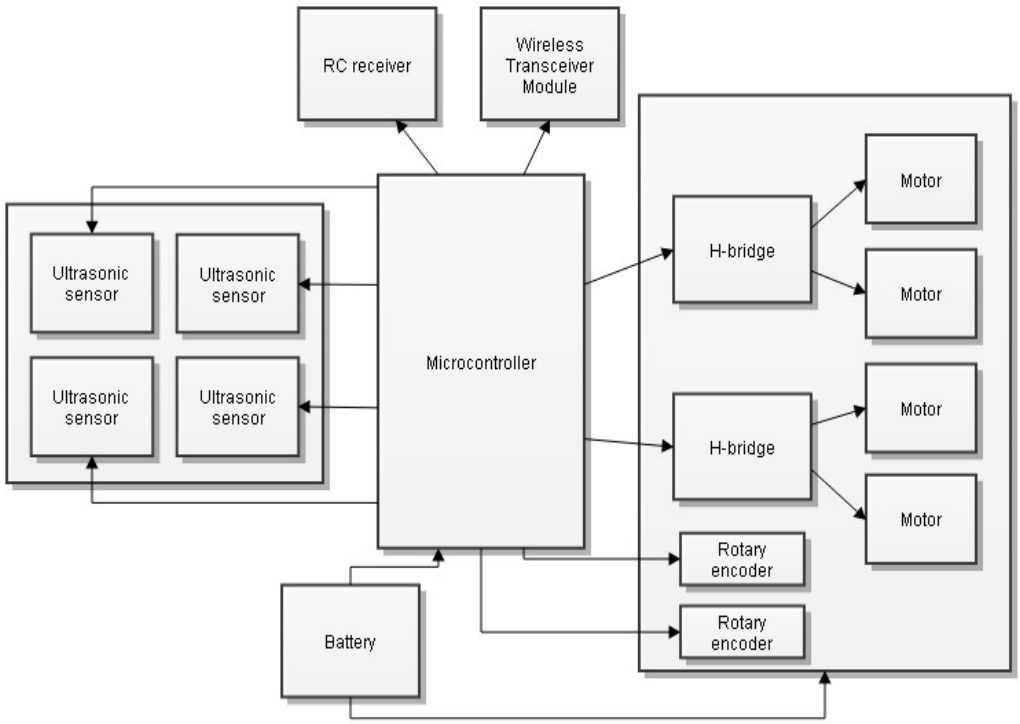


Figure 1. Scheme of the connections between hardware components.

4 Design of control system

The control system of the mapping platform needs to perform five main functions to accomplish autonomous mapping correctly. The first one is the collecting of raw data from sensors and also processing of them. The second function is communication with desktop application and sending

measured data to it for further processing. Another one is a check if anyone is not trying to control it remotely. Performing moves based on measured data or command from remote control is the third function. The final function is a check whether the starting point has been already reached and it should end mapping process or if it does not move in circles. Basic diagram of this mapping process is shown in Figure 2.

Main idea of the design of the autonomous movement is the following of the wall along the right side of the robot. The robot checks measured distances obtained from the sensor on the right side if it is within the limits. If it is too far from the wall, it needs to turn right and get closer to it. If it gets too close, it will turn left and go away from the wall.

The main advantage of the wall following method is its simplicity, but there is also a disadvantage, that the robot can move in circles, so it needs to maintain its own map and check its own track in it.

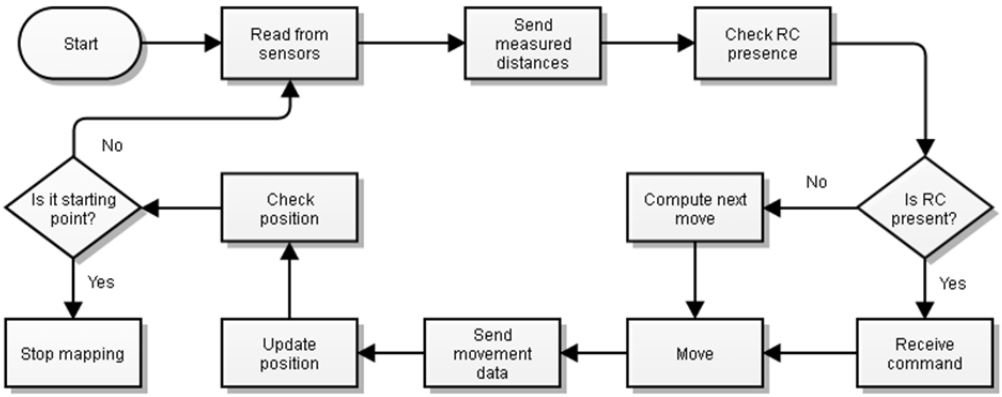


Figure 2. Control system flowchart.

5 Design of desktop application

The map of the environment can be implemented as some grid layout. Every region of this grid has opacity value that can be set according to occupation probability of that concrete region. Regions that are probably occupied are filled with gray or black colour and free regions are coloured white.

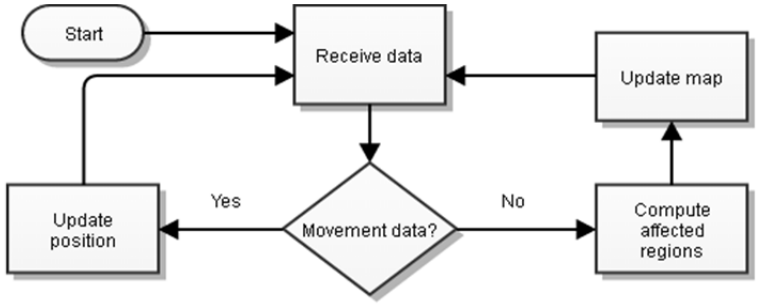


Figure 3. Desktop application flowchart.

An application can consist of many threads. One thread can perform the receiving of the data from the robot and when they are received it creates new thread that would compute affected regions and also update their opacity values, so it can receive another transmitted data. The flowchart of the application can be seen in Figure 3.

6 Implementation

For the construction of a prototype of this embedded system the following components were chosen. To meet the requirements for the microcontroller Arduino Mega 2560[5] seems to be a good choice. It provides enough digital pins as well as many serial interfaces.

As the H-bridges L293DNE were chosen. They are very popular and easy to use in case of controlling motors by Arduino.

The body of the robot consists of Rover 5 chassis. It has four wheels with motors which can be used with or without tracks. Besides there are rotary encoders mounted on all wheels.

The type of the ultrasonic distance sensors is HC-SR04. They are cheap but accurate enough.

To remotely control the robot HobbyKing HK6S can be used. It uses frequency of 2.4 GHz and 6 channels.

For the communication with PC wireless transceivers modules APC220 were used. Their working frequency is about 450 MHz and transmit distance is up to 1000 m what is sufficient for indoor environment.

The capacity of the LiPo battery used to power the robot is 2200 mAh, voltage 11.1 V and it consists of three cells. Hardware configuration of chosen components is shown in Figure 4.

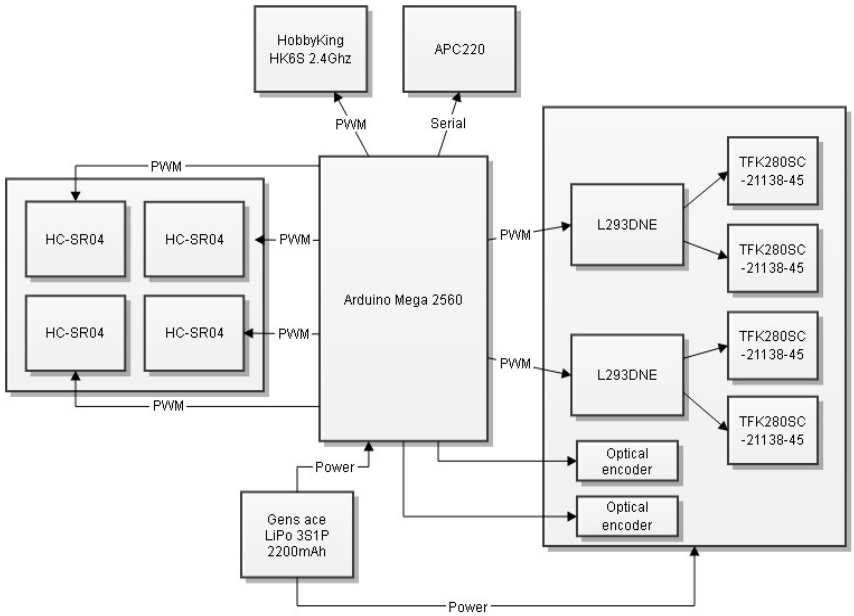


Figure 4. Hardware configuration of used components and their interfaces.

So far, the remote control of the robot is implemented. It can perform basic moves such as rotations to both sides and going forward and backward and it also can move diagonally. The robots reactions are fine and the delay between the command and reaction is about 86 milliseconds.

Autonomous control based on measured data is also partially implemented but it is not complete yet. The prototype of the mapping platform is shown in Figure 5.

7 Conclusion and future work

In this paper we proposed possible hardware configuration of the mapping robot. The robot is operational and can be controlled remotely. It can also gather information from distance sensors

and send them to PC. Used ultrasonic sensors can detect obstacles in the range of 2 cm and 3.8 m. The prototype reaches speeds up to 0.2 m/s. The tracks allow it to rotate pretty fast so it can turn around its own axis in about 6 seconds. The prototype can be powered by battery for approximately 1 hour of continuous mapping. The program for microcontroller was written in special programming language for Arduino boards which is very similar to C++ programming language. Visual Studio 2012 with Visual Micro plug-in was used to write the code.

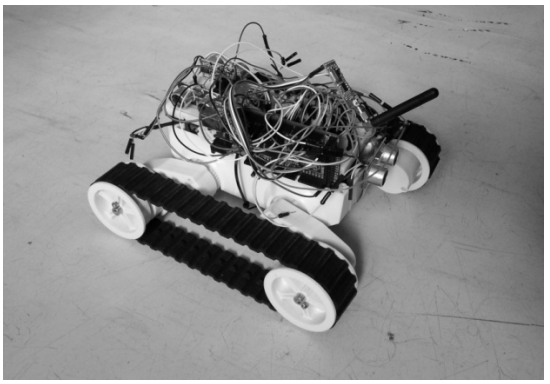


Figure 5. Prototype.

In future work we want to complete autonomous navigation so the robot can map the environment on its own and also localize itself in it. Final goal is to implement application for PC which will process received data and construct a map of the environment. In order to improve the accuracy of the constructed map the ultrasonic sensors can be replaced by rotary lidar sensor. It would also reduce the time needed to perform one measurement. Another possibility is to upgrade the robot by adding the camera and robotic arm so it can be used for transportation of some objects and also stream the video.

Acknowledgement: This work was partially supported by the Science and Technology Assistance Agency under the contract No.VEGA 1/0676/12.

References

[1] Stachniss, C.: *Autonomous Intelligent Systems.* [Online; accessed 14 February 2015]. Available at: http://ais.informatik.uni-freiburg.de/teaching/ws13/mapping/index_en.php.

[2] Thrun, S.: *Robotic Mapping: A Survey.* [Online; accessed 17 February 2015]. Available at: http://robots.stanford.edu/papers/thrun.mapping-tr.pdf.

[3] Font, D. B.: *A Multiagent Approach to Qualitative Navigation in Robotics.* [Online; accessed 16 February 2015]. Available at: http://eia.udg.es/~busquets/thesis/thesis_html/thesis_html.html.

[4] Murphy, R. R.: *Introduction to AI Robotics.* Cambridge, MA, USA : MIT Press, (2000).

[5] Atmel.: *8-bit Atmel Microcontroller with 16/32/64KB In-System Programmable Flash.* [Online; accessed 18 February 2015]. Available at: http://www.atmel.com/Images/Atmel-2549-8-bit-AVR-Microcontroller-ATmega640-1280-1281-2560-2561_datasheet.pdf.

Host Discovery and Monitoring in SDN

Jozef BALÁŽ*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
jozef.balaz.237@gmail.com

Abstract. In this paper we introduce a new way of active monitoring and protecting the local network thanks to a new concept called Software Defined Networking (SDN). The programmable character of SDN [2] enables us to create new solutions for unresolved problems existing in the traditional networks, specifically ARP spoofing attacks and MitM attacks based on them. The main idea is to build an active IDS/IPS solution able not only to detect attacks based on heuristically monitoring hosts connected to the network, but also to further inspect the suspect to confirm the attack, prevent the attacker from further accessing the network and finally minimize the damage by healing the victim's ARP cache.

1 Introduction

Software Defined Networking (SDN) [6] is a relatively new concept changing the way we see computer networks. While in traditional networks all of the intermediary devices worked more or less independently and were mostly built on proprietary software, SDN is different [4]. Here we use a central component called controller, which contains the whole application logic. This controller then manages multiple "dumb" switches. SDN is often misunderstood to mean OpenFlow (OF) [5] however this is not entirely correct. OpenFlow is only the protocol defining an interface between a controller and a switch in SDN.

ARP spoofing attacks are a well-known issue of the traditional networks, which was not yet satisfactory solved. The attacks are being realized by the attacking host sending ARP replies with false MAC/IP mapping pretending to be a different host in the network. These attacks are a way of implementing Man in the Middle attacks, where the attacker is in the middle of a communication between a victim and a legitimate host.

In our paper we propose a new solution based on a pro-active approach without the need for a user to manually pre-configure the mapping. We rather let the controller dynamically gather the necessary information by itself and share it among all of the switches. This way we can utilize the control over the whole network rather than implement it on a per switch basis. Based on this information we are able to identify an attack and further inspect an attacker. In case of a confirmed attack we isolate the attacker and speed up the restore process of ARP cache of the affected hosts with ARP replies.

---

* Master degree study programme in field: Computer Engineering
Supervisor: Pavol Helebrandt, Institute of Computer Systems and Networks, Faculty of Informatics and Information Technologies STU in Bratislava

The rest of this paper is organized as follows: Section 2 analyses the current state of the issue and already existing solutions. Sections 3 explain us the design of our solution. Section 4 describes the implementation of our design together with a testing scenario used for evaluation. Finally in Section 5 we conclude this paper by summarizing the current state of progress and setting future objectives.

2 Related work

The topic of preventing ARP spoofing attacks is not a new one. Existing solutions are mostly based on static pre-configuration of addresses, which is not very practical. With SDN on the rise new attempts have been made to tackle this problem [1, 8]. The main idea is to passively learn information about the hosts from the processed packets and in case the source IP and MAC address don't match the cached information the port gets blocked. This solution has however the following disadvantages: The host learning is done only by passive listening to traffic resulting in an incomplete database. In case of a problem only a warning is issued or the port is disabled without any further investigation on the attack or resolving the damage. They are lacking support for multiple interconnected switches.

Authors in [3] propose a solution in a traditional network based on extending the hosts with a special algorithm, which would first check if the mapping suggested in the ARP reply is correct. The main idea of the algorithm is to generate custom traffic and see if we get a reply or it gets forwarded. In the first case we know that we can trust the host and establish a connection with it adding the mapping to our cache otherwise the reply is ignored.

This algorithm has however one major flaw. The CAM table of the switch must not be changed during our check. This problem is unavoidable in traditional networks making the solution unreliable.

In this paper we propose a solution to these deficits by creating a reliable pro-active security system thanks to the direct access to the switches with a high level of automation thus reducing the need for manual intervention in case of an attack.

3 Solution design

In the following section we will explain the process of discovering the hosts in the network together with our algorithm for selecting target addresses for the discovery. The section then continues with an explanation of our improved method used for detecting, confirming, mitigating the threat and restoring the previous state.

3.1 Network information gathering

After a controller boots up it has no information regarding the underlying topology. In order to gain a complete database of hosts we need to pro-actively start discovering them.

All of the switch ports do not necessarily have to be directly connected to hosts. Some of the ports can be used as trunk ports interconnecting our switches. To get this information the first thing to do is to periodically send discovery packets. A suitable protocol for this purpose is the Link Layer Discovery Protocol (LLDP). During this process the controller does not learn any new information regarding the hosts. After the timer expires we can start learning hosts by sending ARP pings from all the non-trunk interfaces.

3.2 Discovering individual hosts

An effective host discovery has always been a difficult task since the hosts do not always respond to standard ICMP requests or we often do not have enough information regarding the network. An interesting source of inspiration can be considered the propagation of worms in the network [7].

In general we have to expect all possible addresses. Nevertheless there are some options to narrow this scope down. The first thing to consider is that not all address ranges can be used to identify a certain host in a global scope like: loopback, link local, multicast and broadcast addresses. After excluding the addresses there is a second step we can take. The Dynamic Host Configuration Protocol (DHCP) can be a useful source of information. Among other information we can find here the needed network address and network mask. It is enough just to send a DHCP Discover message and extract what is needed from the Offer message without continuing with a Request response.

The worst-case scenario is when we do not have any information at all. In such a case we have to randomly pick from all the available addresses. It might be advisable to prefer private IP addresses due to the fact that they are not unique, thus resulting in a higher probability of them being assigned. After a certain amount of time we should receive our first packet. This is the crucial part because now we have our first IP address and we can start discovering hosts "around" it.

The idea of the algorithm is to take an IP address and start with the mask set to /32. Next we decrease the mask and change the last bit previously covered by the mask. Next we ping all of the addresses included in our newly gained network address range. We repeat this process until we get stopped by the obvious boundaries like for example the private address 10.0.0.0 having the network mask /8.

3.3 Attack detection

After the initial host discovery we should eventually get a list of the individual hosts. The flows have to be implemented so that a Packet In notification is only sent when the source IP, MAC address or port do not match. In case of such an event we have to be able to confirm that we are being attacked.

If we are going to block a host we need to make sure it is really an attacker. One of the ways to do so is the Reverse ARP poisoning [3]. The idea is that in order for the MitM attack to work it is not enough just to poison the victim's cache but also act as a relay.

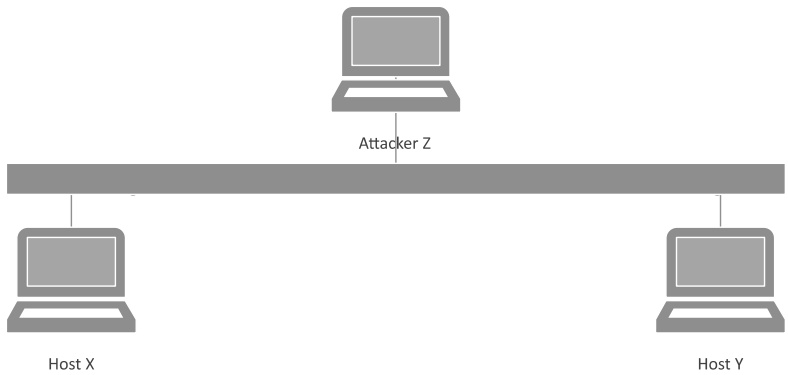


Figure 1. Reverse ARP poisoning topology.

The whole process starts after receiving a suspicious packet with the source IP_Y, MAC_Z on a specific switch port as shown in Figure 1. After detecting such a packet we are sending an ARP reply message towards the potential attacker with the mapping IP_Y and MAC_X. In case our target is really the attacker his ARP cache should be now poisoned, otherwise nothing should change.

The next step is to verify if our poisoning attack has been successful. We can do so by sending an http request with the destination of IP_Y and MAC_Z. In case our request gets

forwarded we can be sure that the host is the attacker and we can follow up with preventing the attacker from accessing the rest of the network.

3.4 Attack prevention

After an attack is detected we need to cut the attacker of the network so that he cannot cause any harm by inserting flows either blocking all of the traffic coming from and to the port or just filtering his source address.

Even though the attack might have been mitigated the victim's ARP cache might still be poisoned. In order to speed up the recovery process we can start sending ARP replies to fix this thanks to our knowledge of the correct mapping.

4 Evaluation

The section starts with a description of the current implementation. Further we can find a testing scenario which is used to prove the fulfillment of functional requirements set by us.

4.1 Implementation

The proposed solution is designed to work as a set of plugins for the open source controller OpenDaylight.

In order for the solution to be able to work with different L2switch implementations our application logic runs within one separate switch table and the forwarding logic in a different one. The forwarding logic is currently simulated by a hub implementation and is intended to be replaced by a switch in the final product for testing.

The first of the plugins is concerned with Host Discovery. After booting up the switch it starts sending LLDP (Link Layer Discovery Protocol) messages to detect and get an overview of the connected switches and their interconnection. Afterwards we start discovering hosts on the non-interconnect ports by sending ARP requests creating a complete list of hosts. A flow is set up for packets from the known host (source port, IP, MAC address) to be further processed by the hub in the second table. Otherwise a notification is sent to the controller with a specific cookie used to identify the source of the problem. Host information is defined as a YANG model and is stored within the Data Store of the switch so that it can be used by other plugins as well like our security plugin. The information is accessible from the outside via REST API.

The security plugin listens for theses notifications and after receiving one it has to determine if we are being attacked. We send an http request towards the attacker and listen for a response. In case we receive our request forwarded we can be sure it is the attacker and can install a flow with a drop action blocking any further communication either on the port or from the specific MAC address. Since the victim's cache is still poisoned we speed up the recovery process by sending multiple ARP replies with the correct mapping learned from our Host Discovery plugin decreasing the recovery time.

4.2 Testing scenario

The testing of the implementation is done on a simulated topology generated by Mininet. Mininet enables us to create custom scripts in which we define the switches, hosts and connections between them. Another important feature is that we can map an interface of our virtual machine to one of the switch interfaces so that we can connect external hosts to our topology, which is useful when simulating the attacks.

The testing is done on the topology shown in Figure 2. In this scenario we have all of the switches under control connected to the controller. Also all of them are interconnected, thus creating a single local network. Ettercap will be used to simulate the ARP spoofing attack.

The first objective is to test the Host Discovery system. In order to do so we generate random IP addresses and assign them to the hosts. Eventually we expect for all of them to be discovered.

After a successful result in the first test we can move to the main part which is the security system implementation testing. At the beginning of the test we check the content of the victim's ARP cache, which should be currently empty. We also verify the connection between the attacker and the future victim. Then we can proceed with a MitM attack with the Ettercap tool. A warning should be logged informing us about a mismatch in the MAC/IP address mapping. By now the victim's ARP cache should be poisoned. After detecting the mismatch in the mapping our request probes should be captured on the attacker's interface, which will be forwarded due to the MitM attack. Based on the forwarded packets communication from the attacker should be dropped by a flow. We also start immediately healing the victim's ARP cache. The test should result into the attacker being unable to communicate with the rest of the network and the host having a correct mapping in the ARP cache.

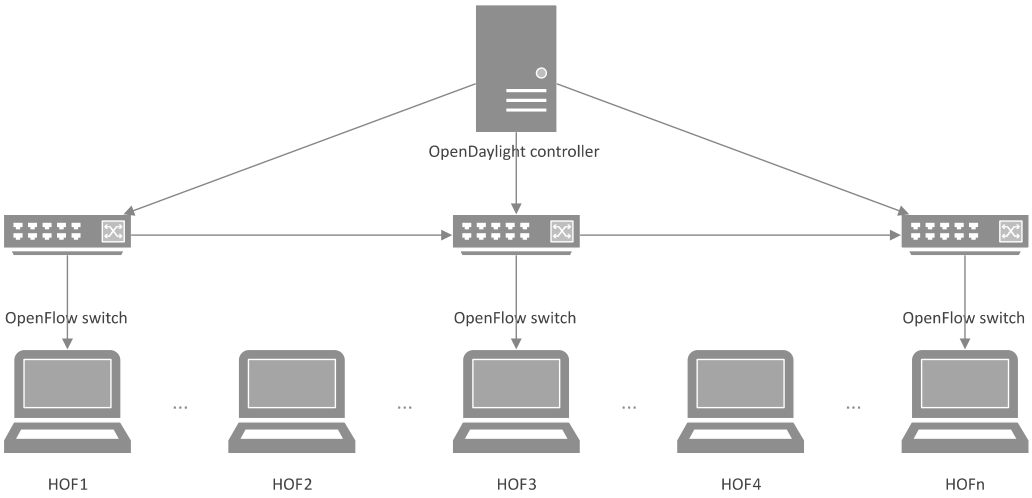


Figure 2. Testing topology consisting entirely of OpenFlow components.

5 Conclusion

In this paper we have introduced the issue of ARP spoofing attacks. We have compared the different approaches in traditional and software defined networks and therefrom emerging new solution to this problem.

We have prepared a design and implementation of a network host discovery system focused on efficiency and low network resources consumption.

The information gathered by the host tracking system enables us to have a complete overview of the network and base on it an Intrusion Detection/Intrusion Prevention system (IDS/IPS), which is currently being finalized.

The finalized implementation will result into a single security system enabling multiple switches to cooperate together using a shared database fully utilizing the complete overview of the network. It will make the work of a system administrator much easier due to the high level of automation. There will be no need for manual rules pre-configuration. The controller will also be able to restore the correct mapping in the ARP cache of the victims faster by itself instead of just waiting for the entry to expire.

References

[1] Crenshaw, A.: Security and Software Defined Networking: Practical Possibilities and Potential Pitfalls [Accessed March 19, 2015]. Available at: http://www.irongeek.com/i.php?page=security/security-and-software-defined-networking-sdn-openflow

[2] Feamster, N., Rexford, J., Zegura, E.: The Road to SDN: An Intellectual History of Programmable Networks. In: *Queue – Large-Scale Implementations.* ACM, New York, NY, USA, (2013), pp. 87–98.

[3] Kalajdzic, K., Patel. A.: Active Detection and Prevention of Sophisticated ARP- Poisoning Man-in-the-Middle Attacks on Switched Ethernet LANs. In: *Proceedings of the Sixth International Workshop on Digital Forensics and Incident Analysis* (WDFIA 2011). Lulu. com, 2011. p. 81.

[4] Nadeau, T.; Gray, K.: Software Defined Networks. First Edition. O'Reilly Media, (2013). ISBN 978-1-449-34230-2

[5] Nygren, A. et al.: OpenFlow Switch Specification. version 1.4.0. 2013. [Accessed May 10, 2014]. Available at: https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-spec-v1.4.0.pdf

[6] Software-Defined Networking: The New Norm for Networks, ONF, White Paper, (2012).

[7] Xiang, Y., Fan, X., Zhu, W.: Propagation of Active Worms: A Survey. In: *International Journal of Computer Systems Science & Engineering*, (2009), 24.3: 157–172.

[8] You, W. et al.: OpenFlow Security Threat Detection and Defense Services. *International Journal of Advanced Networking & Applications*, (2014), 6.3: 2347–2351

Port Control Protocol in Software-Defined Mobile Networks

Kamil BURDA*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
xburdakamil@gmail.com

Abstract. User applications, such as VoIP, have problems traversing NAT gateways or firewalls. To mitigate these problems, applications send keepalive messages through the gateways. The interval of sending keepalives is often unnecessarily short, which increases the network load, especially in mobile networks. Port Control Protocol (PCP) allows the applications to traverse the gateways and to optimize the interval. This paper describes the deployment of PCP in software-defined networks (SDN) and proposes a method to measure keepalive traffic reduction in mobile networks with PCP deployed. The proposed solution reduces traffic overhead in mobile networks and extends battery life of mobile devices.

1 Introduction

User applications that require long-term connections, such as Voice over IP (VoIP), Instant Messaging or online gaming, may have problems establishing connections if hosts running the applications are located behind network address translation (NAT) gateways or firewalls, hereinafter referred to as *middleboxes*.

For each connection, a middlebox contains a mapping entry that is manually configured or dynamically created when the connection is being established. In case of NAT gateways, the mapping entry usually consists of the following fields: internal IP address, external IP address, internal port, external port and mapping lifetime.

If a connection is idle for longer than the corresponding mapping lifetime, the middlebox blocks the connection, without any communicating host being notified. To keep the connection alive, the application sends keepalive messages (such as empty TCP or UDP datagrams) toward the destination host. Because the application does not know the exact connection timeout, keepalives are sent in very short intervals, which increases the network load. The unnecessarily high volume of the keepalive traffic reduces battery lifetime on mobile devices, especially those connected to mobile networks, where each message sent imposes additional overhead in the form of signaling traffic.

---

* Master study programme in field: Computer Engineering
Supervisor: Martin Nagy, Institute of Computer Systems and Networks, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 321–326.

This paper proposes a network architecture to deploy the Port Control Protocol (PCP) in the core of mobile networks and a method to measure the keepalive traffic reduction with PCP in mobile networks.

This paper is structured as follows. Section Related Work briefly reviews existing NAT traversal and keepalive reduction methods. Section Port Control Protocol describes the basics of the PCP protocol and the advantages of the deployment of PCP in software-defined networks. Section Architecture describes the proposed architecture of the core network and its components. Section Evaluation specifies methods to validate the proposed solution and to measure the keepalive traffic reduction in mobile networks with PCP.

2 Related Work

Protocols such as Session Traversal Utilities for NAT (STUN) [18], Traversal Using Relays around NAT (TURN) [7] or Interactive Connectivity Establishment (ICE) [15] can resolve NAT traversal issues for user applications. Additional methods for proper NAT traversal are defined for IPSec ESP [5] and mobile IP [6].

Source [3] proposes a method to reduce the keepalive traffic in mobile IPv4 networks and in IPSec communication by replacing UDP keepalives with the so-called TCP wake-up messages, given the considerably greater mapping lifetime for TCP connections on NAT and firewall devices from popular vendors [4]. The results of the experiments conducted suggest that the keepalive traffic reduction is significant in 2G (GSM) and 3G (WCDMA, HSDPA) networks, but not in IEEE 802.11 Wireless LAN [3].

3 Port Control Protocol

Port Control Protocol (PCP) [1] allows IPv4 and IPv6 hosts to determine or explicitly request network address mapping, port mapping and mapping lifetime directly from middleboxes. From this information, a host behind a middlebox can establish communication with a host in an external network or in another internal network behind another middlebox and can optimize the interval of sending keepalives. PCP does not replace the function of proxy or rendezvous servers to establish connections between hosts in different internal networks. PCP requires that hosts run a PCP client and middleboxes run a PCP server [1].

Based on the existing research [3], the reduction of the keepalive traffic in mobile networks can be considerable. PCP introduces a more universal approach that allows to optimize keepalive traffic for multiple transport protocols (any protocol with 16-bit port numbers) and other upper-layer protocols, such as ICMP or IPSec ESP [1].

PCP may be vulnerable to security attacks such as denial of service or mapping theft [1]. The security of PCP is currently under discussion [17]. an RFC draft specifies an authentication mechanism to control access to middleboxes [14].

3.1 Port Control Protocol in Software-Defined Networks

Software-defined networking (SDN) [8, 9, 11] is a novel approach to managing computer networks which separates the control and data planes of network devices to controllers and forwarders, respectively, and achieves greater network flexibility by allowing to program the network behavior. Existing networks are expected to migrate to SDN given the aforementioned advantages.

With SDN, a PCP server can run on a controller, thereby reducing the processing overhead on middleboxes, increasing vendor compatibility and avoiding the need to upgrade the middleboxes to support PCP server functionality. If multiple middleboxes are placed in an SDN network, mapping lifetime can be determined from the controller instead of every middlebox separately. There is

an ongoing effort to support advanced firewall functionality in SDN networks by introducing new PCP message types [13].

4 Network Architecture

This section describes the architecture of the proposed network, shown in the Figure 1. Only the essential components are shown in the core network of the architecture.

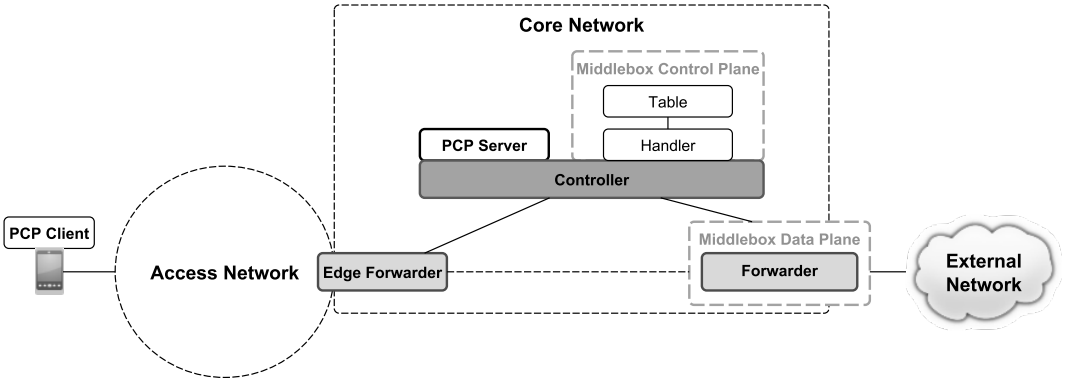


Figure 1. Architecture of the proposed network.

An end host running a user application with a PCP client is located behind an existing access network (such as GERAN in case of 2G networks). The access network connects to the core network via the edge forwarder. The edge forwarder forwards PCP requests to the controller, PCP responses from the controller back to the PCP client and other traffic further through the core network.

In the proposed architecture, the control and the data plane of a middlebox are decoupled. The middlebox data plane resides on another forwarder, placed between the core and the external network (the Internet). The middlebox data plane executes the rules installed by the control plane, such as overwriting IP addresses and transport protocol ports in packets in case of NAT.

The controller runs the middlebox control plane, which is responsible for maintaining mappings stored in a table. The handler accepts requests from the PCP server to create or remove a mapping and instructs the controller to add or remove the corresponding rules on the forwarder.

The PCP server running on the controller receives PCP requests, instructs the middlebox control plane to create a mapping for the client and sends PCP responses back to the client once the middlebox control plane successfully creates a mapping.

4.1 Details

For the implementation, OpenFlow [9, 10] is used as the communication protocol between the controller and the forwarders.

The edge forwarder contains two simple rules, called *flow entries* in the OpenFlow terminology. The first flow entry matches the PCP request – destination IP address of the PCP server and destination UDP port 5351. The second entry matches the PCP response – source IP address of the PCP server and source UDP port 5351. The PCP response is sent out the same access port the corresponding PCP request was received on[1].

[1] The PCP message forwarding is greatly simplified in order to focus on the more essential aspects of the design. Real networks may perform operations such as routing or load balancing.

The PCP server address is assumed to be the address of the default gateway, so PCP clients must use this address to communicate with the PCP server. Dynamic PCP server discovery options [1,2,12] are currently not considered.

In order to verify the proper traversal of packets behind middleboxes and the keepalive traffic reduction, a custom, simple NAT gateway is implemented in the network that supports only IPv4 addresses and TCP and UDP as the upper-layer protocols. a custom firewall, IPv6 or other upper-layer protocols are not implemented in the network, as the verification and evaluation method of the proposed solution is identical and would not affect the results.

The control plane of the NAT is responsible for creating NAT table entries from the configured pool of external IP addresses and ports. Each NAT table entry contains the following items: internal IP address, internal port, external address, external port, upper-layer protocol and mapping lifetime. The NAT data plane is represented as a set of flow tables and entries shown in the Figure 2.

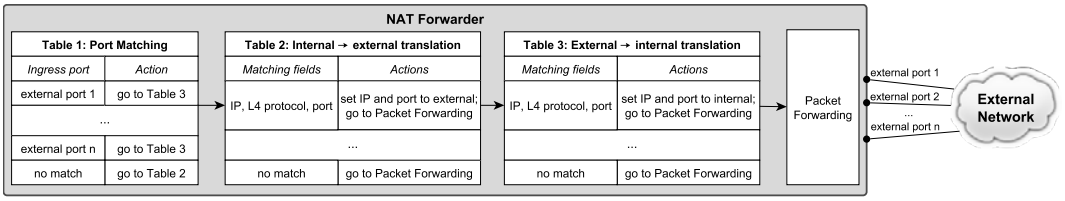


Figure 2. Flow entries in the NAT forwarder.

The design does not address the security of PCP. The authentication mechanism specified in source [14] could be used to control the access to the controller running the PCP server.

5 Evaluation

To verify the implementation of the network, a test topology shown in the Figure 3 is created in a virtual environment, using the *UnifyCore* project[2] [9,16] as the underlying core network infrastructure. For the implementation, software compliant with OpenFlow 1.3 is used -– *ryu controller*[3] for the controller and *ofsoftswitch*[4] for the forwarder.

The edge and the NAT forwarders are merged into one forwarder for the sake of simplicity. The *Packet Forwarding* flow table, as seen in the Figure 2, represents a set of flow entries to forward packets, e.g. based on their MAC or IP addresses. For testing purposes, the table contains two flow entries – one to forward packets from the access port out the external port and another one to forward packets from the external port out the access port.

From the host computer, the user generates a message containing PCP request fields, using the web application from the PCP Testing Tool[5]. *Host 1* runs a PCP client, as the second part of the *PCP Testing Tool*, which receives the message sent from the web application and generates a PCP request. The PCP request is forwarded to the PCP server running on the controller, which instructs the NAT control plane to create a NAT table entry and install corresponding flow entries on the forwarder. Upon successful flow entry installation, the PCP server sends a PCP response back to *Host 1*.

To test the communication through NAT, *Host 2* listens to incoming UDP connections and *Host 1* sends a UDP datagram, e.g. using the Unix `nc` command. On both hosts, `tcpdump` running on both hosts may be used to verify that the source IP address and port were properly translated when

[2] http://www.unifycore.com/

[3] https://github.com/osrg/ryu

[4] https://github.com/CPqD/ofsoftswitch13

[5] http://sourceforge.net/projects/pcptestingsuits/

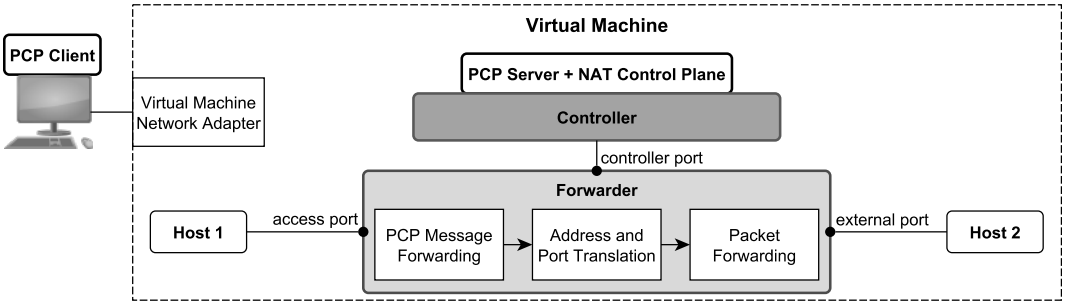


Figure 3. Test topology of the network.

the datagrams are received on *Host 2*. The same experiment from *Host 2* to *Host 1* may be conducted to verify external-to-internal address and port translation.

The same test topology is used to measure the keepalive traffic reduction. The measurement method is specified as follows:

1. Generate network traffic from *Host 1*. From the network traffic, determine the number of bytes of the entire network traffic and the number of bytes comprising the keepalive messages. Perform this step without PCP enabled.

2. Perform step 1 with PCP enabled with the same network traffic (sans the keepalive traffic, which will differ).

3. Compute the percentual reduction of the keepalive traffic: $(k_1/k_2) * 100\ \%$, where k_1 and k_2 are the number of bytes of the keepalive traffic measured in the first and the second test, respectively.

The network traffic will be generated with a network traffic generator that must be programmable in order to integrate a PCP client. The network traffic is assumed to be Ethernet traffic. In order to evaluate the reduction in a specific mobile network, each message sent to the network will be associated with the corresponding signaling traffic generated per each message.

The tests will be conducted with different mapping lifetime values and transport protocols (TCP and UDP). The results can then be interperted as a graph showing the dependency of the keepalive traffic reduction from the mapping lifetime.

The proposed measurement methods may be revised as further research needs to be conducted to determine the background and signaling traffic in chosen mobile networks.

6 Conclusion

This paper described the deployment of the Port Control Protocol in software-defined networks. The described architecture separates the control and the data plane of middleboxes and allows to run the PCP server outside the middleboxes. This improves vendor device compatibility and avoids the processing overhead imposed by running the PCP server. With PCP deployed in the network, mobile devices connected to mobile networks can reduce the amount of keepalive traffic sent, which results in extended battery life of mobile devices. Additionally, the network throughput is increased due to signaling traffic reduction in access networks.

Acknowledgement: This project was partially supported by the Slovak national research grant 1/0676/12 and by the Tatra banka Foundation under the contract No. 2012et011.

References

[1] Boucadair, M., et al.: Port Control Protocol (PCP). RFC 6887, 2013.

[2] Boucadair, M., et al.: DHCP Options for the Port Control Protocol (PCP). RFC 7291, 2014.

[3] Eronen, P.: TCP Wake-Up: Reducing Keep-Alive Traffic in Mobile IPv4 and IPsec NAT Traversal. *Nokia Research Center*, 2008.

[4] Haverinen, H., et al.: Energy Consumption of Always-On Applications in WCDMA Networks. In: *Vehicular Technology Conference, 2007. VTC2007-Spring. IEEE 65th*, 2007, pp. 964–968.

[5] Huttunen, A., et al.: UDP Encapsulation of IPsec ESP Packets. RFC 3948, 2005.

[6] Levkowetz, H., Vaarala, S.: Mobile IP Traversal of Network Address Translation (NAT) Devices. RFC 3519, 2003.

[7] Matthews, P., et al.: Traversal Using Relays around NAT (TURN): Relay Extensions to Session Traversal Utilities for NAT (STUN). RFC 5766, 2010.

[8] Nadeau, T.D., Gray, K.: *SDN: Software Defined Networks*. O'Reilly Media, 2013.

[9] Nagy, M., Kotuliak, I.: Utilizing OpenFlow, SDN and NFV in GPRS Core Network. In Leung, V.C.M., Chen, M., Wan, J., Zhang, Y., eds.: *Testbeds and Research Infrastructure: Development of Networks and Communities*. Number 137 in Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering. Springer International Publishing, 2014, pp. 184–193.

[10] Open Networking Foundation: OpenFlow Switch Specification, version 1.3.0. `https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-spec-v1.3.0.pdf`, [Online; accessed December 5th, 2014].

[11] Open Networking Foundation: Software-Defined Networking: The New Norm for Networks. `https://www.opennetworking.org/images/stories/downloads/sdn-resources/white-papers/wp-sdn-newnorm.pdf`, [Online; accessed November 30th, 2014].

[12] Penno, R., et al.: PCP Anycast Address. Internet-Draft draft-ietf-pcp-anycast-02, 2014.

[13] Reddy, T., et al.: PCP Firewall Control in Managed Networks. Internet-Draft draft-reddy-pcp-sdn-firewall-00, 2014.

[14] Reddy, T., et al.: Port Control Protocol (PCP) Authentication Mechanism. Internet-Draft draft-ietf-pcp-authentication-07, 2014.

[15] Rosenberg, J.: Interactive Connectivity Establishment (ICE): A Methodology for Network Address Translator (NAT) Traversal for Offer/Answer Protocols. RFC 5245, 2010.

[16] Skalný, J. et al.: Application of Software Defined Networking (SDN) in GPRS Network. In: *IIT.SRC 2014: Student research conference in Informatics and Information Technologies*, Bratislava, Nakladateľstvo STU, 2014, pp. 553–558.

[17] Wing, D.: Port Control Protocol. *The Internet Protocol Journal*, vol. 14.

[18] Wing, D., et al.: Session Traversal Utilities for NAT (STUN). RFC 5389, 2008.

Hardware Task Scheduling in Real-Time Systems

Lukáš KOHÚTKA*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
kohutka555@gmail.com

Abstract. Real-time systems often handle a mixture of various types of tasks, not only the real-time tasks. This is a motivation to design a real-time operating system that can handle as many tasks of various types as possible using a robust but fast dynamic scheduling algorithm. Software implementations of task schedulers are either robust and slow or fast and simple. If we want to have a scheduler that is both robust and fast, we need to reduce the time complexity of the scheduling algorithm by using hardware acceleration. In this paper, we propose hardware representation of a robust scheduling algorithm that allows us to schedule any combinations of hard real-time, firm real-time and non real-time tasks in a constant time. Thanks to the constant time complexity, the hardware scheduler is both robust and fast, which leads to better utilisation of a processor.

1 Introduction

Real-time systems represent a category of embedded systems that process real-time tasks, where success of the real-time task depends not only on the result of the computation but depends on the timing of the task too. The real-time task that returns a correct result of the computation but does not meet the deadline may have the same effect as incorrect result of the task. There are several types of tasks that may occur in real-time systems: hard real-time, soft real-time, firm real-time and non real-time tasks. Hard real-time tasks are the most important tasks because a single hard real-time task that does not meet the deadline is considered as a failure of the whole real-time system. Firm real-time tasks are like the hard real-time tasks but occasional missing of the deadline of firm real-time tasks reduces quality of the system but the system does not fail. Soft real-time tasks have two deadlines in comparison to the firm real-time tasks, where completion of the task between these two deadlines leads to worse quality of the result of the task. Non real-time tasks are tasks that do not have any deadline. Even if we use a microcontroller with the highest performance, there is no guarantee that all tasks meet their deadlines. Therefore we use a real-time task scheduler that represents a core of real-time operating systems that handles this problem. The ideal real-time task scheduler creates an optimal ordering of the tasks, where the tasks will be computed and completed on time. The ideal scheduling has no overhead on CPU. The more we

* Master degree study programme in field: Computer Engineering
Supervisor: Martin Vojtko, Institute of Computer Systems and Networks, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 327–332.

use CPU for the scheduling, the less we use the CPU for effective computation and completion of the scheduled tasks.

This paper presents a novel real-time scheduling algorithm suitable for hardware implementations and hardware representation of the new real-time scheduling algorithm. The algorithm is called Flexible Earliest Deadline and is based on the Robust Earliest Deadline. Simulation and implementation results are shown in order to verify the correctness of the designed scheduler. We present a comparison of hardware and software implementation in terms of performance and efficiency of the scheduler.

2 Related work

Most of the schedulers are implemented as a software module of an operating system. There exist static and dynamic schedulers. Static schedulers generate task order during the system initialization and dynamic schedulers generate task order during the system execution. The problem of the software implementations is that the task scheduling cannot be both fast and robust. The simple schedulers are fast but the possibility that the task set will be schedulable is smaller. Most of the software real-time operating systems use the simple scheduling algorithms.

Besides existing software implementations, there exist hardware implementations too. We found that most of these solutions use the EDF scheduling algorithm, which has very good results but the problem is that this algorithm supports only hard real-time tasks. [1-6]

We believe that most of the real-time systems contain heterogeneous task sets so there is a need for an algorithm that can handle the heterogeneous task sets reliably and efficiently.

3 FED – Flexible Earliest Deadline

If we want to support a combination of various types of tasks in our real-time system, we have to use a more robust scheduling algorithm. The second reason to use a more robust algorithm is to predict a better order of tasks to complete.

We propose a novel scheduling algorithm called Flexible Earliest Deadline that is very similar to the existing algorithm Robust Earliest Deadline [7]. In fact, the FED algorithm can be viewed as a special case of the RED algorithm, where one new constraint is added:

$$primary\ deadline = secondary\ deadline \tag{1}$$

Thanks to the fact that the FED is a subset of RED, we can be sure that the FED algorithm works properly because the RED algorithm does.

The added constraint brings us the following advantages in comparison to the RED algorithm. While the RED algorithm supports soft real-time tasks only, the FED algorithm supports hard real-time tasks, firm real-time tasks, non real-time tasks and any combinations of these task types.

The RED algorithm contains a dividing operation, which is used for computation of effectiveness *e* of soft real-time task completion. The computation formula is following:

$$e = \frac{secondary\ deadline - current\ time}{secondary\ deadline - primary\ deadline} \tag{2}$$

However, the FED algorithm has no dividing because if we add the constraint from formula čdividing operation anymore. No dividing leads to lower complexity and area cost of the hardware implementation of the scheduling algorithm.

$$e = \frac{deadline - current\ time}{deadline - deadline} = \frac{deadline - current\ time}{0} = \infty \tag{3}$$

While the RED algorithm does not define the policy of the task sacrificing in order to reduce the overload of the whole system, the FED algorithm describes a recommended method for selection

of such tasks. This method selects the tasks according to 2 attributes, primary and secondary. The primary attribute is a type and priority of the task, where the task with the highest priority number is the best candidate to be sacrificed. The secondary attribute is considered in case that there are multiple tasks with the highest priority. This attribute represents a contribution of the task to overload of the system. Task with the highest primary attribute and then the highest secondary attribute should be chosen to be sacrificed. The term "sacrifice" represents actually a removing of the task from the queue of the scheduled tasks and adding the task to the list of rejected tasks, called rejection list. The list of rejected tasks exists in a hope that some of the tasks being executed will be completed sooner than expected and the processor will be able to complete some of the rejected tasks before the deadline.

4 Hardware architecture

The hardware real-time task scheduler is designed in a form of a coprocessor unit that cooperates with the CPU and thus significantly accelerates the real-time task scheduling. In order to implement a coprocessor unit, we have chosen a soft core processor Nios II. Such architecture represents actually a system on chip (SoC). The designed architecture is platform independent and thus it is applicable in all FPGA and ASIC implementations.

4.1 Functions – custom instructions

The behaviour of the hardware scheduler is described by functions that represent custom instructions of the coprocessor too. These functions are:

- *task_schedule* – creates and adds a new task,
- *task_delete* – removes an existing task,
- *task_sleep* – suspends the task execution due to I/O operations or inter task dependencies,
- *task_awake* – sets to continue a sleeping task.

The task_schedule requires these inputs: ID, type, deadline and worst-case execution time of the task. These inputs are passed to the coprocessor by two universal 32-bit registers. The output of the task_schedule is the ID of the task that should be completed as the first one. The task_delete requires only one input that is the ID of the task that has to be deleted. The task_sleep and task_awake functions require the ID of the desired task as an input.

4.2 System level

In the Figure 1 the architecture of the hardware scheduler is described at the highest level of abstraction (the system level). The control unit decodes the instruction code of the custom instructions and controls all other components of the scheduler according to the decoded instruction code and input data fetched from two 32-bit universal registers. The registers contain these data: ID, type, slack-time and worst-case execution time of the task. Size of the ID is 12 and thus the maximum possible task capacity is 4096. The slack-time and worst-case execution time are both 20-bit numbers. The rest 12 bits are left for the type of tasks and possibility to upgrade the scheduler by adding new features. We do not need to receive the deadline of the tasks because we can simply get the deadline as a sum of the slack-time and worst-case execution time of the task. The scheduled queue contains scheduled tasks that are ready for execution. The acceptation list contains accepted tasks that do not cause overload of the system. The accepted tasks are tasks that are either scheduled or sleeping. The rejection list contains tasks that are rejected ("sacrificed") because of overload of the system. Each task is either accepted or rejected. Each accepted task is either scheduled or sleeping. The bank of all tasks stores important information about all tasks in the system (e.g. task type, worst-case execution time).

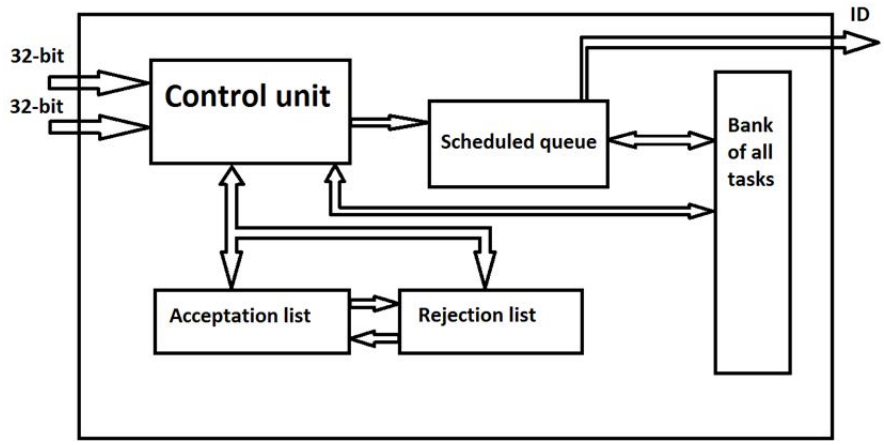


Figure 1. System level architecture of the hardware scheduler.

4.3 Sorting queue

The major part of the hardware task scheduler is represented by structures known as sorting queue. The sorting queue contains tasks that are ordered by some attribute. For example in case of the scheduled queue, this attribute is a deadline of the task.

The scalability of the sorting queue in terms of area cost and performance is extremely important for the scalability of the whole task scheduler because while the control unit has constant size, the size of the sorting queues is highly dependent on the maximum number of the tasks the system can handle. The more tasks we want to support, the bigger structure of each sorting queue is needed.

In order to achieve the best possible scalability of the task scheduler, we propose hybrid architecture of the sorting queue described at register-transfer level (RTL). The proposed hybrid architecture combines two well-known architectures: shift registers and systolic array. The whole idea is quite simple. We use one bigger systolic array that contains multiple smaller shift registers in each cell. This approach gives us constant complexity of the critical path in the circuit, which leads to the frequency of the system that does not depend on the maximum number of tasks. The second advantage (in comparison to simple systolic array approach) is the fact that the area cost of this architecture is much lower. Assuming that N is the maximum number of tasks in the scheduler:

- The area cost complexity of the standard shift registers is N.
- The area cost complexity of the standard systolic array is $2N$.
- The area cost complexity of the proposed architecture is $N + \frac{N}{K}$, where K is the number of shift registers in one cell of the systolic array. We may conclude from the above equations that the hybrid architecture can maintain the same performance as the systolic array architecture but the area cost is much smaller.
- The structure of one task cell is displayed in Figure 2.

As we can see, the task cell operates in two modes: task_add = 1 or task_add = 0. If the task_add is equal to 1, the new task may be inserted to the task cell or the task coming from the right side may be added to the task cell according to the deadline values. If the task_add is equal to 0, the existing task stored in the task cell may be deleted by writing data from the left side to the task cell. The left_deadline and left_task_id are the data coming from the left side. The right_deadline and right_task_id are the data coming from the right side. The input_deadline and input_task_id are the data coming from the input and thus these data represent the new task.

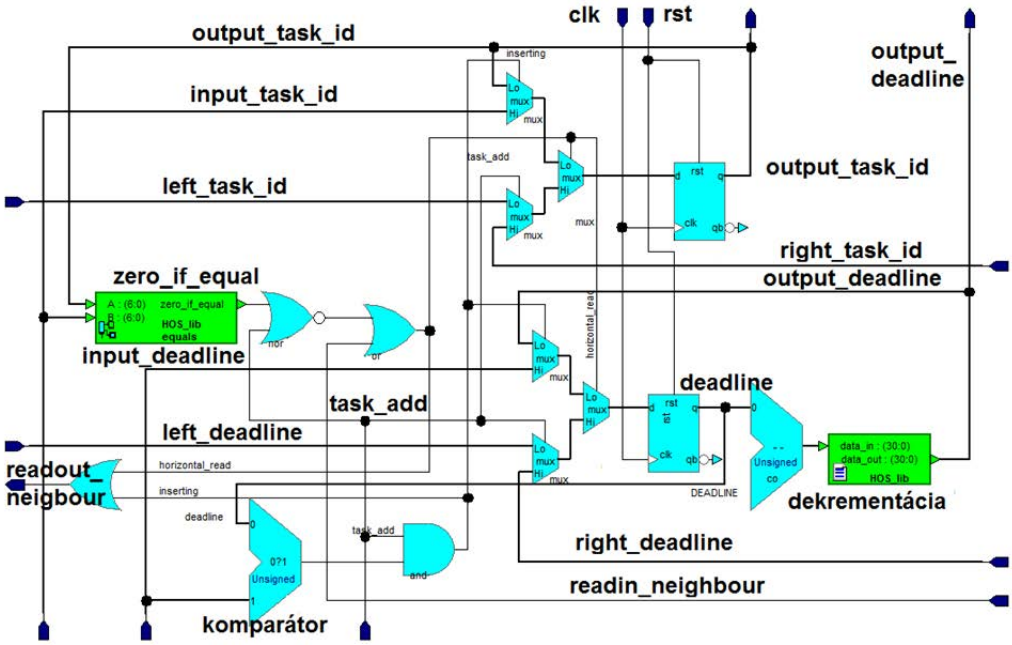


Figure 2. Structure of the task cell.

5 Simulation results

Functionality of every designed circuit was verified by simulations in ModelSim. All of the circuits work as intended. Figure 5 shows one example of such simulations, which verifies functionality of the hybrid architecture described in chapter 4.

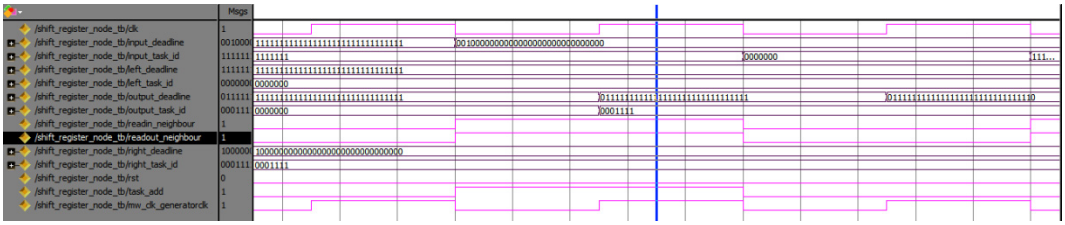


Figure 5. Simulation result example.

6 Conclusion

A novel efficient scheduling algorithm suitable for complex real-time systems was proposed and its correctness was analytically proofed. A hardware architecture that uses the proposed scheduling algorithm was designed. A scalable hardware structure of sorting queue was designed and verified. Further work is planned to implement the designed hardware architecture on FPGA platform. The implemented hardware task scheduler will be compared to the software equivalent of the same scheduling algorithm. The main benefits of this research are the increased predictability, determinism and performance of the real-time systems, which leads to higher number of real-time tasks completed before their deadlines are met. The final effect is an improved quality and reliability of the real-time systems.

Acknowledgement: This work was partially supported by the Slovak Science Grant Agency (VEGA 1/1008/12 “Optimization of low-power design of digital and mixed integrated systems”).

References

[1] Bloom, G., Parmer, G., Narahari, B., Simha, R.: Real-Time Scheduling with Hardware Data Structures, 2010.

[2] Ferreira, C., Oliveira, A. S. R.: Hardware Co-Processor for the OReK Real-Time Executive, 2010.

[3] Chandra, R., Sinnen O.: Improving Application Performance with Hardware Data Structures, 2010, ISSN 11783680.

[4] Ong, S. E., Lee S. C.: Seos: Hardware Implementation of Real-Time Operating System for Adaptability, 2013, ISBN 9781479927951.

[5] Ferreira, C., Oliveira, A. S. R.: Hardware Co-Processor for the OReK Real-Time Executive, 2010.

[6] Kohout, P.: Hardware Support for Real-Time Operating Systems, 2002.

[7] Buttazzo, G. C.: Hard Real-Time Computing Systems: Predictable Scheduling Algorithms and Applications, 2011, ISBN 9781461406754, 9781461406761.

Optimization of Decision Diagrams

Ján LÚČANSKÝ*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
janocansky@gmail.com

Abstract. The primary aim of this paper is to present a new method of reduction for Decision Diagrams (DD) and their free form (Free DD – FDD) [11]. DDs as dynamic structures are extensively used in computer science due to their ability to effectively represent any Boolean function. Optimizing DD can have significant impact on the size, speed or power consumption of circuit represented by DD. The new approach examines the use of residual variable in functional DD and its impact on reduction of number of nodes. Using residual variable in Binary DDs and their Reduced and Ordered version (ROBDD) can lead up to 25,53 % less nodes compared to the original ROBDD [8]. During the time this paper was being written, no papers using residual variable in functional DD has been found. This new approach examines the impact of residual variable used in form of DD called Kronecker Functional Decision Diagram (KFDD).

1 Introduction

In computer science, the primary aim is to create new, more efficient data structures and algorithms that are used in hardware (HW) design. More efficient data structures can lead to better results for already solved problems or have significant impact on solving unsolved problems. Also, as some of the technologies are reaching their fundamental physical limits, it is becoming more important to apply optimization on the design level. In HW design, most of logical circuits can be represented by various data structures. One of the popular data structures used is called Decision Diagrams (DD) [1] or Binary DD (BDD). BDD is a data structure that is used to represent any Boolean function, however is mostly used to represent control functions. In DDs, Boolean function is represented as rooted, directed, acyclic graph with one root node and two terminal nodes, one for logical *0* and one for logical *1*. For the purpose of this paper, each DD considered can have only two child nodes.

Optimizing DDs can lead to various forms of optimizations for their corresponding use [11]. Among many other, DDs and mainly their Reduced and Ordered form (e.g. ROBDD) [11] are often used in functional test generation, data compression, memory representation and other. By reducing the number of nodes in DD, equivalent or close to equivalent reduction can be applied on circuit represented by DD. Smaller circuits often have smaller static power consumption and can perform faster with identical results as their unoptimized form [7].

* Bachelor degree study programme in field: Computer Engineering
Supervisor: Dr. Peter Pištek, Institute of Computer Systems and Networks, Faculty of Informatics and Information Technologies STU in Bratislava

There are many forms of DDs. According to [2], only two types of decomposition have impact on node reduction in DDs – Shannon decomposition and Reed-Muller decomposition. Based on the decomposition used, special form of DD is achieved. Using only Shannon decomposition we get BDD, on the other hand, using only Reed-Muller decomposition, Functional DD [5] can be constructed. Combining these two approaches a special type of DD known as KFDD [5] can be constructed. However, each of this types uses the same order of variables in each level of DD, thus in order to achieve maximum reduction, another form of DD is used – Free DD (FDD) [12]. In FDD, each subtree in diagram (i.e. child of root node is root node of separate diagram) can represent different order of variables. Changing order of variables in each subtree can lead to better predisposition for decomposition methods and can therefore result in higher level of reduction of number of nodes. This paper focuses on optimizing DDs from the point of node count and proposes a new algorithm for DD reduction.

2 Preliminaries

DD optimization can be divided into two categories – ordering and reduction.

Ordering respects a given order of variables for each binary vector and results in ordered version of DD. Order of variables can have great impact on the level of reduction of nodes' number achieved in DD. Two types of ordering are considered in this paper – order of variables and order of decompositions used for each variable.

Reduction decreases the number of nodes in DD and results in reduced version of DD. Types of reductions used in reduction process are shown in Figure 1. Reduction type *I* can be applied any node regardless of the decomposition type used. Reduction type *S* can only be used on nodes with Shannon decomposition. Reduction type *D* can only be used on nodes with Reed-Muller (sometimes called Davio) decomposition [9].

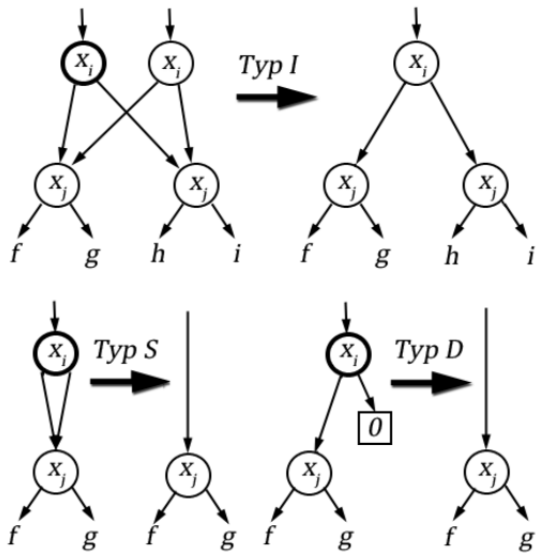


Figure 1. Reduction methods used in Reduced Ordered DD (RODD) [9].

Each RODD has these two attributes:

- Uniqueness – no two nodes have the same left and right child and represent the same variable,

- Non-redundancy
  - No two nodes representing the same variable have identical children (*Typ I* in Figure 1),
  - Shannon – no node has identical left and right child (*Typ S* in Figure 1),
  - Reed-Muller – no node has the right child equal to terminal node with logical *0* (*Typ D* in Figure 1).

3 Genetic algorithm

DD ordering is a complex problem. With *n* being the number of input variables, total number of all possible variable orders is equal to factorial of *n* and the total number of all possible orders of decompositions (i.e. Decomposition Type List – DTL) is equal to 3^n. By combining these two method together, we get a $n!.3^n$ complexity without using FDD (remember each node can have different order of variables for each following child). Using FDD increases this complexity even further and for each child we get $(n-1)!.3^{n-1}$ complexity. Therefore, genetic algorithm (GA) presented in [6] and Modified Memetic Algorithm (MMA) presented in [10] is used with order of variables and DTL to lower the complexity of DD ordering.

The core term of a GA is called population (generation) – in this case order of variables and/or DTL (each variable has one decomposition, i.e. the length of order of variables is equal to the length of order of decompositions). Each population consists of chromosomes (one particular variable order) and each chromosome consists of genes (variables or variable indices).

In each step of GA, genetic operators are applied on each chromosome. Two operators used are:

- Crossover – combine attributes from multiple (two) parents (shown in Table 1) [6],
- Mutation – Swap genes of particular individual (chromosome) in generation (applied on each chromosome separately) (shown in Table 2) [6].

By combining these two operators on generation (parents), new population is achieved (successors). Mutation and crossover are applied on selected chromosomes with predefined probability. To ensure best results, technique called *Elitism* [8] is applied – subsequent generations will never provide less fitting results than previous generation.

Table 1 shows example of Crossover for DTL. DTL_{11}toDTL_{16} represents previous generation of DTLs. DTL_{11} to $DTL_{15}$are selected as parents for new genereation and based on thei values, Successor A and Successor B is achieved using Crossoves operator.

Table 2 shows example of Mutation operator. Single parent from previous generation is *mutated* and new Successor is achieved. Values *S*, *P* and *N* in both tables represent form of decomposition(S – Shannon, P – positive Davio, N – negative Davio).

Table 1. Crossover for DTL.

		P_1	P_2	P_3	P_4	P_5	P_6	
Previous generation	DTL_{11}	S	S	S	S	S	S	Parent A
	DTL_{12}	P	N	P	N	P	N	
	DTL_{13}	S	P	N	S	P	N	
	DTL_{14}	S	S	P	P	N	N	
	DTL_{15}	P	P	P	N	N	N	Parent B
	DTL_{16}	S	N	N	P	S	N	
Subsequent generation	DTL_{21}	S	S	P	N	S	S	Successor A
	DTL_{22}	P	P	S	S	N	N	Successor B

Table 2. Mutation for DTL.

	P_1	P_2	P_3	P_4	P_5	P_6
Parent	S	S	P	N	S	S
Successor	S	P	S	N	S	S

4 Residual variable

In this paper we combined FDD and KFDD and applied new method of reduction called residual variable (RV). RV was previously used only in BDD and lead up to 25% less nodes compared to the original ROBDD [9]. Use of RV in KFDD has no beed previously tested, therefore we design this method as new approach in KFDD optimization.

Function (represented by Binary vector) decomposition f for $x_i = c$ $(c \in \{0,1\})$and for order of variables $x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots x_n$ is based on following function [8]:

$$f|_{x_i=c}(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots x_n) := f(x_1, \dots, x_{i-1}, c, x_{i+1}, \dots x_n) \quad (1)$$

In DD construction for *n* variables, Shannon or Reed-Muller decompositions were applied until the level defined by following function has been reached [8]:

$$f|_{x_n=c}(x_1, \dots, x_n) := f((vector\ \ n-1), c) \quad (2)$$

If the decomposition is stopped an iteration earlier, function (3) can be obtained:

$$f|_{x_{n-1}=c}(x_1, \dots, x_{n-1}, x_n) := f((vector\ n-2), c, x_n) \quad (3)$$

where $c \in \{0,1\}$ and (*vector n-1*) contains corresponding logical values of*0*and *1*(based on order of variables$x_1, \dots x_{n-1}$ in upper levels of DD) [9]. Using this approach we get residual functions for x_i shown in Table 3. Figure 2 shows transformation of common DD node to residual variable.

Table 3. Residual functions for canonical matrix.

Left child	Right child	Generated residual function
0	0	0
0	0	x_i
1	1	$\overline{x_i}$
1	1	1

5 Proposed algorithm

New approach consists of existing methods and algorithms combined together and applied on FDD. Scientific value of this paper lies in examining the benefits of residual variable in FDD. Up to the publication date of this paper, no other works implementing this method were found.

In the first step, GA is applied on input order of variables and DTL. Changing the input order of variables results in different input Binary vector (BV), therefore for each subsequent generation, new BV must be computed by reordering the original input BV according to the new order of variables. Maximum number of generated populations is stated beforehand (e.g. maximum of 300 populations will be generated).

In the second step, we verify the input BV and fill don't care bits, usually marked with letter *X* in vector (note that bit with logical value *0* is marked with 0 and logical value *1* is marked with 1). In order to achieve the highest level of reduction, don't care bits of BV are filled with value based on the value of their immediate neighbor. Firstly, BDD is constructed from the input BV using only Shannon decomposition. Nodes with value corresponding to don't care bit are then

replaced with the value of their neighbor, preferably of the neighbor sharing the same parent node (if the neighbor has a value of don't care bit, the next node with value other than value *X* is used).

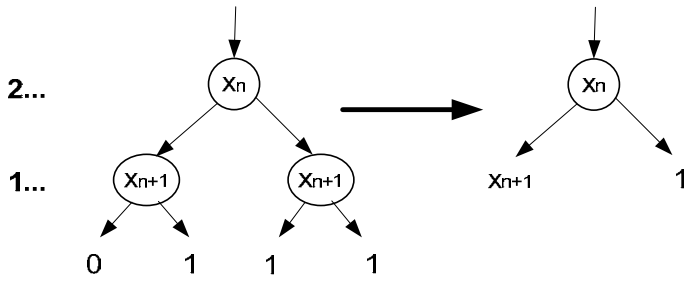


Figure 2. Transformation of common DD node to node with residual variable [9].

In the next step, complete, unreduced Ordered KFDD is constructed. Reduction is applied from bottom to top, i.e. from terminal nodes to root node. RV is then applied on OKFDD, removing all terminal nodes. After that, reduction methods are applied, resulting in Reduced Ordered KFDD (ROKFDD).

In the final step, logic of FDD is applied. Basically all previous steps are repeated, from GA (BV verification can be skipped, previous values of don't care bits can be reused) to RV. The difference is only in the node that is considered to be the root. Each of the previous steps is applied on the next node from the top to the bottom until last but one variable. At the beginning, GA, BV verification and reduction (including RV) are applied on the root node of the entire DD (first variable in input order of variables). After computing ROFDD (Reduced Ordered FDD) for the selected order, all steps are applied on left subtree (and subsequently on right subtree), i.e. left child of root node is root node of new diagram (diagram without the original root). By recursively applying these steps, each subtree of the original DD can have different order of variables.

6 Experimental results

Up to this day, implementation has not been fully completed yet. The proposed algorithm will be tested on benchmark sets like LGSynth93 [4] with focus on the impact of RV. First run of tests will be done using only ROKFDD logic with GA. This will be compared to ROKFDD with RV, FDD without RV and FDD with RV. Each of the tests will then be compared to each other by percentage of reduction compared to the original ROKDD without RV. Results will later be compared to recent measurements done in the same area, for example results shown in [8] or [9].

Some other areas of testing will be considered, for example all previously mentioned tests can be influenced by the maximum number of populations in GA. Changing this value can have significant impact on time of computation as well as the level of reduction achieved. On the other hand, all tests can be compared to results achieved by computing all possible variations ($(n-1)!.3^{n-1}$complexity) with the same criteria.

7 Conclusion

A novel algorithm for DD optimization is proposed. Since the implementation has not been fully completed so far, no testing was conducted. Nevertheless, there are a few positive expectations, mainly in the level of reduction. Using FDD logic, the most optimal results can be achieved with chosen decomposition methods in comparison to logic used in BDD or KFDD. Also, using residual variable in FDD is expected to reduce the number of nodes even further, as

this method has not been tested so far as well as residual variable in KFDD. Examining the impact of residual variable used in KFDD and FDD is the main asset of this paper.

One of the downsides of the proposed methods is that the time of computation of final ROFDD can be significantly higher in comparison to ROKFDD or ROBDD. In this case, GA is applied on each of the children of the root node, thus creating computations for a new DD with *n-1* variables.

Results achieved by the form of testing described in previous chapter can help identify the type of function, as BDDs are more suitable to represent control functions and functional DDs are suitable for data functions [11]. Based on this, future researchers can appropriately chose benchmark sets for their design.

References

[1] Akers, S. B.: Binary Decision Diagrams. In: *IEEE Transaction on Computers*, IEEE, 1978, *C-27*, no. 6, pp. 509–516.

[2] Becker B., Drechsler, R.: How many decomposition types do we need?. In: *Proc. Eur. Design & Test Conf.*, (1995), pp. 438–44.

[3] Henessy, J., Patterson, D.: *Computer Organization and Design: The Hardware/ Software Interface.* Morgan Kaufmann Publishers, San Mateo California, (1994).

[4] BenchmarkCircuits: *LGSynth93*, [Online; accessed March 24, 2015]. Available at: http://www.cbl.ncsu.edu:16080/benchmarks/LGSynth93/, (1993).

[5] Drechsler, R., Sarabi, A, Theobald M., Becker B., Perkowski M.: Efficient Representation and Manipulation of Switching Functions Based on OKFDD. In: *IEEE Transactions on Computer-Aided design of integrated circuits and systems*, (1998), Vol. 17, No. 10.

[6] Drechsler, R., Jung, M., Lee, G., Park, S.: *Minimization of OPKFDDs Using Genetic Algorithms*. Technical Report, Chonnam Natl University, Kwangju, Korea, IEEE, (2001).

[7] Lindgren, P., Kerttu, M., Thornton, M., Drechsler, D..: Low Power Optimization Technique for BDD Mapped Circuits, In: *Design automation conference*, IEEE, (2001).

[8] Maruniak, M.: Metódy Optimalizácie Multiplexorových stromov. 2014, Master's thesis, (2014), pp. 46–56.

[9] Pištek, P.: *Nové metódy syntézy logických obvodov s multiplexormi.* PhD. thesis, Slovak University of Technology in Bratislava, (2010), pp. 6–46.

[10] Sanisha, R, Bansal, M.: Performance Comparison among Different Evolutionary Algorithms in terms of Node Count Reduction in BDDs. In*: International Journal of VLSI and Embedded Systems-IJVES*, (2013), pp. 491–496.

[11] Towhidi, F., Lashkari, A. H., Hosseini, R. S.: Binary Decision Diagram*.* In: *International Conference on Future Computer and Communication*, (2009).

[12] Gergov J., Meinel C.: Efiicient Boolean manipulation with OBDD's can be extended to FBDD's. In: *IEEE Transaction on Computing*. (1994), vol. 43, no. 10, pp. 1197–1209.

Real-Time Performance Comparison of FreeRTOS and User-Written Scheduler

Michal VALIČEK*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
michal@valicek.sk

Abstract. Real-time systems are depending on deterministic timing and effective use of system resources. So there is always hard decision whether to use real-time operating system – RTOS or to compose light-weight scheduler, especially in case of resource-restricted architectures. In this paper, we are bringing a simple method and metrics to determine timing performance of RT system. The method is used to compare simple scheduler to FreeRTOS, which is one the most popular RT operating systems. All tests are performed by external systems, so there is no device or system-specific routine or overhead to target device.

1 Introduction

Hard real-time implementations are depending on deterministic way of task scheduling and handling [2]. There are plenty of certified, even free to use systems to fulfill these needs. One of them is FreeRTOS, which is the most popular system and its popularity is increasing rapidly every year. It is ported on complex and also low-end platforms [3]. On the other hand even basic kernel – supporting set of periodic and aperiodic tasks plus preemption – is utilizing significant part of program memory and computation time. That is the main reason for programmers to think about other, more memory and computational effective ways of task handling. In some implementations, set of real-time tasks can be very simple and doesn't always require advanced OS features like semaphores, separate queues, mutexes etc.

In case of not using out-of-box system, it is crucial to test real-time aspects of the system [2]. Referring to this, we have developed such a method and presented it on Advanced Virtual RISC (AVR) microcontroller in a way to help developers in decision to choose right alternative [6]. The method is system and architecture-independent so it's possible to use it in any other analysis.

The presented method is inspired by paper published in 2002. Chris Tacke and Lawrence Ricci benchmarked real-time determinism in Microsoft Windows CE on the industry standard strongARM platform. The point was to inject stimuli to interrupt input of driving microcontroller (sleep button of handheld) and measure delay of dependent task. The task concurrency was simulated by running various numbers of polygon calculating program instances in background.

* Master degree study programme in field: Computer Engineering
Supervisor: Assoc. Professor Tibor Krajčovič, Institute of Computer Systems and Networks, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 339–344.

We are appreciating the simplicity of the method and need of only basic laboratory equipment – signal generator, counters and oscilloscope – to perform tests [7]. On the other hand only one scenario has been taken into account and results have not been evaluated in any standardized way.

The paper is split into sections. Second section is dealing with testing method, chosen operating systems and used metrics. The measured data are shown in third section and finally in last, fourth section we are bringing a summary and further work.

2 Testing method

The way of testing is straight-forward and as simple as possible. We are generating and injecting aperiodic stimuli in any time-slice of scheduler cycle to microcontroller and measuring the time of delay in environment of other concurrent tasks. We are dealing with 3 scenarios:

- Reschedule between T1 and T2, having the same priority,
- task T2 state is changed to ready during execution of task T1, time of pre-emption is measured,
- the aperiodic task driven by external stimuli is set as the highest priority so the ability of preemption and rescheduling is measured (time of ISR handling and twice the time of rescheduling).

The values are read and evaluated by an oscilloscope. Minimal and maximal delays are measured in persistent mode. The delay distribution in time is evaluated in average mode (when significant jitter occurs).

2.1 RIOS

We chose a simple task scheduler, developed at the Irvine University by Frank Vahid and Bailer Miller [8]. It's clearly written and easy to modify in order to add other functionalities. The scheduler is written to support set of independent periodic tasks and schedule them according to their static, user-defined priorities. We have added the support of aperiodic event handling, debugging mode and sporadic server (prevents starving of low priority tasks). The whole code is 213 lines long (in comparison to 117 lines of original code). The ability of the modified system was tested and it meets criteria required for RTOS even in the worst case scenario [9].

2.2 FreeRTOS

FreeRTOS is a market leading real-time kernel, which is ported to 25 architectures. It has been developing for over a 12 year in close cooperation with microcontroller producers. We have chosen it because it is becoming an industrial standard in the field of the embedded systems. It is the number one in the market survey of EETimes in two categories: The RTOS kernel currently used and the RTOS kernel considered for the next project. Finally, it is free of charge even in commercial application without a need of sharing user code and its memory overhead is highly optimized [3].

2.3 Used timing metrics – RhealStone

In 1989, Rabindra P. Kar and Kent Porter published a paper describing a way of RT performance testing – RhealStone. In that time there were metrics to compare performance of computational systems like xFLOPS or more sophisticated ones, Dhrystone or Wheatstone, but none of them was focused on system aspects relevant for RT computation. The goal is to quantify numerically the predictability and system ability in order to meet timing constraints and not the calculation performance. Rhealstone is based on 6 different characteristics:

- average task switch time – average time for switching between two independent active processes (tasks) of the same priority (i.e. effectiveness of (re-) storing process contexts) – t_1,
- average pre-emption time – average time for switching to a previously inactive higher priority process (i.e. effectiveness of schedule-dispatch algorithm) – t_2,
- average interrupt latency – average time between the occurrence of an interrupt at the CPU and the execution of the corresponding interrupt service routine's first instruction (i.e. effectiveness of dynamic priority assignment and scheduling) – t_3,
- semaphore shuffle time – an average time between a request to lock a semaphore, which is locked by another process and a time when its request is granted – t_4,
- deadlock break time – an average time to resolve a deadlock that occurs when a lower priority process holding a resource is pre-empted by a higher-priority process also requiring the same resource – t_5,
- inter-task message latency – average delay that occurs when sending a message between two processes – t_6[5].

In result each of the variables is multiplied by constants a_1, $a_2 \ldots a_6$, chosen by designer-defined weight so the benchmark is very flexible and able to truly reflect performance in any specific application [1]:

$$RhealStoneindex = \left(\frac{a_1.t_1+a_2.t_2+a_3.t_3+a_4.t_4+a_5.t_5+a_6.t_6}{6}\right)^{-1} [RhealStones/s] \qquad (1)$$

In this paper we have tested3 out of 6 characteristics (remaining 3 have weight of zero). We have chosen the ones with the highest impact on most RT systems, such as time to reschedule between same priority tasks, time of preemption, time of ISR delay handling.

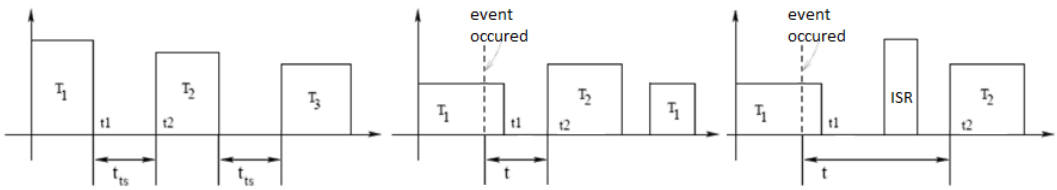


Figure 1. Visualisation of chosen RhealStone metrics.

All chosen metrics are visualised in the Figure 1. We have not considered message latency, semaphore shuffle time and deadlock break time (priority inversion), since those aspects are too specific for the application, but in particular cases there is a possibility to take those variables into account and measure them with the same method as described in this paper.

3 Tests outputs

In this section we are dealing with practical outputs and data measured by using proposed method.

3.1 Timing performance comparison

We have used the oscilloscope to evaluate and quantify metrics. In most cases only maximal reaction time is crucial and all system properties are informational, however it depends on the distribution of efficiency of the result in time (it is possible to calculate with mean values when the target appliance and its tasks has character of soft RT system).

Time delay of the task switch is shown in the Figure 2. In the case of RIOS the response is always around 10.1 µs without any jitter. FreeRTOS delay is reaching field from 37.4 µs to 38.4 µs, with peak around 38.1 µs. The ability of system to store and restore content is better in the case of RIOS.

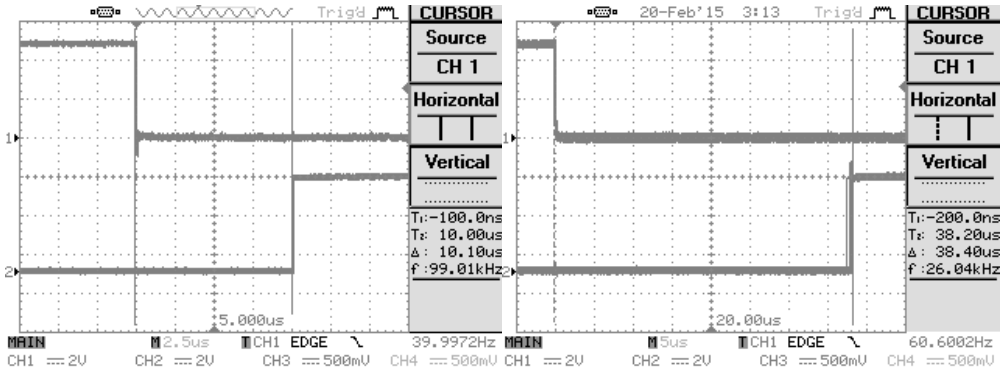


Figure 2. Measured time to reschedule between tasks with the same priority.

The pre-emption delay of systems is shown in the Figure 3. Lower priority task is toggling output, so it is possible to see its active stave. In the moment when the event occurs – higher priority task is ready and scheduler is woken up – the toggled output becomes stable. The metric is measured since this moment up to rising edge of higher priority task is detected. The time necessary to interrupt running task and start executing of prioritized task is distributed from 10.1 µs to 10.2 µs in case of RIOS. The FreeRTOS best response is 37.2 µs and worst one is 38.6 µs without any peaks.

The pre-emption delay is very similar to task switch delay in both cases, even different edges were read from oscilloscope. In case of task switch the delay was counted from falling edge of ending task to a rising edge of woken task. Pre-emption time was read between moment of toggling output stabilization and rising edge of higher priority task.

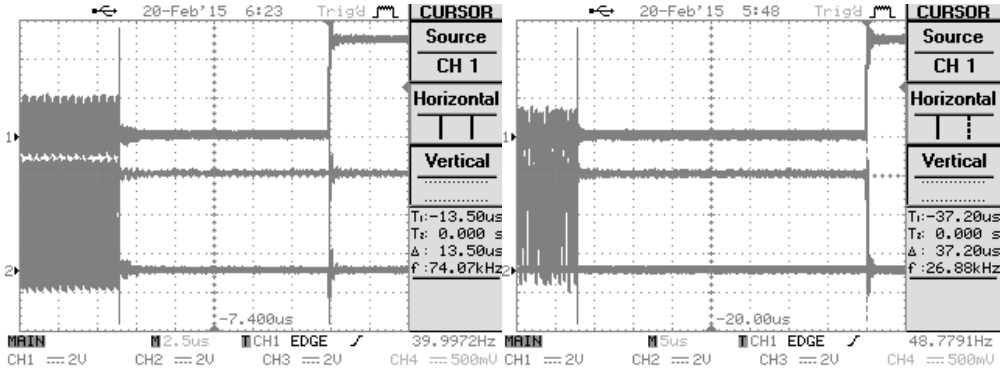


Figure 3. Measured time of preemption.

In the Figure 4, both pictures merge two measurements. Firstly the one with oscilloscope's persist mode is turned on, so there are all edges sequentially added in each iteration visible in the picture, thus the fastest and slowest system reaction is readable (minimum and maximum delay, jitter). Second measurement is done when oscilloscope's average mode is turned on, so count ratio and delay median are readable.

This last observed characteristic is the most complex one. In this case the system has to store running task, process all required steps to handle interruption and to start event handling task, thus there is a time of pre-emption, ISR and task switch included. The fastest response of RIOS is 5.6 µs and the slowest one is 22.6 µs with noticeable peak aroud 7.5 µs. The time response of FreeRTOS is mostly around 125 µs, but there are also events with worse delay up to 3 220 µs. The best response is 40 µs. If usage of the FreeRTOS as a hard RT system is considered, this characteristic is the weakest one (even it depends of specific time constrains of the appliance).

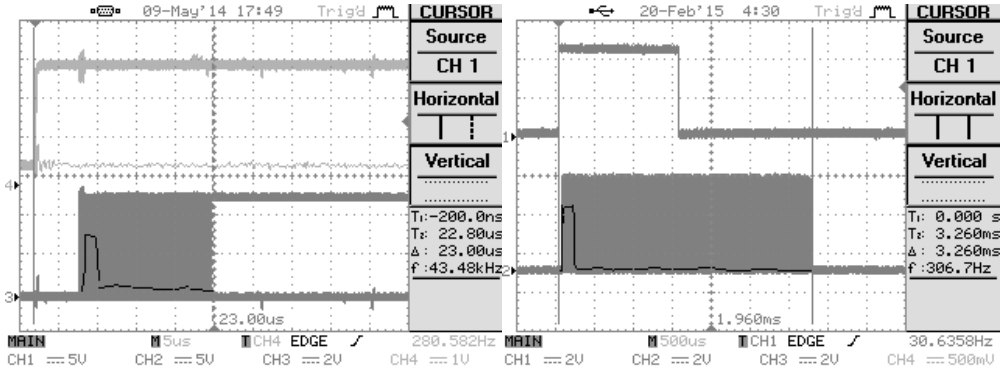


Figure 4. Time of measured ISR handling.

All measured data are reviewed in the Table 1:

Table 1. Tests outputs.

	RIOS (user-written scheduler)				FreeRTOS (RTOS)			
	T_{min}	T_{max}	T_{jitter}	T_{median}	T_{min}	T_{max}	T_{jitter}	T_{median}
Taskswitch	10.1 µs	10.1 µs	0 µs	10.1 µs	37.4 µs	38.4 µs	1 µs	38.1 µs
Pre-emption	10.1 µs	10.2 µs	0.1 µs	10.15 µs	37.2 µs	38.6 µs	1.4 µs	37.9 µs
Interrupt latency	5.6 µs	22.6 µs	17 µs	10 µs	40 µs	3.26 ms	3.22 ms	150 µs

In the Table 2 we are summarising overall score of tested architectures calculated by formula (1). In case of hard RT systems, it is needed to evaluate index using worst case values – T_{max}[2]. In case of soft RT systems, it is possible to calculate index using median values – T_{median}.

Table 2. Overall achieved score.

	RIOS (user-written scheduler)	FreeRTOS (RTOS)
Hard RT system score	139 860 [RhealStones / s]	1 800 [RhealStones / s]
Soft RT system score	198 348 [RhealStones / s]	26 549 [RhealStones / s]

3.2 Memory utilization comparison

As predicted, the memory resources are more utilized by complex FreeRTOS, even after removing all unnecessary functions. Avr-size tool has been used to quantify memory utilization. The chosen ATmega32 microcontroller is equipped by 32 KB of flash memory for program and separate 2 KB of static RAM for data.

As shown in the Table 3, program memory is utilized by 2.7 % in case of RIOS and keeps most of the resource for users code. On the other hand FreeRTOS needs 11.3 % to schedule periodic and aperiodic tasks, it's about 400 % increase in usage. Data memory comparison is meaningless for FreeRTOS, because it includes heaps of all tasks (RIOS doesn't allocate data memory of all tasks, compiler takes care of this, thus there is also no memory leakage check in the

case of RIOS). As a result of primitive way of data memory manipulation in RIOS there is no reason to penalise FreeRTOS. There is even higher risk of nondeterministic behaviour in a case of simple scheduler, thus there are other requirements for testing such an implementation.

Table 3. Memory utilization.

	Program memory (flash)	Data memory (RWM)
RIOS	886 bytes (2.7 %)	49 bytes (2.4 %)
FreeRTOS	3698 (11.3 %)	1624 bytes (79.3 %)

4 Conclusions

We have proposed a method to test RT performance of embedded systems based on microcontroller. The method is derived from paper on Microsoft Windows CE determinism test. We have demonstrated the possibility of RhealStone metrics to be used with this method, thus we have removed the lack of reusability of test outputs and enhanced the flexibility.

Performed measures showed results, that is the user-written scheduler reaches better delay times and lower jitter, thus it provides better determinism. On the other hand, there is a negative aspect of time necessary to develop any more complex OS features (RIOS only schedules set of periodic tasks, we have implemented aperiodic task handling and external interrupts support and much other functionality is usually needed even in simple appliances).

When mean values are taken into account, the performance of both systems is very similar. In case of hard RT systems, it is necessary to consider worst case scenario values – the maximal delays of system and then the RhealStone index of FreeRTOS is 80 times worse. This is due to general purpose and safe-oriented functionality of FreeRTOS.

In output data there is also a great example of a possible optimization of OS when it is developed by target appliance designers. The RIOS is handling external stimulus faster than task switch. It is because very efficient sporadic server to collect interrupts was implemented, but it was possible only because the functionality was added in way of following specific target characteristics, thus it is unsuitable for general purposes.

References

[1] Boger, T. J.: RhealStone Benchmarking of FreeRTOS and the Xilinx Zynq Extensible Processing Platform. Master's Thesis, (2013).

[2] Buttazzo, G. C.: *Hard Real-Time Computing Systems*. Springer Science, (2011).

[3] FreeRTOS: www.freertos.org

[4] Gumzej, R.: Real-time Systems' Quality of Service: Introducing Quality of Service Considerations in the Life Cycle of Real-time Systems. Springer (2010).

[5] Kar, R. P.: Implementing the Rhealstone real-time benchmark. In: *Dr. Dobb's Journal*. (1990), Vol. 15 Issue 4, pp. 46–55.

[6] Mazidi, M. A., Naimi, S., Naimi, S.: *AVR Microcontroller and Embedded Systems: Using Assembly and C*. Prentice Hall, (2010), 792 p.

[7] Ricci, L., Tacke, Ch.: Benchmarking *Real-time Determinism in Microsoft Windows CE*. [Online; accessed March 24, 2015]. Available at: http://msdn.microsoft.com/en-us/library/ms836535.aspx

[8] RIOS Scheduler. [Online; accessed March 24, 2015]. Available at: www.riosscheduler.org

[9] Valicek, M.: Implementácia súbežných procesov vo vnorených systémoch. Bachelor's thesis, (2014).

Dynamic Processes Controlled by PID Controller

František KUDLAČÁK*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
frantisek.kudlacak@stuba.sk

Abstract[1]

Control systems are common part of our lives. They can be found in vehicles, industry, computers and machines of daily use. Control systems can be divided into two main groups. The open loop systems and the closed loop systems. The open loop systems are simpler and they do not utilize advantage of feedback, but they can be easily implemented. Closed loop systems react on their output, so these systems are interconnected into the cycle by feedback. Most common closed loop control system in industrial controllers utilize Proportional-Integral-Derivative (PID) control logic. PID controller calculates error of system and in each loop PID controller attempts to decrease error value of system by adjusting control input to process.

PID controllers can be tuned with tuning methods. There are two main types of tuning methods, first is offline tuning methods and the second is online tuning methods. Offline tuning methods are based on setting parameters before planting process into execution. Parameters are not changing during execution of process, they can be changed only if execution of process is interrupted and setting parameters does not reflect behaviour of process automatically. If set parameters are wrong or they are not sufficient, process is interrupted and the parameters have to be set again. If process is running and parameters are set during execution, these methods are called online tuning methods. They can react to changes in process and adjust PID parameters according to changes. If process is dynamic or there is no description of process, offline tuning methods are not sufficient.

This paper deals with current state of art in field of tuning methods for PID controllers, and shows how output error depend on dynamicity of the process. Controlled process is dynamic, therefore it can change over with various dynamicity. These dependency leads to adaptive controllers, and new tuning methods for them. In the first part of the paper is presented general description of PID controller and tuning algorithms. There are mentioned online and offline tuning methods. In the second part of the paper, is mentioned first experimental result, which shows dependency of output error on dynamicity of process.

Acknowledgement: This work was partially supported by the Science and Technology Assistance Agency under the contract No. VEGA 1/0676/12.

---

* Doctoral degree study programme in field: Applied Informatics
Supervisor: Assoc. Professor Tibor Krajčovič, Institute of Computer Systems and Networks, Faculty of Informatics and Information Technologies STU in Bratislava

[1] Full paper available in printed proceedings, pages 345–352.

Distributed Firewall in Mobile Ad Hoc Networks

Jozef FILIPEK*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
jozef.filipek@stuba.sk

Extended Abstract

Traditional networks consist of central point, router, and other devices, hosts, which connect to it. Networks like that are called centralized or fixed, because router is a central point which other hosts communicate with. This system is used, when you can effectively protect your central point, physically and otherwise. But what if you need to immediately deploy whole network, where would be central point demerit or where would not be even possible to deploy centralized device? This is where MANETs (Mobile Ad Hoc Networks) come in. Every node in MANET acts as a router. This allows to spread network as far as most distant node is and nodes can act as intermediary devices. MANET is dynamic, decentralized and does not have well-defined perimeter. These properties make MANETs also vulnerable against a lot of attacks: eavesdropping, compromising nodes, man in the middle attack and so on.

Our approach is using firewall to increase security of the network. Basic concept of the firewall is to deploy it on entry/exit point of the network. Since MANET does not have fixed infrastructure and every node acts as an entry/exit point, every node has to have some kind of firewall mechanism. To make it more secure, we are using distributed public key infrastructure designed specifically for MANETs [1].

In the last years, several approaches have been introduced, which are securing MANET with a firewall: Distributed Firewall for MANETs [2], High Performance Firewalls in MANETs [3] and Fully distributed dynamically configurable firewall to resist DOS attacks in MANET [4]. Detailed analysis is omitted due to its length, but biggest disadvantages amongst these solutions were:

- Containing single point of failure type of node
- Not encrypting communication between secured nodes
- Lack of protection against compromising and compromised nodes

Fully distributed dynamically configurable firewall to resist DOS attacks in MANET [4] – this solution uses MPR to exchange information with other MPR enabled nodes. This information contains reject list and black list. Nodes in these lists are denied communication. Firewall does not deal with compromised nodes or securing communication between nodes.

After the analysis we were able to identify critical parts of our firewall concept for MANETs:

---

* Doctoral degree study programme in field: Applied Informatics
Supervisor: Assoc. Professor Ladislav Hudec, Institute of Applied Informatics, Faculty of Informatics and Information Technologies STU in Bratislava

1. Distribution of firewall mechanism – for this purpose we decided to include in our architecture distributed PKI for MANETs [1]. This PKI architecture is based on BATMAN routing protocol and uses complete public key cryptography. Asymmetric cryptography – symmetric keys will be sent throughout network with the help of RSA algorithm. This information will be signed and encrypted using asymmetric algorithm.
2. Symmetric cryptography – all data will be encrypted using AES algorithm. This encryption is less power consuming and faster than RSA algorithm. Firewall mechanism which will be encrypting traffic will be deployed between network and transport layer.
3. Network capabilities – capabilities serve as firewall mechanisms to allow communication between secured nodes. They define how much bandwidth nodes can use and what services they can use. Other constrictions can be added to capabilities. Capabilities are issued by competent nodes, each with different impact on the network. Every capability is signed by node which handed it out. Everything that is not permitted is denied.
4. Secured communication – routing protocol will be secured with PKI infrastructure, every routing information will be validated before accepting. Communication of upper layers between nodes will be encrypted using symmetric cryptography.
5. Intrusion detection system – while IDS/IPS system is not in the scope of this work, every node will need to have basic IDS/IPS capabilities to detect malicious behavior of the nodes in close proximity.
6. Any communication which is not allowed is by default denied.

We have presented concept of distributed firewall for general MANETs with IDS/IPS capabilities. Our solution adds another layers of protection to the network and protects it against attacks, which could cause collapse of the whole network or attacks, which would breach confidentiality between nodes. We used existing solution of distributed PKI [1] and added our own version of firewall mechanism on top of it. Added security comes with a price. Securing routing information and data exchange means increased processing power for devices and storing information associated with it requires use of more storage capacity. All this produces overhead on the network, since routing information is signed and used for sending certificates and all data are encrypted. Our concept will be implemented and tested in network simulator Omnet++ [5]. We will use several scenarios with static, moving and attacking nodes to thoroughly test our solution.

Acknowledgement: The paper was prepared with support of research grant VEGA 1/0361/15 entitled “Security in distributed computer systems”.

References

[1] Vilhan, P., Hudec, L.: Building Public Key Infrastructure for MANET with Help of B.A.T.M.A.N. Advanced. In: *Modelling Symposium (EMS)*, Manchester, IEEE, (2013), pp. 566–571.

[2] Alicherry, M., Keromytis, A.D., Stavrou, A. Bangalore: Evaluating a collaborative defense architecture for MANETs. In: *International Conference on Internet Multimedia Services Architecture and Applications (IMSAA)*, IEEE, (2009). s. 1–6.

[3] Zhao, H., Hangzhou, S.M.: High Performance Firewalls in MANETs. In: *Sixth International Conference on Mobile Ad-hoc and Sensor Networks (MSN)*, IEEE, (2010). s. 154–160.

[4] Akram, S., Zubair, I., Islam, M.H.: Fully distributed dynamically configurable firewall to resist DOS attacks in MANET. First International Conference on Networked Digital Technologies, IEEE, (2009). pp. 547–549.

[5] *OMNeT++ Discrete Event Simulator – Home.* [Online] [Dátum: 19. February 2015.] http://omnetpp.org.

Innovative Applications and Technologies

Kinlík: A New Game for Leapkin Therapeutic System

Matúš DEMKO*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
demko.mat@gmail.com

Abstract. In this paper we describemovement disorders causing movements disabilities and our addition to an existing therapeutic system. "Kinlík" is a game for rehabilitation of children with movement disabilities which allows patients to exercise their upper body. This game is a part of already existing therapeutic system "LeapKin" developed by students of Slovak University of Technology in Bratislava Faculty of Informatics and Information Technologies. Compared to other therapeutic systems, LeapKin is unique because of the cooperative gaming system, which helps patients to socialize while exercising. It also keeps track of patients movement characteristics and evaluates the progress of patient conditions. These elements are the key features of this therapeutic game. The cooperation of patients is achieved using the MS Kinect and Leap Motion sensors.

1 Introduction

There are manydisorders causing motor disabilities that can bring many problems into everyday life of the disabled person. People all around the world have been trying to help people with these disorders. Not only the experts, but also the parents, relatives and friends. These efforts resulted into various methods of early recognition and reparation of damage caused by these disorders. The most common form of these methods is classic physical rehabilitation. Physical rehabilitation includes stretching of musclesor exercising with some kind of tool or even a toy. This form of rehabilitation is not attractive for children because they get bored while exercising the same movement each and every therapeutic sessionand then don't pay much attention to it. The therapists are also required to oversee the rehabilitation process.

Nowadays there are many modern approaches such as therapeutic projects using computer games to treat disabled children and help the therapist (e.g. Innomotion, Speekle[1] etc.). Games that are used for higher purposes are called serious games. This work is a part of such project which is developed by students of Slovak University of Technology in Bratislava Faculty of Informatics and Information Technologies (STU FIIT). It is called "LeapKin" and is designed to help patients suffering mainly from cerebral palsy, apraxia and dyspraxia by playing games which require

* Bachelor degree study programme in field: Computer and Communication Systems and Networks
Supervisor: Martin Nagy, Institute of Computer Systems and Networks, Faculty of Informatics and Information Technologies STU in Bratislava

[1] http://speekle.sk

various movements to control games [7]. These games along with the game descibed in this paper are developed using the Unity game engine[2].

1.1 Movement disabilities

Movement disabilities may be a consequence of brain damage or other impairment of the central nervous system. Their character is very different with each individual patient. Some of the disabilities can be effectively supressed by rehabilitation but are not competely curable. In this paper we will discuss cerebral palsy, apraxia and dyspraxia.

"Cerebral palsy (CP) is a disorder considered to be the most common cause of movement disabilities" [1]. CPmainly causesphysical disability but isalso often accompanied by intelectual handicap.

CPcan have two forms: spastic or non spastic. Spastic form(spasticity) is characterized by involuntary contractionof muscles and non spastic form by very weak tension of muscles. This implies the need of exercising the muscles by rehabilitation. CP can have several symptoms which makes it difficult to focus on every one of them at once. These are for example:spasms, weak and stiff muscles, mental retardation in some cases, problems with articulation, patients can have difficulties sitting(weak back muscles). There are also other issues arising from the physical or mental disadvantages e.g. social denial, low confidence or self-esteem[8]. Therefore it is good to address these issues too.

Apraxiais described as an inability of a person to perform voluntary and skillful movements of one or more body parts. Dyspraxia refers to less severe forms of apraxia [2].

There are several types of apraxia:

- Ideational – difficulties executing complex movements that consist of simple movements which patients have no problem with;
- Ideo-motor – inability of gesture creation (fist closure, touching nose);
- Limb-kinetic – inaccurate limb movements;
- Orofacial (Apraxia of speech) – difficulties with use of facial muscles and speech creation [3, 4].

Both of mentioned disorders can effectively be treated by rehabilitation in their easier forms. Rehabilitation is performed in form of physiotherapy and various exercises (e.g. stretching exercises). Rehabilitation is most efficient in the early ages, specifically first years of life.

Besides the classic forms of rehabilitation there are also alternative and experimental forms. The most common alternative methods are Hippotherapy for younger children (riding on a horseback) and electrical stimulation for older children.

Serious games are a form of experimental method of rehabilitation using computer game principles for higher purposes. It is making the rehabilitation more fun and thus making the patients forget that they are actually exercising. This new method has been made possible by the modern inventions of motion sensors such as data gloves or other wearable body sensors and cameras.

2 Related work

Many researchers have created various computer games and systems for patients with movement disabilities. This type of game development has become more interesting to the developers by the arrival of the affordable motion sensors with software development kits available. Nevertheless rehabilitation games had been developed with expensive custom sensors or improvised motion sensors even before affordable motion sensors came to market.

[2] http://unity3d.com

A great example is a game “Catching dishes”, which is a part of a group of games developed for patients with spastic type of movement disorders such as cerebral palsy [5]. The goal of this game is to catch dishes flying towards the player usnig only the players hands. Hand movements arecaptured using web camera and a piece of cloth which the player holds in his hand. The camera is then used to determine the position of the hand based on the position of the cloth (shown in the Figure 1). The dishes stop at the end of their flight for a preset time, to make the catching easier to the player. This time is one of parameters that can be adjusted individually for each patient.

Figure 1. Gameplay screenshot from the game Catching dishes [5].

New devices allow easier data processing for a fair price. Leap Motion and MS Kinect are a great example of such devices. For example, Leap Motion device is used to capture hand movements in InnoMotion. InnoMotion is a rehabilitation system focused on exercising fine motor skills of hands and allows patients to exercise directly from home [6]. Main part of this system is an Internet application/web framework in which patients can perform various hand and finger movements including grabbing, finger lifting or wrist rotations etc. Data from the gaming sessions are captured and stored on the patients profile.

Students of STU FIIT have also been developing a therapeutic system called LeapKin, which will be extended by the game described in this paper. LeapKin therapheutic system is designed to help patients suffering from CP, apraxia and dyspraxia. Games in this system are designed to be very adaptable to specific needs of patients and should help them to overcome social and emotional barriers by making the players play the games together and cooperate. It also helps the therapist by storing the data collected from the patients game plays [7].

From mostof the projects oriented on rehabilitation of pacients with movement disabilities we get the same conclusion. It is necessary to provide adaptable difficulty of motions required to play the game properly.

3 Game

In this section we describe featuresrequired for the game, based on the previous research and the nature of the Leapkin system. We also describe the game principles and control mechanics.

Using computer games for treatment of movement disabilities is a complex task. It is necessary to implement several features:

- Rehabilitation aspects (movement of body and limbs) in a fun way,
- cooperation of players,
- adjustability of difficulty and accuracy of movements,
- game story to increase motivation to play the game,
- children friendly environment (e.g. no rough violence).

Players control the main character called “Kinlík” who is flying through the city on a hovering board (shown in the Figure 2). The goal is to collect specific game objects scattered throughout the level and avoid collision with various horizontal and vertical obstacles.

The game is played by two players. First player controls the movement of the game character. Second player has to protect Kinlík from hostile characters by activating a shield and reducing the damage done by the enemy or destroying the projectile if the shield is strong enough.

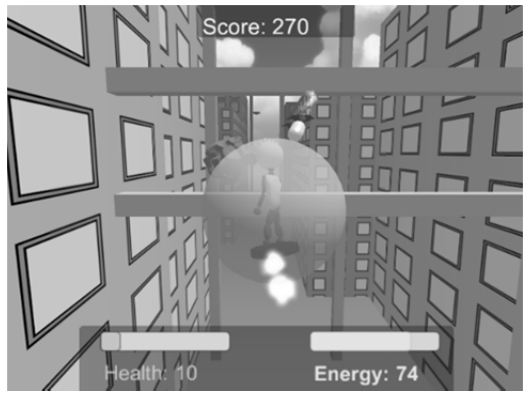


Figure 2. Gameplay screenshot from the game Kinlík.

Gross motor skills will be exercised by the player in front of the MS Kinect device:

- tilting the torso forward and backwards will affect the altitude of flight,
- tilting the torso sideways with outstretched arms (imitating an airplane) will move the player sideways.

These movements (shown in the Figure 3) strengthen the muscles of upper limbs and the back. It also improves the patient’s balance and orientation in space.

Figure 3. Sample of movements for movement control using MS Kinect device.

Fine motor skills will be exercised by the player using Leap Motion device:

- oscillating the forearm or palm up and down, will refill the energy for the shield,
- closing the hand will activate the shield(strength of the shield is determined by the fist closure strength).

These movements (shown in the Figure 4) strengthen the finger and arm muscles. In some cases a combination of both movements resembles the movement of picking up an object which is useful in daily life.

There are several factors of difficulty which can be adjusted in the options menu to fit each patient needs:

- frequency of the collectable game object (how often it is spawned),

- frequency of enemies,
- player movement speed,
- projectile and obstacle damage,
- accuracy and difficulty of required movements.

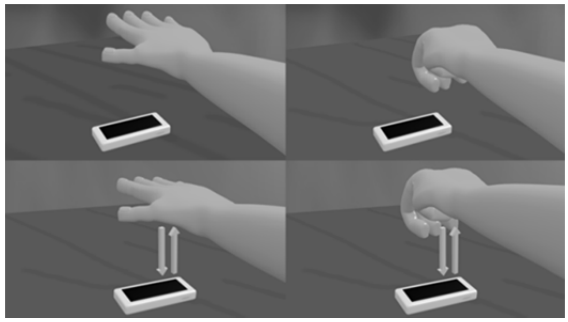


Figure 4. Set of movements recognized by the game using Leap Motion device.

Frequency of the collected game object is the frequency at which the collectable game objects are spawned. This influences the periodicity of making the decision by the player in front of Kinect, whether to move the body or stand still.

Frequency of enemies specifies the period at which the enemies attack the player, which sets the period of movements performed by the second player holding his hand over the Leap Motion sensor.

Player movement speed defines how much time the player has to perform the required movements. Projectile and obstacle damage defines the punishment for player performing wrong or slow movements.

Accuracy and difficulty of required movements is very important since the nature of disability and personal needs are different with each patient. Therefore it is necessary to provide an easy way of adjusting parameters like minimum angle of torso tilt or minimal amplitude of hand oscillation.

4 Evaluation

The game introduced in this paper will be evaluated in Research Institute of Child Psychology and Pathopsychology, Children's Center in Bratislava. Evaluation will consist of three parts. In the first part we will perform control tests to determine the level of movement skills of children such as screw test which is focusing on soft motor skills. In this test the therapist measures the time it takes a child to unscrewbolts from one board and then screw them to second board [8]. In the second part the children will regularly play the game Kinlík. In the third part we will gather feedback from children about the comfort during the play and ease of understanding the game. At the end we will perform the control tests again to identify improvement of movement skills.

The game will collect data from the gameplay such as various times (e.g. length of longest performed grip or maximum and minimum angle of torso tilt), score, number of hits by projectiles and collisions with terrain and obstacles. Additional statistics will be added to the system according to consultations and requests by the therapists from the Research Institute of Child Psychology and Pathopsychology.

5 Conclusions and future work

In this paper we have presented the new game for the therapeutic system LeapKin. We have introduced that the classic ways of rehabilitation are not sufficient due to lack of amusement for children. Moreover we introduced various projects which use game based rehabilitation systems.

We also presented design of the prototype (shown in Figure 2) of the game which will extend the LeapKin therapeutic system. The game is focusing on the upper body parts including torso, arms, forearms and fingers. It specifically helps with stretching and strengthening these muscles and exercise of grabbing and picking up objects. Players playing the game also have to cooperate to be successful.

Future plans include implementing features which will help patients to play the game – visual aids, slowing time in critical areas and automated evaluation of game score which will then change game settings either to make the game easier to prevent frustration or harder to challenge the players. We are planning to add specifically designed levels to focus on exact sequences of movements and new type of movement in level passage where players have to cooperate to move objects from one place to another.

Acknowledgement: This project was partially supported by Tatrabanka foundation under the contract No. 2013et029 and Slovak national research grant 1/0676/12. We also thank Lenka Lesayová and Jana Bieliková for guidance regarding the psychological aspect of the game and the staff of the Research Institute of Child Psychology and Pathopsychology, Children's Center (Výskumný ústav detskej psychológie a patopsychológie, Detské centrum) for guidance regarding the physical aspects of the therapy.

References

[1] Chamberlin, S. L., Narins, B.: *The Gale Encyclopedia of Neurological Disorders Vol. 1*. A-L. Farmington hills MI : Thomson Gale, (2005). ISBN: 0-7876-9160-7

[2] Latham, S.: Dyspraxia. In *Encyclopedia of Autism Spectrum Disorders*. Volkmar, FredR. Springer, New York, (2013). 1023-1023 ISBN: 978-1-4419-1697-6 DOI=10.1007/978-1-4419-1698-3_340http://dx.doi.org/10.1007/978-1-4419-1698-3_340

[3] Lexikón medicíny. Asklepios. [Online; accessed January 24, 2014]. Available at: www.datasolution.sk/pdfenc/a7.pdf.

[4] GoldmannGross, R., Grossman, M.: Update on Apraxia. *Current neurology and neuroscience reports*, (2008). Vol. 8, no. 6, pp. 490–496.

[5] Geurts, :, Vanden, V. A., Husson, J, Windey, F, Van Overveldt, M., Annema, J.H., Desmet, S.:Digital games for physical therapy: fulfilling the need for calibration and adaptation. In *Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction* (TEI '11), (2010) ACM, New York, NY, USA, pp. 117–124.

[6] Chen, L., Yan, N., Kiang, M., Muth, A.S., Krishna, K.S.: Innomotion: a web-based rehabilitation system helping patients recover and gain self-awareness of their body away from the clinic. In: *CHI '14 Extended Abstracts on Human Factors in Computing Systems (CHI EA '14)*, (2014). ACM, New York, NY, USA, pp. 233–238.

[7] Burda, K.,Grežo, R., Hasin, M., Kohútka, L.: Therapeutic System for Children with Movement Disorders. In: *IIT.SRC 2014, 10th Students Research Conference in Informatics and Information Technologies Bratislava*, (2014). Proceedings. 1. vyd. Bratislava: Nakladateľstvo STU, 2014, pp. 565–570. ISBN 978-80-227-4153-8.

[8] Vančová, J.: Využitie videohier s aplikáciami virtuálnej reality v stimulačnom programe pre deti s detskou mozgovou obrnou. TRNAVA: Trnavská univerzita v Trnave, Filozofická fakulta, Katedra psychológie. Bachelor thesis, (2010).

Reservation System for Coach Transportation

Marián ČARNOKÝ*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
majo.carnoky@gmail.com

Extended Abstract

Most of us are using smartphones every day and we have these devices in our pockets all day long. Nowadays there are people that rather use it [1] as use desktop computers or laptops because they are lighter, smaller and they have almost same functions as computers. However it is not used for bus ticket booking as often as buying ticket directly on bus.

Our system consists of several parts which must be interrelated to make system operates and performs its function. Some of these parts are located directly in the bus and others are located outside the bus. System architecture is shown in Figure 1. Architecture components are also located in FIIT Molpir s.r.o. laboratory at FIIT STU.

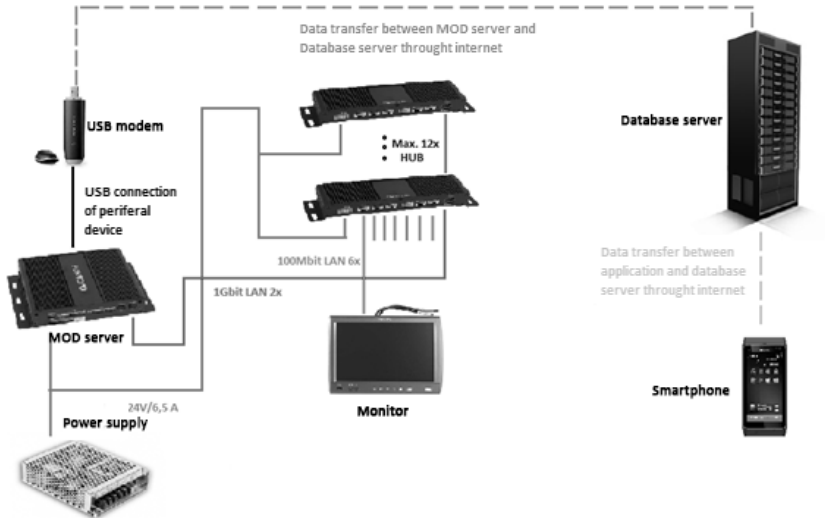


Figure 1. System architecture.

First user needs to start the application for ticket booking and then he needs to login or register a new account if he does not have one yet. After login, he could press the "buy ticket" button.

* Bachelor study programme in field: Informatics
Supervisor: Dr. Peter Pištek, Institute of Computer Systems and Networks, Faculty of Informatics and Information Technologies STU in Bratislava

After pressing the “buy ticket” button, application will open up a window with ticket buy form. At first, user needs to select starting destination, final destination, date and time of travel. The next step is to select ticket type from the list and then choose a seat place. After selecting seats, the summary of ticket booking will appear.

The bus driver has to check passenger’s purchased tickets while they are boarding. Tickets will be able to check when the driver starts the application and login to the system. The driver uses the same application to manage tickets as passengers use to purchase tickets, it has different functions and permissions. At first the bus driver needs to enter a bus ride number. Only drivers in duty can view and edit purchased tickets of bus ride. Now he is able to view all tickets of his ride and he has some options for manage. If boarding passengers do not have purchased ticket, the bus driver will be able to sell them one via system.

In Table 1 is a comparison of our solution (SmartBus) and competitive solution [2, 3, 4]. There is a comparison of features and functionality of each solution.

The rows which names are not self-explanatory in table 1 are:

1. Android app. – system implemented for Android platform.
2. History – view history of purchases.

Table 1. Functionality comparison of solutions.

	SmartBus	**Student Agency**	**redBus**	**Slovak Lines**
Android app.	✓*	✓	✓	×
Ticket reservation	✓	✓	✓	✓
Cancel ticket in system	✓	✓	✓	×
History	✓	✓	✓	✓
No need to print ticket	✓	✓	✓	×
Seat reservation	✓*	✓	✓	×
Benefit points	✓*	×	×	✓

* feature will be implemented in near future.

History and ticket reservation is implemented by all solutions. Slovak Lines did not make a system for Android platform from compared solutions, but they implemented a benefit point feature which will be also implemented in our solution. Other features are used by all solutions except Slovak Lines.

Acknowledgement: This work was partially supported by the Slovak Science Grant Agency (VEGA 1/1008/12 “Optimization of low-power design of digital and mixed integrated systems“).

References

[1] Smith A.: 17% of cell phone owners do most of their online browsing on their phone, rather than a computer or other device. Pew Research Center’s Internet & American Life Project, (2012).

[2] Student Agency: *Cestovné lístky*. [Online; accessed February 19, 2015]. Available at: https://play.google.com/store/apps/details?id=cz.studio9.sa

[3] Slovak Lines*: Kúpiť cez Internet.* [Online; accessed February 19, 2015]. Available at: http://www.slovaklines.sk/moja-cesta-so-slovaklines/predaj-listkov/kupit-cez-internet.html

[4] redBus: *redBus - Bus Ticket Booking.* [Online; accessed February 19, 2015]. Available at: https://play.google.com/store/apps/details?id=in.redbus.android

Multimedia System for Public Train Transportation

Ondrej GONO*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
gono.ondrej@gmail.com

Extended Abstract

The field of information technology is rapidly expanding and it affects our lives in many different ways; travelling is not an exception. Transport operators are constantly trying to enhance the comfort and to offer new services to the passengers by which they open the door to the new information technology. At the present time, developers of multimedia systems in public transportation focus on information systems with or without central screens, MOD (media on demand) systems and comprehensive systems with many offered services.

On the market, there are some MOD systems solutions that are very similar. Differences lie in the services they offer, for example an option to play satellite TV, to access the Internet, to play games or built-in passenger information system. There are often differences in system parameters like resolution, screen size, available applications, performance of server and number of touch screens which can be connected to one MOD server.

In MOD systems, the passengers use integrated devices; some systems do not have them so the passengers use their own mobile phones, tablets or notebooks. The passengers can connect themselves to the network in those vehicles which provide access to the Internet, television or other multimedia content. Moreover, some systems provide much more varied possibilities than entertainment services for customers, for example passenger information system, real-time video surveillance system or they gather train information and send them to the transport operator center.

Our work is focused on multimedia system for train transportation that can play multimedia content on central screens in the carriages and it also offers options to manage the whole system. The aim of our work is to improve the existing *Train Multimedia System* [1] by adding the central carriage component which will use the Funtoro facility. We designed software for this component in order to play multimedia content and to communicate with the implemented elements of the system. Moreover, we have to ensure that the system will work properly even when the connection is lost. The whole design of the system brings a solution for open chair cars with central screens for which there is only little competition in the world.

Our suggested solution provides central management of several trains. Each train has a main train component on which multimedia content is stored. This component communicates with the central carriage components that are in carriages. They ensure playing of video on central screens even if the connection between the main train component and the central carriage components is

* Bachelor study programme in field: Computer and Communication Systems and Networks
Supervisor: Dr. Peter Pištek, Institute of Computer Systems and Networks, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 365–366.

lost. The system we have designed can be divided into three logic sections – controllers (control panel, web interface), main train component and central carriage component.

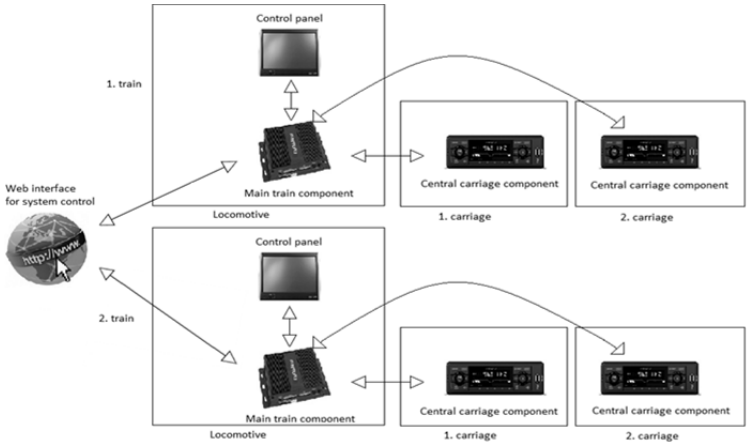


Figure 1. Scheme of suggested solution.

Controllers and main train component are already implemented and we will make changes in these components only if necessary, to ensure that the whole system will work properly. The main aim of our work is to implement software for Funtoro server which will work like the central carriage component. We defined these functions for central carriage component:

- Control of multimedia content – server gets playlist from the main train component and then it downloads items from this component. It has to manage removing of useless content, too.
- Play content – server plays downloaded content on central screens following the playlist.
- Play advertisement – server downloads advertisement from the main train component and play it on central screens in carriage according to the advertisement settings.
- Store information about played items – server has to make a history of played items and played advertisement or to send these information to the main train component.
- Continuation of the work after the connection is lost – if central carriage component and the main train component loose connection, the central carriage component will continue with playing items according to the playlist.
- Synchronization – after re-connection, the central carriage component will synchronize with the main train component. In case of need, it sends information about played content through the system breakdown; it gets new playlist and downloads new content.

We presented the domain of multimedia in public train transportation where these services increasing comfort of the passengers are not so common. Solution, as we designed it, brings to the transport operator an opportunity to enhance offered services with relatively small amount of resources and in the way that is very easy to install to vehicles in use.

Acknowledgement: This work was partially supported by the Slovak Science Grant Agency (VEGA 1/1008/12 "Optimization of low-power design of digital and mixed integrated systems").

References

[1] Ignačák, M., Kincel, A., Maruniak, M., Minárik, M., Straka, V.: Aplikácia pre platformu Funtoro. Team Project, Slovak University of Technology in Bratislava, 2013.

Information System for Public Transport

Martin HUDEC, Martin BAŇAS*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
xhudecm1@stuba.sk, xbanasm@stuba.sk

Extended Abstract

In this paper we focus on introducing an improved way of travelling via public transport, which becomes the backbone of many bigger towns and cities. Nowadays people drive by cars for short distances, often alone in a car. This environmentally unfriendly situation may be improved by raising the frequency of public transport usage. We designed a system, which provides users of public transport with multimedia content and real-time information such as line delays and current positions of vehicles. The content and information can be easily accessible via mobile devices and information displays built-in public transport vehicles. Good information system may provide popularity boost in public transport. There are many possibilities to enhance current applications used by travelers by making them more intuitive and easier to use [1].

Determining the user's position can be considered the main breaking point, which offers developers the ability to build applications, which can provide more accurate information to the end users and extend the use of their mobile devices potential [2]. A problem with creating this kind of system is the way of getting the real-time data from the vehicle. To solve this problem and be able to create a system, which is independent from the vehicle manufacturer we have chosen Funtoro Infotainment server. This box has standardized 1 DIN size, runs Android OS and has 1TB storage capacity for multimedia content. The server uses the connected GPS antenna to collect location data and transmits it in real-time to the server using the connected 3G module. Collected data can be used for real-time traffic monitoring or later for statistical purposes. "This project is divided into three parts: the mobile device application (MA), the display application (DA) installed in vehicle and server application (SA) which provides communication also with database.

The DA is designed as information and entertaining center in vehicles. It provides basic services like information about upcoming operated stations, date and time, transfer lines at stops and announcements of transport company. This information is displayed on monitors built in vehicles (see Figure 1). Thanks to known position, monitors also provide calculated delay and real-time arrivals to the bus stops. The main purpose of the DA is to show predefined multimedia content as the vehicle drives through the specified geographical areas, known as point-of-interest (POI). If the content is video sequence the application reduces the section with stops in order to get sufficient space to play it. In case of audio sequence the section with stops stays unchanged. The vehicle's position information is sent to SA.

Server application is the main node of the system. It provides the communication among the applications and the database. The application also runs the web server, which is proposed for management of the POIs and it uses API to talk to MA and DA. All requested data is converted to JSON structures and transmitted using HTTPS connection for improved security. The origin of

* Bachelor degree study program in field: Computer and Communication Systems and Networks
Supervisor: Dr. Peter Pištek, Institute of Computer Systems and Networks, Faculty of Informatics and Information Technologies STU in Bratislava

transit information is the GTFS (General Transit Feed Specification) stored in MySQL database. The GTFS standard provides the speed and simplicity in transit content lookup.

Figure 1. Display application on the left, mobile application on the right.

For the mobile application, Relative Layout of Android OS has been chosen. It enables us to specify the location of it's child objects so they can be put anywhere on the screen. The main control part of the application is the Sliding Tab Panel, which redirects the user to 3 different views from which the first one is shown in Figure 1. It is map view, where the nearest stops and accurate locations of vehicles are shown based on user's location. The second one allows user to find information about chosen stop together with arrival times and delays. The last one provides user with nearest stops and possible transfers acquired by his current position. MA requires Internet connection and also active GPS or network localization (Cell-ID Aided Positioning System) [2]. The application is developed for the newest Android version 5.0 Lollipop, but it is also compatible with devices running older versions of Android OS. Since the Google's Material Design was defined alongside with the Android 5.0, the developed application will be based on it.

Providing relevant information to passengers of public transport should be the most important functionality of this information system. There are also ways in which this project's functionality can be enhanced. In case of the mobile application there can be features like predefined stops, which would be available in special, constantly updated, view. Passengers not having the vehicle's displays in direct visibility or having vision deficiency may also see current content related to their line on their mobile devices using the mobile application. The display application's functionality can be extended by providing departure times of subsequent lines on the next operated bus stop. Vehicles headed to transportation nodes (i.e. bus or railway stations, airports, etc.) may display departure or arrival times of important lines. According to the fact, that this system is developed simultaneously as mobile application for personal mobile devices and display application for vehicle information panels, it brings the benefit of one consistent system on the visual and also functional dimension.

Acknowledgement: This work was partially supported by the Slovak Science Grant Agency (VEGA 1/1008/12 "Optimization of low-power design of digital and mixed integrated systems").

References

[1] McWherter, J., Gowell, S.: Professional Mobile Application Development. John Wiley & Sons, ISBN: 1118240685, 9781118240687, ITPro collection (2012)

[2] Paek, J. et al.: Energy-Efficient Positioning for Smartphones using Cell-ID Sequence Matching, 2011, http://enl.usc.edu/~jpaek/docs/mobisys11_caps.pdf

Integration of Smart Technologies to Support Fitness Activities

Roman KOPŠO*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
kopsor@gmail.com

Extended Abstract

The rising number of sedentary jobs and overuse of technology made us lazier thus neglecting the basic need of our bodies for some physical activity. Therefore, the primary aim of this work is to use the advantages of already existing smart devices to develop a fitness application that would counter the negative effects of our unhealthy lifestyle.

This application uses the integration of three smart devices, namely, smart phone, smart watch and smart TV. The main feature of this system is to monitor user's heart rate during the workout and display it on the TV screen, while also enabling the user to watch TV channels. Apart from the information about the heart beat, the user can also view how many calories he or she has burned. Each one of the devices plays an important role and all of them have to be used.

The smart watch is primarily used for measuring the heart rate thanks to its heart rate sensor. We have used the LG G Watch R (W110) Black for this purpose. This watch runs on Android Wear OS. The software on watch uses Google Data API for formatting data into Data Items, which provide storage for information that is automatically synchronized between the watch and the phone via Bluetooth connection. Every time when the speed of heartbeat changes, for instance, from 75 beats per minute to 78, the packet of data is sent. The only thing the user has to do for the application to work is to tap the watch display and choose the fitness app icon. Application will then automatically start measuring heart rate and simultaneously send data into the mobile phone.

The smart phone receives the data about heart rate and then sends it via wifi connection to smart TV. Sony Ericson Xperia Mini ST15i smart phone has been used as a gateway between the watch and the TV. The highest official supported Android version for this mobile phone is 4.0.4 and smart watch could only be synchronized with smart phones that are running on Android 4.3 or higher. To solve this problem, we have installed Cyanogenmod 10.2[1] which is an open-source operating system based on Android 4.3 platform, hence making it possible for us to pair the phone with the watch. We are working with Samsung smart TV, so we had to use the technology they provided for development. We have used Samsung's multi screen SDK[2]. This SDK is used for communication in local area network between smart TV and devices like smart phones or tablets. In our phone application, we have implemented the functionality for sending JSON messages

---

* Bachelor degree study programme in field: Computer Engineering
Supervisor: Dr. Peter Trúchly, Institute of Computer Systems and Networks, Faculty of Informatics and Information Technologies STU in Bratislava

[1] http://www.cyanogenmod.org/

[2] http://multiscreen.samsung.com/

between the TV and the phone. Then we have added Wearable Listener Service for receiving the data from the watch. This data is sent periodically in three second intervals to the smart TV using Samsung's multiscreen SDK. The user has to turn on his Bluetooth and wifi connections on his phone. Then he has to sync the watch and the phone using Android Wear app and finally start our app, where he detects available smart TV in LAN and activates the sending process.

The television is the last and main aspect of this whole networked system. We have used Samsung ue55h6470 Smart TV. The development of applications for Samsung smart TV uses javascript, css and html. It is similar to making an interactive website. The smart TV variation of our fitness application uses functionality from Samsung's multi screen SDK for receiving JSON messages. It is able to obtain incoming data about heart rate sent by phone and show the output on the TV screen afterwards. Currently, the application is capable of showing the output of the heart rate in number and graph format which updates every three seconds. Features which will be added are two pop-up windows, one for help and the other for user input information. This information (gender, age, weight, resting heart rate), will be used for calculation of five heart zones[3] and the quantities of burned calories. Heart zones are intervals of the speed of heart beats per minute. Each of these zones has different effect on our health and our app will inform the user in which heart zone he is (and should be) situated during workout. Karvonen formula [1] together with maximum heart rate formula [2] will be used for the calculation of heart zones. Energy expenditure equations [3] also with the maximum heart rate formula [2] will be used for computing the amount of burned calories. The general layout of the app will be comprised of four blocks in total. User can choose to hide these blocks and watch the TV channel with the app running in background. The hide and show function will be possible with the use of A, B, C and D keys on the remote control. Block A will be in the far left side showing info about the burned calories. Block B will be in the middle taking the largest part of the screen showing the dynamic heart rate graph and info about heart zones together with buttons for pop-up windows. Block C will be in the far right side showing the heart rate data. Finally, block D will be a pop-up window of smaller frame showing the overall status of the workout.

This work was partly inspired by the projects carried out by [4, 5]. The application that we created can be considered as innovative because nothing similar has been created so far. With our work, we have contributed to the more efficient fitness exercise indoors.

Acknowledgement: This work was partially supported by Slovak national project MINCA KEGA 047STU-4/2013.

References

[1] Karvonen, J., Vuorimaa, T.: Heart Rate and Exercise Intensity During Sports Activities. *Sports Medicine*, (1988), vol. 5, no. 5, pp. 303–311.

[2] Tanaka, H., Monahan, K.D., Seals, D.R.: Age-predicted maximal heart rate revisited. *Journal of the American College of Cardiology*, (2001), vol. 37, no. 1, pp.153–156.

[3] Keytel, L.R., Goedecke, J.H., Noakes, T.D., Hiiloskorpi, H., Laukkanen, R., Merwe, L., Lambert, E.V.: Prediction of energy expenditure from heart rate monitoring during submaximal exercise. *Journal of Sports Sciences*, (2005), vol. 23, no. 3, pp. 289–297.

[4] Sornanathan, L., Khalil, I.: Fitness Monitoring System Based on Heart Rate and Sp02 Level. In: *Information Technology and Applications in Biomedicine (ITAB), 2010 10th IEEE International Conference on*, IEEE, (2010), pp. 1–5.

[5] Seetharamu, V.K., Bose, J., Sunkara, S., Tigga, N.: TV Remote Control via Wearable Smart Watch Device. In: *India Conference (INDICON), 2014 Annual IEEE*, IEEE, (2014), pp. 1–6.

[3] http://www.polar.com/us-en/support/tips/Polar_Sport_Zones

Mobile Application for Monitoring Athletes and Patients

Zuzana KUCHARÍKOVÁ*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
xkucharikovaz@stuba.sk

Extended Abstract

The main goal of this work is to enable connection between a coach and a athlete/patient and provide online information about training through mobile application. Thanks to highly precious information about athlete's/patient's performance and heart rate during training.

Mentioned application is developed primary for fitness area which is very popular in these days. The coach tracks performance of top athlete and is able to manage trainings. The application can be also used in more useful sphere. For example patients who needs rehabilitation exercises or overweight people. It's very important to follow regularity of trainings to achieve the desired progress in treatment.

There are many existing solutions being used by mobile application aimed at tracking user's health and performance activities. Also we can find many applications on today's marketplace designed for coach that support and simplify analysis of sports performance. However, none of these applications connects coach and athletes/patients directly and cannot display real time data of athletes/patients performances on coach's device.

The application distinguishes two types of users. The coach, who organizes trainings and athlete/patient, who have no permissions to create new users or trainings but they can evaluate each exercise. Athlete/patient cannot edit data which have been entered by the coach.

Users log in to the application using username and password which are generated by the application. After entering credentials application sends request to the database. Database finds hashed password according to username and sends it back. Then the application compare hashed password from database and hashed password from entered credentials. After successful login, main menu is displayed to user.

Coach is able to use this features:

- Add training's unit. Each training belongs to some specific unit of individual athlete. This unit include individual athlete/patients, specific group of whole team created by the coach.
- Add training. This feature includes name, type of a group, maximum heart rate, duration, difficulty, description and other information such as date and time.
- Calendar. Each user is able to view calendar with trainings related to them.

---

* Bachelor degree study programme in field: Informatics
Supervisor: Dr. Tomás Kováčik, Institute of Computer Systems and Networks, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 371–372.

- Biorhythm chart. This chart shows daily high reached goals of athletes/patients during individual workouts. Athletes/patients can be sorted by the filter included in feature.
- Percentage chart. The percentage of the power performed during the training. Reflects if was the required performance achieved and to what measure.

Athletes/patients are able to use Calendar, Biorhythm chart and Percentage chart. Charts are displaying data that belongs only to certain athlete/patient.

In co-operation with Faculty of electrical engineering and information technology, we would like to broaden the concept of the mobile application with external device. This monitoring element is able to scan various physical functions of human body such as ECG and respiration. Additional sensors are included in the probe such as accelerometer, gyroscope and barometer.

Basic communication between the application and external device utilizes Bluetooth (2.0) technology. Security is ensured by specific protocol defined by developers and implemented in the probe. In order to make this communication in correct way the protocol must be implemented also in mobile application.

One of advantage of this system is immediate connection between user's and the possibility of feedback. Another advantage is communication with reliable external device which offers plenty of precise information about cardiac activity of user.

References

[1] Gartner, I.: Gartner Says Annual Smartphone Sales Surpassed Sales of Feature Phones for the First Time in 2013. 13. 2. 2014. [Online]. Available: http://www.gartner.com/id/2665715. [accessed 1. 12. 2014].

[2] Ďásek, M.: Biometrika. 26. 2. 2003. [Online]. Available: https://akela.mendelu.cz/~lidak/bif/dasek.html. [accessed 7. 12. 2014].

[3] Statista, I.: Number of apps available in leading app stores as of July 2014. 7/2014. [Online]. Available: http://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/. [accessed 10. 12. 2014].

[4] N. I. o. Health: National Heart, Lung, and Blood Institute. 9. 7. 2012. [Online]. Available: http://www.nhlbi.nih.gov/health/health-topics/topics/hb/understanding. [accessed 19. 3. 2015].

[5] Gibbs, S.: Fitbit Surge review: a fitness tracking watch that's not quite super. 13. 3. 2015. [Online]. Available: http://www.theguardian.com/technology/2015/mar/13/fitbit-surge-review-fitness-tracking-watch. [accessed 19. 3. 2015].

Gesture Library and its Application in a Game

Marián KURUC*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
mariankuruc1@gmail.com

Abstract. The use of hand gestures in applications is a technology that is still in its early stages. The gestures are most of the time not very well thought out and users are having hard time remembering them easily. Additionally a lot of the gestures are not designed sufficiently different, which in turn often makes the application perform undesired results. Individual applications use different sets of gestures, which in the end makes the user confused, if he uses more than one application. In this paper we provide a solution to these problems by defining our own gesture library and present its usability in a game. The gesture library evolves while users play the game and constantly improves based on the inputs from the players. The game was implemented in a 3D world, uses modern graphics and elaborate algorithms of artificial intelligence. The designed system provides recognition of multiple types of gestures that are interpreted by a game, which as results performs the desired action. Recognition of the gestures is accomplished with the Leap Motion device.

1 Introduction

Human-computer interaction (HCI) provides multiple ways to communicate with computers, while one of the continuously evolving fields is the one that deals with alternative methods of input. In this paper we focus on the use of hand gestures to control a game, which are received from a gesture recognition device.

As of today, there is no extensive library of hand gestures that is being used application-wide. The lack of this makes it a problem addressed in this paper. In our project we utilize the gesture recognition device, Leap Motion controller, which is able to track movements of all ten fingers simultaneously. The precision and speed of this device makes it one of the leading gesture recognition devices in today's world.

We have designed a library that contains significantly unique gestures by which we ensure that the system encounters less to none problems, while detecting the gestures. A lot of attention has been also put towards making sure that the gestures are intuitive enough for users to learn and remember. Because of this, we have designed a solution to let the users help us mark the gestures that are intuitive, without making it too intrusive. A game with a purpose is one of the best solutions to this problem. The game ensures that players are having fun, while playing with

* Master degree study programme in field: Information Systems
Supervisor: Dr. Juraj Štefanovič, Institute of Applied Informatics, Faculty of Informatics and Information Technologies STU in Bratislava

different kinds of hand gestures. During the time the players play the game, the system collects data about the most used gestures, and also gestures that are players having problems to perform. Based on these calculations the system evolves the library, improving gestures for each specific action a player performs in a given game.

2 Related work

There are several examples of other people's work, where they try to design games using the gamification process to extract the needed information. One of the examples, where a game has been used successfully with a goal of acquiring information, is the Google's attempt to label images. Several applications such as search engines and accessibility programs require accurate image and video labels. However, there are no common principles about writing appropriate text labels for the millions of videos on the web, and computer vision is not yet capable of determining their content accurately.

This problem was solved by designing a multiplayer game, called Google Image Labeler, where two people were shown the same image and their goal was to label it until they used the same label. After that the game has moved onto a different image and the people have repeated the process. In the light of this process the company has by now acquired labels for an enormous set of images.

At University of California in Berkeley, several students have participated on a project, where they have designed a lexicon containing a large set of gestures. These gestures were created by the authors and had several test subjects to manually rate them on their effectiveness, preference, match and ease. The results were based on the preferences of the people rating the gestures, and due to the small diversity of the testers, these were not correlated.

The Leap Motion device has been also used in several projects that help people with their problems. One of them is a design of a system using Australian Sign Language, to help people with a disability to speak and hear to communicate. In the study, they have performed several tests to recognize various letters of the Australian Sign Language, and discovered that due to the limited recognition angle of the device, various symbols are impossible to recognize. Since the device is unable to recognize gestures with a titled hand around the Z axis, it was impossible to recognize most of the numbers. Additionally, any attempt of recognition where the fingers are pressed too close together, makes the device not recognize the gestures properly.

3 Our method

An important part of this paper is the gesture-based interaction in recognizing gestures. Recognition of 2D patterns, drawn by users with their finger in front of the Leap Motion device, requires capturing the drawn points and storing them in a data structure. We had to make the algorithm to perform fast search and were willing to sacrifice certain precision of recognition for the speed of the recognition. It was needed due to the fact the gesture recognizer is located in the game and players need to receive the output fast, without losing the dynamic perception of the atmosphere.

The algorithm utilizes unistroke instance-based nearest neighbor classifier with Euclidean scoring and uses several training samples that are stored in a data structure. At runtime, an unknown gesture is then compared against these samples and the most similar instance is used as the output of the algorithm, while the similarity implies how precise the result is. It is hard to foresee what gestures would like the user to add and what would they look like. Fortunately, the algorithm is able to recognize most of the shapes and is usable on a variable dataset. However, the algorithm is not able to identify gestures that depend on specific orientations, aspect ratio or location, so the system avoids using gestures with these attributes. Despite of that, the algorithm offers recognition for a large scale of symbols that can be found in our game.

3.1 2D Gestures

The designed algorithm needs to pre-process the input it receives before it tries to match the drawn gesture to one of the gestures in the library. There are several pre-processing techniques the algorithm uses described in this chapter.

3.1.1 Resampling

Gestures in user interfaces are captured at a rate determined by the input device and the software used for capturing. Along with that the movement speed has a clear effect on how many input points are captured when recording the gesture. If we move our finger slower around the screen, the system will capture more points, while the faster movement will capture less recognition points. Because we need to compare the individual gesture strokes that can be captured at different movement speeds, we need to resample this input afterwards with an N of equidistantly spaced points. If we use a larger N we increase the time it takes to compare the gesture, but we also increase the precision. We have found that the average number of N equal to 64 is adequate for most recognition attempts.

Resampling the captured stroke requires to calculate the total length of the drawn gesture and divide it by the (N-1), where N is the number of equidistantly spaced points. This gives us the length between each increment I, between N new points. We then search through this new path and if the distance covered exceeds I we add a new point to the gesture through linear interpolation.

3.1.2 Rotate based on the Indicative Angle

Each time the gesture is drawn it has an offset in the angle. To correct the angle at which the gesture was drawn, the algorithm rotates the whole shape based on the Indicative Angle. The Indicative Angle is an angle formed between the centroid of the gesture and the gesture's first point. After its calculation, the gesture is rotated to be aligned at zero degrees of the angle.

3.1.3 Scale and Translate

After rotating the gesture the algorithm needs to scale it into a reference square, to keep all gestures at the same aspect ratio. There is also a need to translate the gesture to a reference point. To keep it simple, it translates the gesture to keep its centroid at zero (x = 0, y = 0).

3.1.4 Compare

To compare the drawn gesture against the gestures that are stored in our data structure, we use the cosine similarity, for which we need to calculate the optimal cosine distance. The cosine distance gives an angular distance between two sets of vectors. Before we can calculate the optimal cosine distance we need to calculate the vector lengths (a, b) as follows:

$$a = \sum_{i=0}^{n}(v1[i] \times v2[i]) + (v1[i+1] \times v2[i+1]) \quad (1)$$

$$b = \sum_{i=0}^{n}(v1[i] \times v2[i+1]) + (v1[i+1] \times v2[i]) \quad (2)$$

where v1 and v2 are arrays of 2D vectors and n is the length of the array. With the vector lengths (a, b) we are able to calculate the optimal angle:

$$\theta_{optimal} = arctan\frac{b}{a} \quad (3)$$

Finally, we can easily find the maximum similarity between these two sets of vectors and score how similar is the unknown drawn gesture to the gesture in a library template. The score is calculated with the formula:

$$s = \cos^{-1}(a \cdot \cos\theta_{optimal} + b \cdot \sin\theta_{optimal}) \quad (4)$$

For each gesture in the library we calculate the score of similarity to the unknown gesture and the gesture with the best score is returned as most similar. Since the gesture recognition does not have to be precise at all times we have also implemented a threshold of the score, which needs to be surpassed, otherwise the gesture will not be recognized. In the following figure we present the initial gestures we would like to put in our library. The threshold is currently set at the 85 % of the similarity for the gesture to be recognized.

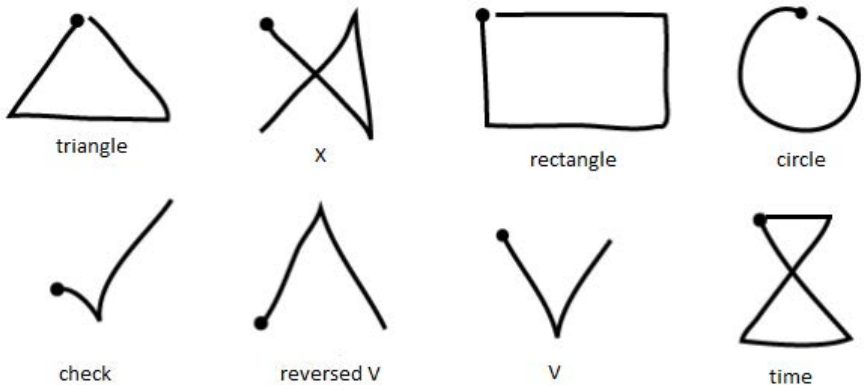


Figure 1. Gestures in the Library.

3.2 3D Gestures

The 3D gestures can be performed with both of the hands and are based on the movement of the fingers and palms. These gestures are set by us and cannot be changed in the game. We have performed several tests to determine the ideal controlling model for these 3D gestures.

Our algorithm determines whether player performs the gestures with the right hand or the left hand, and based on that outputs different actions. One of the gestures requires the player to extend all five fingers on a given hand, then hide all fingers and extend them again as can be seen in the figure 2. This casts a currently assigned spell from the given hand in the game. Each hand has a different spell assigned to it, which allows the players to get slightly different results with the same gestures on different hands.

Figure 2. Set of gestures required to perform the action of casting a spell.

3.3 Gamification

The game gains an output from the players, whether the gesture they draw is similar to the one the game wants them to draw, or not. To do that the players are asked to complete a survey in the game every time they stop playing, either after they die in the game or when they complete it.

Players often do not want to complete such surveys and are generally annoyed by them. Because of that a solution had to be designed to make them complete them, which would allow the gestures to evolve. In our case we have implemented a reward system that rewards the players with a virtual currency. Each survey a user even partially fills out, gives him several coins in exchange, which he can later use to improve his gaming experience in various ways. There is

a lot of ways on how to spend the currency in the game, for example it is possible to buy additional virtual lives, buy the option to re-spawn at the checkpoint where they die at, and unlock additional content. The additional content may vary from new spells they can use to the modification of the existing ones, but also to acquire new visuals for the enemies in the game.

To increase the amount of the players interested in the game and by that increasing the amount of gestures captured and recorded by these players, we had to come up with an interesting concept. We have designed a game which uses a magic system that depends on performing various gestures with the Leap Motion device to control the game. This concept makes it possible to be both intriguing for the players by trying something new, while also directly linked to evolve the gesture library.

We have compared the results gained by the users playing our game with the reward and without it and have determined a significant boost in interest. In the following figure we can see the results of several tests performed in our test session. We have performed a test on five people, each of them playing the game for 15 minutes straight. The first column showcases how many surveys have been filled across all these people, when the reward was not present. In the second part of the test, we have changed the surveys to include a reward for filling them and as can be seen in the figure, people have filled them a lot more often. The third column showcases how many of these gestures were correctly recognized, which is based on the survey, where people selected that the recognized gesture performed the action they desired. The last column showcases how many surveys were fully filled, after the person has decided to start filling them.

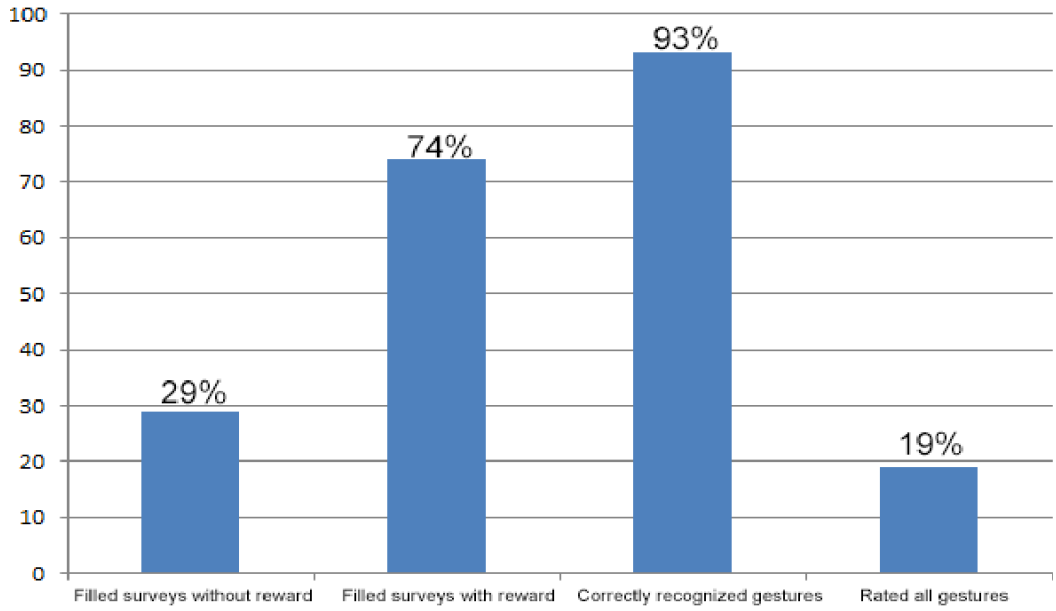


Figure 3. Results of several surveys.

4 Conclusions

In this paper we analyze the use of alternative controls in various applications used to provide an intuitive Human Computer Interaction. We have designed a library that contains various 2D gestures recognized by our algorithm, while the library also contains several 3D gestures. The library evolves the 2D gestures the more it is being taught by the players in the game. To make players help us improve the gestures in the game we introduced several gamification mechanics that reward player for completing short surveys.

We have tested the implemented solution on several users to determine the effectiveness of our solution. The test targeted people with experience in playing games, with various levels of the experience across different genders. Based on the results of this test we can now say that the gesture library recognizes the gestures correctly and users are able to successfully extend and evolve the existing set of gestures in the library. In the process of testing the game we have also encountered a problem with learnability of the game by the players, due to its rather unorthodox controls with the Leap Motion device, which requires us to implement various tutorials for the various mechanics in the game in the future.

In the future work we would also like to focus more on the gestures of the library and improve the game by adding more content. We would also like to let the player to experiment with more types of the gestures, and that is why we plan to extend the system to allow the player to add any gesture directly in the game. There is also a possibility of further experimenting with the controls of the game by introducing a virtual reality device, Oculus Rift.

References

[1] Dutta, T.: *Evaluation of the Kinect sensor for 3D kinematic measurement in the workplace*. Applied Ergonomics. Article. July 2012.

[2] McNeill, D.: *Language and Gesture*. Volume 2 of Language Culture and Cognition. Cambridge University Press. 2000. ISBN 9780521777612.

[3] Nielsen, M., Storring, M., Moeslund, T., Granun, E.: *A Procedure for Developing Intuitive and Ergonomic Gesture Interfaces for Man-Machine Interaction*. Technical Report CVMT 03-01. Aalborg University, Aalborg, Denmark. March 2003.

[4] Pereira, A., Wachs, J., Park, K., Rempel, D.: *A User-Developed 3-D Hand Gesture Set for Human-Computer Interaction*. University of California, Berkley. Department of Bioengineering. 2014.

[5] Potter, L., Araullo, J., Carter, L.: *The Leap Motion controller: A view on sign language*. Proceedings of the 25th Australian Computer-Human Interaction Conference. ISBN 9781450325257.

[6] Reichert, F., Bachmann, D., Rudak, B. Fisseler, D.: *Analysis of the Accuracy and Robustness of the Leap Motion Controller*. Article Sensors (Basel). May 2013. DOI 10.3390/s130506380.

[7] Tashev, I.: *Recent advances in human-machine interfaces for gaming and entertainment*. Information Technology Security. Volume 3. 2011.

[8] Leap Motion Developer Website. [Online; accessed Feburary 18, 2015] Available at: https://developer.leapmotion.com

IPv6 Security

Matej RYBÁR*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
mat.rybar@gmail.com

Abstract. Project focuses on a new generation of internet protocol IPv6 with a detailed orientation on an analysis and practical security testing in order to refer to its limitations. We identify security problems and selected vulnerabilities, threats and attacks are analysed in detail. The feasible tools for implementation of attacks under IPv6 network are discussed as well. Key element is to implement analysed attacks and based on the results of the tested attacks indicate vulnerabilities and security threats of the IPv6 protocol. The main goal is to create security testing system, which proposes the possibility of preventing security problems.

1 Introduction

Internet protocol IPv6 should become a new generation internet protocol. This protocol is the clash of many conflicting views on its safety and functionality. Decreasing number of available IPv4 addresses is the main source of discussions about its implementation. The main reason for delaying the implementation was transition to routing based on network masks and mechanism NAT. *IPv6 has many security risks known from Internet Protocol IPv4.*

IPv6 in compare to IPv4 uses extension headers used for packet additional identification, fragmentation is used at least. IPv6 uses multicast addresses instead of broadcast addresses and ICMPv6 protocol plays an important role in sending important information messages in the network. Taking into account these differences we can identify other new security threats.

With the increasing number of attacks many companies care about implementation of security systems. This was the motivation for creation of the *security testing system*. The system will involve implemented selected types of attacks and will be able to advise a security mechanism to eliminate the threat. Existing systems are not in sufficient numbers are expensive and parallel to that not suitable for a wider group of users.

2 Internet protocol version 6

The Internet has become the inherent part of our work and personal life. This fact has caused the exhaustion of IPv4 addresses. Therefore, a new Internet protocol IPv6 was created.

---

* Master degree study programme in field: Computer Engineering
Supervisor: Assoc. Professor Ladislav Hudec, Institute of Applied Informatics, Faculty of Informatics and Information Technologies STU in Bratislava

While IPv4 contain 32-bit addresses and 4.2 billion addressable nodes, IPv6 has 128-bit addresses and is able to address *3.4 * 10^{38} nodes*. This address space cannot be exhausted [1].
Main features and characteristics of IPv6:

- MTU size is at least 1280 B
- Uniform addressing scheme for internal networks and the Internet
- Using extension headings mechanism
- Fragmentation works only at the end nodes of the network
- IPv6 address type: x: x: x: x: x: x: x: x, where x are the16-bit hexadecimal fields
- Each address has its own range, period of validity and type
- Stateless Address Autoconfiguration
- The main types of addresses are: multicast, unicast and anycast
- Support jumbogram packet – size to 232 B
- Checksum is not necessary, IPsec security mechanism, support of mobility (portable devices)

The main feature of the IPv6 is that the standard header contains only the most essential fields, and therefore its size is fixed at a *constant value – 40 B*. Any other additional or optional data are transferred to the *extension headers*.

Since the IPv6 address space is very large, the solution has been sought to derive the host part of the address. A good solution is to use an already existing link-layer address, MAC address. The result of the mapping 48 bit MAC address on a 64 bit host part of an IPv6 address is the address of the end device. The algorithm for the creation of these addresses is called *IPv6 EUI-64*.

ICMPv6 represents informative protocol. It announces error conditions, tests reachability and is used to transmit communications information. ICMPv6 has very important role in the Internet Protocol IPv6. It serves as an information protocol for almost all mechanisms such as discovering neighbors, multicast addressing, information about nodes, mobility etc. ICMPv6 messages include these fields [2]:

- *Type* – specifies the basic type of ICMP messages divided into two classes: Error and Information,
- *Code* – sub-base type message.

For communication with another computer is important to know the *IP addresses* and the corresponding *MAC address*. *Neighbor Discovery Protocol* is used in the IPv6 protocol for search link – layer address PC on the other side of the communication channel. There are many other ways for use such as quick update invalid entries, search routers in the network, forwarding, verifying reachability neighboring routers, duplicate address detection, identification data for automatic configuration of addresses etc. It uses two types of ICMP messages: *Neighbor Solicitation and Neighbor Advertisement* [3].

The mechanism of automatic configuration of addresses can be used at the end nodes of the network. We know two types:

- *Stateless Address Autoconfiguration (SLAAC)* – by random intervals all nodes in the network segment are notified about the network prefix, and also about which router should be used when routing to the internet. It uses two types of messages: ICMPv6 Router Advertisement and Router Solicitation [4].
- *Stateful Address Autoconfiguration* – an additional server, the same mechanism as DHCP.

3 Security threats

In computer security the threat is a possible danger that might exploit a vulnerability to breach security and thus cause possible harm.

The first step for an attacker is to detect important information about the target network. These attacks belong to the category of *horizontal scanning network*. Attacker can use detected information as an advantage and successfully carry out an attack on the target device. Implementation of security mechanisms that protect the network as part of the initial scanning, can avoid an attacker to find the necessary information.

3.1 Horizontal scanning network

Scanning network is conducted in order to determine the topology and active nodes in the network. Whereas in the IPv6 network is not able to provide a brute force attack, in the process of identifying information the following procedures could be chosen:

- Finding a live IPv6 addresses on the local network
- Finding information on the routers
- Detection and identification of new IPv6 addresses on the local network
- DNS AAAA records
- Whois database

3.2 Types of threats

In case that some of these attacks will be successfully fulfilled, it may cause serious damage to a network or its decommissioning.

- *Spoofing records in the cache memory of neighboring devices* – each node in the network maintains pair of IPv6 and MAC address in cache memory. The aim is to persuade the victim (target PC) that just the MAC address of a computer belonging to asked IPv6 address.
- *Spoofing router advertisement messages* – If the attacker wants to impersonate a router, it is necessary to falsify RA messages. In this case, the attacker as a gateway must ensure that the communication passed through him and he acted transparently. This type of attack is called MITM (Man-in-the-middle) [5].
- *Denial of Service by detecting duplicates* – Ipv6 allows find the same address in the network through the mechanism of Duplicate Address Detection (DAD). The attacker will respond NA messages, which means that the address is already used. Then the interface cannot be assigned IPv6 address [6].
- *Overloading neighbor's table* – neighbor's table contains a pair of IPv6 addresses and MAC. Because of the large address space, an attacker can fill the table neighbors.
- *Attacks on fragmentation* – rearrangement fragments, indefinite presence information of higher protocol, sending a large amount of fragments, lack of end fragment, overlapping fragments, etc. [7].
- *Denial of Service attacks* – flooding ICMPv6 messages (Router Advertisement, Router Solicitation, Neighbor Solicitation, Neighbor Advertisement), Teardrop attack, Peer – to peer attack, Distributed DoS, Smurf attack.

4 Security mechanisms

The main security solution currently remains *IPSec protocol*, which provides two basic services:

- *Authentication* – guarantees connectionless integrity and data origin authentication of IP packets. What is more, it also verifies that the content of the datagram was not changed by the transmission via network.
- *Encryption* – encryption of transmitted content datagram.

The whole process of encryption and authentication requires additional support mechanisms such as security association, regular exchange of keys or management certification paths. IPSec can operate in two modes – *transport mode and tunnel mode* [8].

IPSec protocol has a number of features that significantly reduce or prevent the following attacks: *Sniffers,* data modification, identity spoofing, man-in-the-middle attacks, denial-of-service attacks.

Another mechanism is *SeND (Secure Neighbor Discovery)*, which is actually in testing phase and not used in implementations routers. It uses cryptographically generated address (CGA), digital signatures and certification routers to secure the network. It is expected to deploy to the new generation of IPv6 networks.

5 Security testing system

To create a comprehensive system for security testing we will use described attacks and security mechanisms. These attacks will be called tests that allow us to test the target network. Using the open source existing package *THC IPv6 attack toolkit* tests will be implemented to the system. The system will be designed as a Linux application using the web interface and Unix scripting. The target group of the security testing system will be wider layer of users, and therefore we will choose a simple, clear and intuitive interface. The facility will include not only a set of tests, but also will help in the appropriate security solutions. The system will propose solutions to identified safety threats and guidance to improve safety. It will be available via the Linux Live CD, which will be executable on any PC. System testing will be conducted on real networks in a laboratory environment.

At the time of writing this article the first part of system was created – the *implementation of tests on the selected network*. In the second phase the *user interface* and a proposal for possible security on the basis of the acquired test results will be created. The tests were carried out on an experimental network in the simulation environment with using programs GNS3 and Virtual Box.

5.1 Scanning IPv6 address

The possibility of obtaining important information by an attacker is tested. Scanning IPv6 address is classified to the category of *horizontal scanning network*. The system of this test *detects active IPv6 addresses* on the local network. The basis is to generate the ICMPv6 message including echo request and echo reply (ping) for a detection of live interface. ICMPv6 message uses multicast FF02::1 as a destination address. We see the system output in *figure 1*, where *p7p1* is an identifier of our computer interface. We are connected to target router by this interface.

```
[root@localhost thc-ipv6-2.5]# ./alive6 p7p1
Alive: 1111:1111:1111:1111:c200:11ff:fe14:0 [ICMP echo-reply]

Scanned 1 address and found 1 system alive
```

Figure 1. Scanning IPv6 address.

5.2 Information about router

The system detects the availability of information of the router. This test belongs to the category of *horizontal scanning network.* It controls *IPv6 address* of the interface that is connected to our PC, *MAC address, router priority, maximum number of hops, MTU* and also a *net prefix* for the

network in which it is located. The aim is to reduce or completely eliminate the availability of this information for attacker.

5.3 Detection of new IPv6 address

The security testing system is also able to *detect a new device* added to the network or a reconfiguration of existing IPv6 address (see Figure 2). Information about new IPv6 address is located in ICMPv6 message "*Neighbor Advertisement*" in the Neighbor Discovery protocol.

```
[root@localhost thc-ipv6-2.5]# ./detect-new-ip6 p7p1
Started ICMP6 DAD detection (Press Control-C to end) ...
Detected new ip6 address: 1111:1111:1111:1111:c200:11ff:fe14:1
Detected new ip6 address: 1111:1111:1111:1111:c200:11ff:fe14:2
Detected new ip6 address: 1111:1111:1111:1111:c200:11ff:fe14:3
Detected new ip6 address: 1111:1111:1111:1111:c200:11ff:fe14:4
Detected new ip6 address: 1111:1111:1111:1111:c200:11ff:fe14:5
```

Figure 2. Detection of new IPv6 address.

5.4 Duplicate Address Detection

In case of using autoconfiguration there is a high possibility that more stations will appear with the same address in the IPv6 network. To avoid this situation, the mechanism for *Duplicate Address Detection* is used. The principle of detection is that the terminal station sends a message *Neighbor Solicitation* with the same destination address which intends to assign. If it gets the message *Neighbor Advertisement*, it means that the address is not available and it could cause a conflict. If it doesn't get the message, the address is free and can assign it. The security testing system generates false messages and pretends an existence of IPv6 address. The system can recognize that the attacker knows how to use this vulnerability in the network. This test is classified as a *Denial of Service by detecting duplicates.*

5.5 Flooding attacks

Flooding attacks are based on generating of multiple messages flooding the destination node in the network. These attacks belong to the category of *Denial of Service attacks.* This device will be more overloaded and it leads to slowing-down the performance. Sometimes it can result to complete elimination of device from running.

- *Overloading neighbor's table* – In the testing topology the system found out the router overloading on 50–60 % and slowing-down the performance. The testing system generated 37 764 packets for 11 seconds. The processor overloading increased by 23 %.
- *Flooding Neighbor Advertisement messages* – the system generated 218 299 ICMPv6 packets of type NA for 67 seconds. The overloading of the processor increased by 14 % without sensible effect.
- *Flooding Router Advertisement by messages* – are generated by every router in the network for the reporting of net prefix. The system generated 105 000 messages for 30 seconds. The router was overloaded on 100% and it was not possible to continue to work with that.
- *Flooding ICMPv6 Echo Request messages* – flooding the target stations by requests for a ping. The test may cause an accumulation of processor overloading of the target station and also may cause slowing-down.

The security testing system carries out all described attacks and displays ascertained results. This tool also suggests possible solutions. On the basis of results a user may take steps to eliminate vulnerabilities and improve security of the system.

6 Conclusions and future work

The project purpose is to point out to the vulnerability of the protocol of new generation and enable to test networks for the general public. For the implementation of tests the selected tool THC IPv6 Attack Toolkit was chosen and it seems the most suitable. After verifying the results of the test functionalities attention should be paid onto the creation of appropriate user interface with a possibility to propose an elimination of vulnerabilities that were found. By that the environment of testing should become more accessible for more people and it would be easier for drawing conclusions of individual attacks.

The system of the real network in laboratory environment will be tested. The test results namely showed that without additional security of IPv6 network the vulnerability is very high and therefore it is not appropriate to implement unsecured network into real environment. The system generating test results along with the proposal and instruction for elimination of security threats would become a complex tool for improving the security of the IPv6 network.

Acknowledgement: This work was partially supported by research grant VEGA 1/0361/15 entitled "Security in distributed computer systems".

References

[1] *CZ.NIC*: *Seriál Pohneme s IPv6*. [Online; accessed 2011]. Available at: http://www.lupa.cz/serialy/pohneme-s-ipv6/

[2] RFC 4443: *ICMPv6 for IPv6*. IETF Working Group. [Online; accessed 2006]. Available at: http://tools.ietf.org/html/rfc4443

[3] RFC 4861: *Neighbor Discovery for IPv6*. IETF Working Group. [Online; accessed 2007]. Available at: http://tools.ietf.org/html/rfc4861

[4] RFC 4862: *IPv6 Stateless Address Autoconfiguration*. IETF Working Group. [Online; accessed 2007]. Available at: http://tools.ietf.org/html/rfc4862

[5] SANS Institute: *A Complete Guide on IPv6 Attack and defense*. [Online; accessed 2011]. Available at: http://www.sans.org/reading-room/whitepapers/detection/complete-guide-ipv6-attack-defense-33904

[6] RFC 3756: *IPv6 Neighbour Discovery Trust Models and Threats*. IETF Working Group. [Online; accessed 2004]. Available at: http://www.ietf.org/rfc/rfc3756.txt

[7] HOGG: *IPv6 Security*. Cisco Press. No. 5133 [Online; accessed 2009]. Available at: http://dark.ellende.eu/public/Cisco.Press.IPv6.Security.2009.pdf

[8] RFC 4301: *Security Architecture for the Internet Protocol*. IETF Working Group. [Online; accessed 2005]. Available at: http://tools.ietf.org/html/rfc4301

Efficient Processing of Data Transport on Mobile Devices

Miroslav MOLNÁR*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
xmolnarm7@stuba.sk

Extended Abstract

Using of public transport brings reduction of environmental pollution, more effective transport, danger reduction of human life and reduction of travel expenses. Disadvantage of public transport is its complexity. City public transport often consists of different types of vehicles and number of connections. Transport organization can be changed according to the changing conditions. Understanding the transport connections is difficult. There are many applications planning public transport routes. A lot of applications are developed for mobile devices because there is much higher usage and flexibility during the traveling.

There are several structures of recording the traffic data. We decided to focus on The General Transit Feed Specification (GTFS). Advantages of this data structure are good readability, possibility to edit it by users and especially its expandability [1]. The disadvantage is a big requirement for the storage space. Mass transport data of a large city can have hundreds of megabytes. These data are not suitable to process on mobile devices. Our goal is to change the existing structure and to verify the efficiency on mobile devices.

Although, the hardware equipment of mobile devices has undergone much development we are still limited by the capacity of resource, availability of the storage space, by the memory and by the processing power. These aspects should be taken into account by developing the mobile application.

Current applications solve this topic different. One of the most used approaches is client – server architecture. This architecture has many advantages such as lower requirement for storage space and processing power on mobile device. This approach is used for example by the application iTransit [2]. A big deficiency is necessity of client connection to the server. The connection can be interrupted because of many reasons. In a case the service is unavailable the mobile application with the application server is unusable. The second approach is to save data in relational databases. It is used for example by the application Grand River Transit bus system [3]. This approach is easy, it offers a great functionality to work with data and little improvement of memory resources.

In this work we focused on two ways of data transformation, transformation into relational database system and transformation into binary form. Transformation into the database is easier. The brand new data have lower requirement for storage space but higher than by the

* Master degree study programme in field: Information Systems
Supervisor: Assoc. Professor Michal Čerňanský, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 385–386.

transformation into binary form. Available SQL systems offer technologies for quick work with data. Size of GTFS data is by the larger cities too big and searching the transport connection is inefficient.

The transformation into binary form is more difficult to be implemented. It brings better reduction of storage requirements. This transformation brings bigger options of the structure editing. It is possible to connect data information into the structures and to use the binary operations effectively [4].

GTFS structure consists of text files written in the form of CSV tables (Comma-separated values). GTFS structure does not strictly define many values. Often, the same data from different transport companies can be written different or values are entered ineffective. To change the data structure is necessary to know the data and their values. For this purpose, GTFSParser application was proposed which is able not only to find the information about the data but as well to provide the function for working with data and to propose a data type for the group data. The application picks one column and searches all the data in it. Then it returns the data type which suits all data in the column.

The main problem of GTFS structure is inefficient data entry. The data are represented as text chains. Work with these kinds of data is very inefficient. We have proposed a change of the data structure according to the steps: Data normalization, Transformation of text representation into binary form, Change of the structure for quick data searching.

Data normalization of GTFS structure consists of files entered in tabular structures. Each GTFS file represents one GTFS table. Tables contain large data redundancy and therefore the normalization is necessary.

Transformation of text representation into binary form reduces the size of requirement for storage space and speeds up the work with data. All data are able to be transformed efficiently in another way. The best results are achieved by the number or sign values. The worst results are by the transformation of text values when the size of resulting data does not decrease.

Adjustment of the structure for quick searching over the data is very important. Searching transport connections is difficult and also after the adjustment of GTFS after normalization and transformation into binary form is very inefficient for using in mobile devices. Various algorithms can use different data in different orders, and intensity of use. Also the data structure determines the type of processing. Therefore it is important to identify not only the intensity of work with data but as well the order of loading and way of processing. Frequently used data are separated from the less frequently used. The data processed in a certain order are arranged to avoid re-reading the whole data structures. The structure is divided into smaller data structures to make the work much efficient.

The aim of further research is to verify the correctness of the proposed structure for a mobile application to use GTFS data searching connection. In another research it would be good to explore the options of graph databases which for the best searching way in the graph achieve interesting results.

References

[1] Google Developers: Transit, (2012), https://developers.google.com/transit/gtfs/

[2] Už nikdy nezmeškaj spoj!, http://www.itransit.apptives.com/vlastnosti.php.

[3] Brenna, W., GTFSOffline, (2014), https://github.com/wbrenna/GTFSOffline

[4] Szincsak, T., Vagner, A., Data Structure to Store GTFS Data Efficiently on Mobile Devices, Journal Of Computer Science And Software Application, Volume 1, Number 1, (2014)

ReTSys – Real-Time Tracking System

Ján ONDER, Tomáš TRÁVNIČEK*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
Funteam2@googlegroups.com

Extended abstract

Many companies have a lot of cars or trucks which are difficult to keep track of and maintain. It is also hard and time consuming to make logbooks manually for them. Logbooks are nowadays created after every ride even when it is just in one town. After certain period (usually a month), they have to be rewritten into digital form so this data can be used for tax purposes.

Our goal is to make this process easier, more enjoyable, accurate and better for employer than it is now. With series of interesting statistics it is possible to lower company expenses on vehicles, because drivers will be rated according to their driving style and also their average fuel consumption.

All this is possible with new complex devices, which provide very accurate position information from GPS satellites and a wide range of telemetry data from vehicle (through OBD-II CAN bus, which is common in most of modern vehicles).

In Figure 1 below you can see the architecture of our system. It consists of three main parts. The basic part is composed of various tracking devices of which some are personal and some can be connected to vehicle.

Some of the devices for vehicles can connect to the car's OBDII port and get information from its CAN bus. The other devices have digital input and output ports which can be utilized for various uses. For example to connect an SOS button or DALAS key reader for authenticating the driver. Then there are devices for long term monitoring of valuables. Their battery can last, depending on settings, for more than a year.

All of these devices send GPS and other information to the second part of our architecture and that is the parsing program located on our server. This program processes the data and saves it to database for later use. If the device sends specific message like some alarm then the parsing program can generate a warning email or SMS message and send it to specific user.

The third part is the main Web application located on our web server. This application takes the data saved in database by the parsing program and shows it to authenticated users.

With this architecture, we would also like to introduce new services, like automatic detection of theft or tow. Also very important feature we are focusing on is automatic reading and notifying responsible person about errors and faults through CAN bus [2] of vehicles. Monitoring of personnel or family members, like children. With series of interesting statistics it is also possible to lower company expenses on vehicles, because drivers will be rated according to their driving style and also their average fuel consumption.

---

* Master degree study programme in field:Computer Engineering
Supervisor: Dr. Peter Pištek, Institute of Computer Systems and Networks, Faculty of Informatics and Information Technologies STU in Bratislava

For best user experience and comfort it is appreciated to be able to do at least basic functions to be viewed by mobile phone or other smart pocket device. Also our system allows identifying driver of vehicle based on his personal chip, called Dallas key, so the only thing he or she would need to fill to the logbook is purpose of travel.

For person responsible for maintenance the company fleet, we provide an ability to configure tracking devices inside vehicles directly through SMS, or by specialized form on portal.

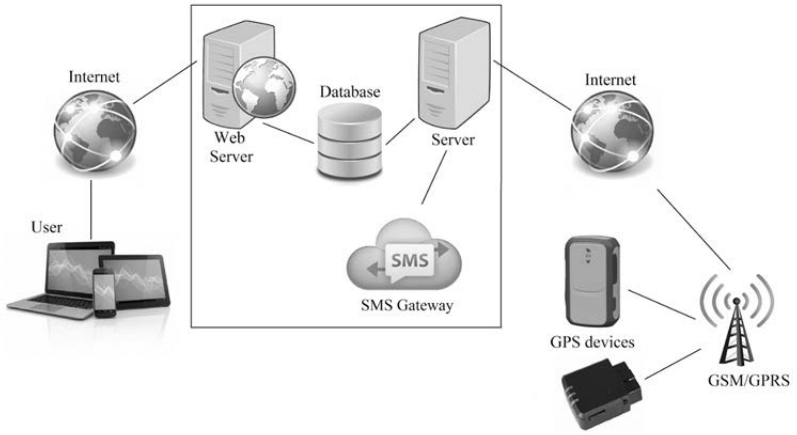


Figure 1. ReTSys system architecture.

Proposed solution with all functionalities will be implemented in form of ReTSys tracking system. The main user interaction with all of the services, besides installing the hardware in vehicle, or other asset will be via our web portal. Through it, users can easily view current state and position of vehicles, check the statistics, fill or lock logbook for current month, configure devices or manage personal and company settings, based on their user roles. For example, administrator of fleet will be able to check every vehicle on the map, accountant to fill or lock current month's logbook, but not check current position of vehicles. Common users – drivers, couriers, will only be able to check tracks, or fill logbook entries of vehicles, they have permission to drive.

We are also focusing on possibilities of implementing exact hierarchy of every company, so that whole company can be divided to branches, branches to sections and so on. We developed a complex tracking system that contains both programming and hardware components. Our system compared to our competitors contain more objects of tracking. For example, asset tracking for long term tracking, or vehicle tracking scaling from few vehicles to whole fleets of them. Our webpage is also optimized for mobile devices so you can carry the provided information with you. We also provide automatic logbook creation for company cars. A company can create an infinite number of branches. We are also open to adjusting our product to suit the needs of other companies.

All of our solutions are widely tested across Europe and we are still incorporating the feedback into the final product. From this testing we gathered 1.3 million position messages so far.

Acknowledgement: This work was partially supported by the Slovak Science Grant Agency (VEGA 1/1008/12 "Optimization of low-power design of digital and mixed integrated systems").

References

[1] Národná rada SR: *595/2003 Z. z. o dani z príjmov.* Zbierka zákonov, SR, (2003).

[2] ESAcademy: *CAN bus*. [Online; accessed February 15, 2015]. Available at: http://www.canbus.us/.

Multimedia Systems for Taxi Services

Martin POLÁK*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
xpolakm5@stuba.sk

Extended Abstract

Taxi services are common type of transportation and they have not changed a lot in last years. In some countries there are problems with overpriced taxi drives just because driver used longer road or because he used device which will measure longer distance travelled than the distance really was.

Related systems are usually just for playing commercials or just for watching TV. For example one of the best systems is sold in India and it is called Tabbie [1]. There is possible to play music, watch movies and also watch map where the taxi is currently located. But there is no possibility to see whole route so customer does not know if the driver is driving by optimal route or if the distance for which customer is paying is no longer than the actual travelled distance.

This system (TickTack taxi) is here to solve problems when user is inside vehicle with the use of GPS and Google Maps inside tablet just for customer. As an additional features system will be able to provide Internet access for passengers, play multimedia files – movies, television and music, which is possible to play directly into car audio system.

The system consists of two tablets (customers and drivers) and they communicate to each other. For this communication is used wireless network created by MiFi router with direct connection to the Internet by 3G mobile network. TCP protocol is used for sending commands for activation, deactivation and sending ID of the commercial. The use of UDP protocol is in position sending, because if sometimes packets get lost it is not a serious problem.

Every driver has its own "controlling" tablet, which can send commands to "customer" tablet and activate it, deactivate it, send which commercial should play and send location data periodically, because "customer" tablet does not have built-in GPS module. When driver activates "customer" tablet, the data packet is send from "controlling" tablet. Data packet contains driver's name, locations of starting and end position and also ID of first commercial, which will be played right after activation of tablet. When the "customer" tablet is deactivated, customer cannot do anything with it – the application is disabled.

This system is based on more applications which can be used by customer – there is one main application which can run the others (customized launcher). Each multimedia application (television, movies, music and Internet) has its own list application where are multimedia files listed.

Multimedia data inside tablet are possible to synchronize by Wi-Fi or by USB flash disk. Wireless synchronization is used for easier update of data (commercials, movies, music, settings

* Master degree study programme in field: Computer Engineering
Supervisor: Dr. Peter Pištek, Institute of Computer Systems and Networks, Faculty of Informatics and Information Technologies STU in Bratislava

etc.) stored in a tablet. Wireless connection makes possible to synchronize more vehicles at once by connecting to one access point. USB flash disc synchronization has to be done manually inside vehicles. First installation has to be done with USB flash disc.

Main application is first application that customer can see – right after commercial is played (commercials are optional and it is on driver if it is played or not). After this part he or she can watch his or her actual travelled route and optimal route on small map, see daily newsletter downloaded by RSS and run other applications for entertainment. Also on the top there are additional information about driver's name, names of streets (start and end of the route), estimated length of route according to Google Maps API and actual travelled length of route so far.

Figure 1. Graphics interface of main application.

TickTack taxi solution is deployed on approximately 150 vehicles of taxi service in Prague. Every one of these vehicles is equipped with the same set of applications and the same configuration. If any crash happens, it appears on Crashlytics site. This testing is made by real users in real world hundreds of times a day.

This solution is unique in way that it contains combination of competitors' features with addition of tracking travelled route with comparison of optimal route according to Google Maps API. This way customers will not be lied to and can be sure that price is right and they are paying just for route they actually travelled.

In our best knowledge, there is no other system with describer features for cabs. In comparison with similar solutions it is possible to see this approach offers the best functionality for customers.

Acknowledgement: This work was partially supported by the Slovak Science Grant Agency (VEGA 1/1008/12 "Optimization of low-power design of digital and mixed integrated systems").

References

[1] Luminosity Technologies: *Tabbie (Tablet in a Cab)*. [Online; accessed March 18, 2015]. Available at: http://luminositydigital.com/products.html

Application of Funtoro MOD Embedded System as Tourist Guide

Tomáš ŠKRIEČKA*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
tomas.skriecka@gmail.com

Extended Abstract

Convergence and advance in technology affects every area in nowadays. One of them is the automotive industry. Consumers, in this case the passengers have constantly increasing demands for comfort, information as well as entertainment while traveling. This is because it is trendy for the passengers to be constantly online and they want to share their travel experience. The trips are generally longer so the entertainment is also important. Part of this problem is the fact that half a year in the development of applications for smartphones is eternity, but in the development of the automotive industry it is a blink of an eye.

Some technical innovations in the automotive industry increase comfort and safety (such as smart seat [1] or traction and electronic stability control), while others information about the current state of the vehicle or arrange for entertaining passengers. Increasingly in this area begins to expand the concept of "infotainment" systems, i.e. systems combining information and entertainment. These systems include a system of "Media on Demand", which is increasingly applied in aircraft, buses or trains. Similar solutions mainly focus on the displaying current position on the map but omit the provision of information about surrounding area based on GPS position. Provision of information about surrounding area is very interesting from the carriers view because it enables giving more complex information and they can use this part as the opening of advertisement segment. Our goal is to create innovative multimedia applications for this system and thereby contribute to the comfort, information and entertainment for passengers.

Our aim is to offer solution that improves the complexity of provided information, which is achievable with Funtoro MOD system [2]. It is most common and most advanced system in this automotive segment. Our tourist guide solution is divided into two parts. First part is Android application for Funtoro MOD system with automatic/manual playing sound/video records when passing near certain points of interest (POI) or displaying relevant web links. The second part is tool for searching and selecting POI directly from the map and generating necessary XML files.

Main application will be using three different information sources:

- Internet for communication with Google servers and downloading map tiles using Google Maps API v2 [3],

* Master degree study programme in field: Computer Engineering
Supervisor: Dr. Peter Pištek, Institute of Computer Systems and Networks, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 391–392.

- GPS satellites for coordinates and other information from NMEA messages,
- IP cameras for showing real-time video from outside or inside of vehicle.

For better providing of comfort and information for users our application was designed into two modes: drive mode and watch mode.

In the drive mode the MOD server connects to Google servers to download and display map. It shows the position of the vehicle on map using GPS and also shows points of interest loaded from XML file. If during the movement the vehicle is within range of certain point of interest, the media file is played either automatically or after manual choice (based on settings). User interface of application besides actual position also shows time, speed and altitude.

In the watch mode user selects from available camera sources and appropriate camera output is displayed in real time.

Concurrent solutions omit provision of information about surrounding area and other aspects of complex information awareness so our approach with its open options for further carrier's business development has the potential to be unique in this area. Our approach could also be assumed as universal for every type of transport after modularity consideration (camera mode may not be always possible or desirable).

In this paper we presented our solution of "media on demand" system for buses using infotainment methods, where we focused on tourist guide application with combination of camera system of the vehicle. In a further development we have to ensure proper trigger of very close points of interest and we also must appropriately and correctly handle transitions between the two application modes (drive, watch).

Acknowledgement: This work was partially supported by the Slovak Science Grant Agency (VEGA 1/1008/12 "Optimization of low-power design of digital and mixed integrated systems").

References

[1] European Commission: Smart Technologies for stress free AiR Travel. [Online; accessed February 16, 2015]. Available at: http://ec.europa.eu/research/ transport/projects/items/ seat_en.htm

[2] Funtoro INC: HIGH-END solution – Digital MOD server. [Online; accessed May 5, 2014]. Available at: http://www.funtoroeurope.com/sk/riesenia/c13/autobusy-vlaky-lode.html

[3] Google INC: Google Maps Android API v2. [Online; accessed May 5, 2014]. Available at: https://developers.google.com/maps/documentation/android/

Using Parallel Web Browsing Patterns on Adaptive Web

Martin TOMA*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
martin.toma.svk@gmail.com

Extended Abstract

The possibility to use browser tabs as a tool for parallel web browsing is definitely not new. In recent years, however, more and more people tend to use this feature every day. Despite that, little research has been done to deeply analyze why, and more importantly how, people use this browsing mechanism. Even less of it is aimed to utilize this information in a way, which would further enhance the browsing experience for the Web browser users themselves.

Parallel browsing occurs when the browser user is using browser mechanisms that allow him or her to open and browse multiple Web pages at the same time [3]. Browser windows and tabs are examples of such mechanisms. Users have many reasons why they prefer to use tabs and why they practice parallel browsing in general, some of those reasons were mentioned in a study performed on users of Mozilla browser [1]: short-term, visual bookmark; parallel searching (branching from search result page); opening interesting link on background without interrupting the current process.

Users have decided to use tabs as the best tool to perform parallel browsing. According to research from 2010 [2], 57 % of all browsing sessions involve some type of parallel browsing and users are multitasking by splitting their browsing activity into different tabs rather than viewing more pages overall. The same research also expressed the belief that this type of behavior has been growing recently and we foresee this to continue gaining popularity. Research also shows that studies focused on this type of navigation behavior have been lacking in Web and hypertext communities, despite the fact that parallel browsing covers fundamental interactions with hyper-links.

Our first goal was to design and implement a method, which will be used to collect accurate browser usage data with the primary focus on browser tabs. Existing solutions like Brumo browser extension framework[1] lack the accuracy in some cases, because they are not utilizing the modern browsers APIs to the full extent. Collecting accurate data will lead to more accurate parallel web browsing analysis. We have implemented a Web browser extension called TabRec[2]. While using

---

* Master study programme in field: Information Systems
Supervisor: Martin Labaj, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

[1] http://brumo.fiit.stuba.sk

[2] http://tabber.fiit.stuba.sk

the improved Chrome API we are able to confidently detect eight different tab events. TabRec is already live and is capturing parallel browsing activity produced by about 25 users.

We have already collected almost half million of usage logs. Utilization of the most common patterns via the browser actions recommendation was our second goal. There are many techniques dealing with sequential data analysis and pattern extraction. To be able to use some of these techniques, we firstly have to deal with data preprocessing. We group several specific sequences into more meaningful items representing concrete user actions. This not only allow us to use those techniques and algorithms, which normally ignores duplicate items in transactions, but also express the real user intention much clearly. After preprocessing, we use Generalized Sequential Pattern algorithm to find the most frequent sequences. We also consult the results with our custom sequence extraction process to help us identify the most common patterns more confidently.

The most commonly shared sequences can be considered trivial and consist of repeating subsequences each in form of one or two events. Tab close and activate or create and update are examples of such trivial sequences. One of the most common non-trivial sequence we found, consists of five events. User firstly creates a new tab, then focuses it and navigate to a certain web address. This is actually a sequence of three events (create, activate, update). Sequence continues by switching to another, already opened tab and navigating there to another web address. This might represent a typical parallel browsing activity. User is keeping both tabs and regularly updates their content. Another example starts with user focusing an existing tab, then closing it and focusing another two tabs. This sequence continues by closing current tab and focusing another one. Because the time gap between these actions is only five seconds, we can imply that the purpose of this particular sequence was to find those tabs, user don't need anymore.

We will now utilize some of these most common patterns and sequences we discovered to enhance users browsing experience. This can be achieved by recommendation of the most appropriate tab action for specific situation or by detecting some characteristic sequences, for example multiple activation events, which can signalize that users is probably searching for specific tab and then provide helpful actions, in this case rearranging tabs by Web domains. Currently we focus on identifying the characteristics of such situations and detecting the corresponding sequences in our data.

As a last step, TabRec extension will be enhanced to perform a real-time detection of selected sequences and will have the ability to notify users and execute the appropriate action. We will record the implicit feedback from accepting or rejecting these actions, which will be used as the main metric in our evaluation process.

Acknowledgement: This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG 1/0646/15.

References

[1] Dubroy, P., Balakrishnan, R.: A Study of Tabbed Browsing Among Mozilla Firefox Users. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '10, ACM, 2010, pp. 673–682.

[2] Huang, J., White, R.W.: Parallel Browsing Behavior on the Web. In: *Proceedings of the 21st ACM Conference on Hypertext and Hypermedia*. HT '10, ACM, 2010, pp. 13–18.

[3] Srivastava, J., Cooley, R., Deshpande, M., Tan, P.N.: Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data. *SIGKDD Explor. Newsl.*, 2000, vol. 1, no. 2, pp. 12–23.

Television Content Management

Dávid URBÁN*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
garfieldnagyur@gmail.com

Extended Abstract

Issue of the playing the multimedia content simultaneously and automated in the several television devices is not very widely and openly discussed. However, the current state of technologies offers a large amount of possibilities and methods, to design and to implement satisfactory up-to-date solutions for the television content management. The base of the success in this field is the perfect understanding of the UPnP protocol, which is spread in the smart TVs over the whole world. Main idea of the design solution is the network based communication between user and television devices. The primary benefits of the concept should be saving the user time spent by the multimedia content customization and centralized access to the system using a web user interface accessible via the Internet. Except the managing of the multimedia content, system should save, interpret and visualize statistical information, obtained by mapping of the network and the available devices.

Currently, copying the video by hubs, personal computers and multimedia players are the well-known solutions for the presenting advertising messages, playing of the multimedia content, lecturing or presenting at the various events or different places. From the perspective of time, these solutions are ineffective for managing, which is absolutely necessary to change the content.

Another solution, created by the ATEN Company[1] called ATEN Seamless Switch, is based on the hardware device *8x8 HDMI Matrix Switch with Scaler*. This solution allows centralized control over video sources and displays from a single chassis which accommodates up to 8 input and 8 output devices simultaneously. This hardware is built with a bi-directional RS-232 serial port that connects to a high-end system which can control the unit via serial commands [1].

Everybody wants to have a comfort while controlling their devices from the one point with the saving the time and the effort. Therefore aim of the work is to utilize the potential of the network improvements. Centralized access to the system saves the time. Web user interface accessible via Internet offers the comfort, and management of the devices and the multimedia content. Statistical information provides an overview of all system including TVs. All of these features should be the part of the satisfactory up-to-date design of the TV Content Management.

In achieving these goals will be necessarily helpful UPnP protocol, which creates peer-to-peer network connections between network devices. Establishing a connection between the devices does not require any configuration [2].

---

* Master degree study programme in field: Computer Engineering
Supervisor: Dr. Tomáš Kováčik, Institute of Computer Systems and Networks, Faculty of Informatics and Information Technologies STU in Bratislava

[1] www.aten.com

The core of the design system consists of a media server with the database. Television devices have to be connected to the media server via the local area network. Connection may be via an Ethernet cable for more bandwidth or wirelessly for the ability to place the televisions in a location where the Ethernet connection is unavailable. For the media sever content management and indirect control of television devices is designed a web server. Administrator always has access to the web server from the local area network. If the web server is allowed to use the public IP address, for the administrator will be available remote management of the system via the obtained public IP address. Figure 1 describes system architecture.

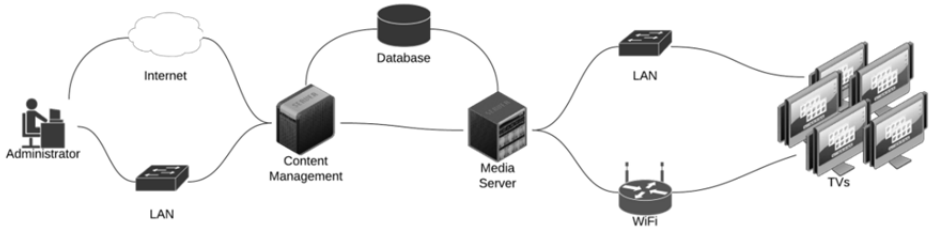


Figure 1. System architecture.

The media server consists of the storage for multimedia content, audio-visual content streamer and control unit. The server must be accessible to all network television devices, web server and database. The control unit is designed to manage the entire system on the basis of predefined tasks assigned by the user via a web interface. The role of management is based on the following points: (1) searching UPnP devices in the local area network, (2) on the basis of playlists stored in the database provide playing of the multimedia content in each device, (3) remote control of the devices, (4) monitoring system status and providing this information to the web user interface, (5) providing management of devices in the network via web user interface.

Web application is used for a user interaction with the designed system. These functions are available via this interface: (1) management of multimedia content, (2) playlist management, (3) management of television devices, (4) system reporting.

The communication between the web server and media server is provided by signal channel. Indeed, it is a network connection that uses standard TCP connection with the possibility of predefining ports.

In this paper is presented new design of the television content management, which is based on the UPnP protocol. An important advantage is that a lot of smart televisions natively support UPnP protocol. On other side, complexity of UPnP protocol, and identification of parameters and their values, which are necessary for the correct communication with television devices, were challenging task.

Acknowledgement: This work was partially supported by the Slovak Science Grant Agency (VEGA 1/1008/12 “Optimization of low-power design of digital and mixed integrated systems”).

References

[1] ATEN International Co., Ltd.: *ATEN Seamless Switch™ Delivers Real-time Program Video in a Movie Theater.* [Online; accessed February 20, 2015]. Available at: http://www.aten.com/data/CKUpload_files/files/Case-Studies/2014_CaseStudy_Entertain_Theater_VM5808H_Japan.pdf

[2] Presser, Alan et al.: *UPnP™ Device Architecture 1.1. UPnP Forum.* [Online; accessed November 28, 2013]. Available at: http://upnp.org/specs/arch/UPnP-arch-DeviceArchitecture-v1.1.pdf

Experimental Gateway between CAN and Ethernet Networks

Ján LAŠTINEC*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
jan.lastinec@stuba.sk

Abstract. This paper deals with the realization of a configurable gateway node between widely used automotive bus – Controller Area Network (CAN) and traditional Ethernet network. The gateway is implemented on an ARM-based dual-core single board computer with Linux operating system. Furthermore the article contains experimental evaluation of processing time of the described gateway as well as measurements of round-trip time (RTT) of CAN communication "tunnelled" through Ethernet network. The contribution of this work is the creation of a test-bed that can be used for further research in the area of securing CAN communication using Ethernet/IP technology.

1 Introduction

In-vehicle network of today's car has a complex topology where several sub-networks operating different communication technologies are interconnected by a central gateway. Due to integrating more and more innovative and advanced features in modern vehicles, the complexity of the car's electrical and electronic systems increases. Modern top-class cars accommodate up to 100 Electronic Control Units (ECUs) that are exchanging thousands of signals and run several gigabytes of control software. Because of the communication needs of such a high amount of nodes, current in-vehicle network architecture is approaching its limits.

It is widely accepted among car manufacturers that Ethernet is a promising candidate for future automotive networking. The vision is a hierarchical domain-based network with powerful "domain controller" ECUs acting as gateways between traditional automotive buses and Ethernet backbone [1].

The main goal of this work is realization of an experimental CAN-Ethernet gateway that will provide configurable system that is easily extensible to support various Ethernet/IP protocols.

The paper is organized as follows. Sections 2 and 3 briefly introduce basics of Controller Area Network and Ethernet network respectively. Section 4 deals with the related work in the area of interconnecting CAN and Ethernet. The realization of the proposed experimental gateway is described in Section 5 and the following Section describes experimental evaluation of the realized gateway. Finally the Section 7 concludes the results.

---

* Doctoral degree study programme in field: Applied Informatics
Supervisor: Assoc. Professor Ladislav Hudec, Institute of Applied Informatics, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 397–403.

2 Controller Area Network

Controller Area Network (CAN) [2] is the most used in-vehicle network in today's vehicles. It was designed by Bosch in 1980's to multiplex communication between ECUs to reduce the wiring harness and is now standardized as ISO 11898 [3]. It is used in areas where bounded response times for frames and timely delivery of control commands to actuators are needed (e.g. engine and transmission control, ABS, etc.) but also in driving-unrelated domains. It is common for cars to have multiple CAN networks – high-speed (500 Kb/s – 1 Mb/s) variants for driving related data and low-speed variants (125 – 250 Kb/s) for body and comfort functions.

2.1 Concept of communication

CAN is a serial communication protocol designed for real time control systems. It operates over twisted-pair copper wires using differential signal and provides high integrity of transmitted data and high fault tolerance. It is a multi-master bus where each node can receive and transmit broadcast messages and therefore a single message is processed by multiple receivers. CAN nodes do not have any addresses – messages are addressed by their content using CAN Identifier (CAN ID). The identifier also determines the priority of the message (lower value equals higher priority).

Access to medium is maintained by the *Carrier Sense Multiple Access with Collision Detection and Arbitration on Message Priority (CSMA/CD + AMP)* scheme. When a transmitting node detects mismatch between transmitted and sensed bit value, it stops the transmission node in favour of a higher priority, dominant bit (logical 0). This means that in case of simultaneous transmission a message with higher priority CAN ID "rewrites" a lower priority message without collision.

2.2 Frame format

CAN 2.0 specification defines four types of frames:

- Data frame – conducts data from transmitter to receivers (max. payload is 8 bytes).
- Remote frame – requests data frame with the same CAN ID (e.g. measured value).
- Error frame – transmitted when an error is detected to notify other nodes.
- Overload frame – signals that a receiving node is overloaded with incoming messages.

Data frame and Remote frame have practically identical structure. The only differences are that Remote frame has zero-sized data field and is signalled by a specific bit (RTR) in header set to 1. The format is depicted in Figure 1.

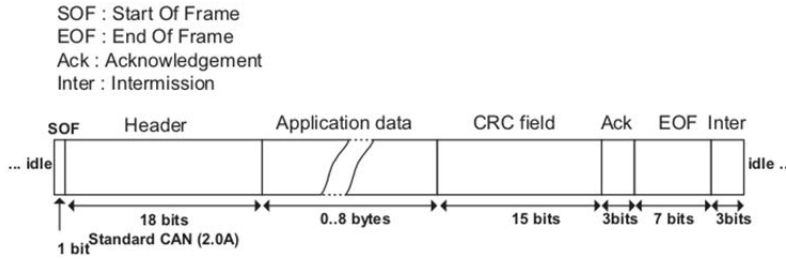


Figure 1. CAN data and remote frame format [1].

3 Ethernet

This Section briefly describes the concept of Ethernet technology and then gives an overview of its usage in automotive area. It is a proven standard used in many IT and also industrial areas.

3.1 Concept and frame format

Ethernet is a serial communication technology that uses twisted pair copper wires consisting of four conductor pairs. It is defined in IEEE 802.3 standard [4].

For medium access, *Carrier Sense Multiple Access (CSMA/CD)* method is used. A transmitting node firstly checks the medium and starts transmitting only if is available (idle). Collisions occur in case of multiple simultaneous transmissions which introduce additional delays and results in non-deterministic communication.

The standard format of 802.3 frame is shown in Figure 2. Receivers are synchronized using the preamble field and 48-bit addresses are used to identify both sender and receiver node of the frame. The specification defines minimum and maximum payload size that equals 46 and 1500 bytes respectively.

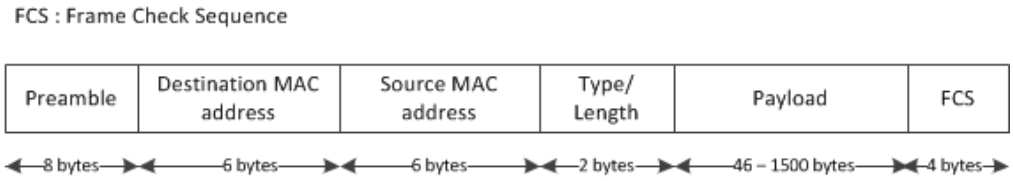


Figure 2. Ethernet frame format.

3.2 Automotive Ethernet

Ethernet is not widely used in current vehicles however many car manufacturers show interest in Ethernet as a future in-vehicle network. The main benefits are higher bandwidth, less complex network architecture and also its openness and widespread use in other areas. Today's implementations are limited mainly to diagnostics multimedia and driver assistance network domains to provide functions such as surround view for parking assistant.

3.2.1 Diagnostics over IP

The first standardization of Ethernet communications in vehicle diagnostics is a set of ISO 13400 standards referred to as *Diagnostic communication over Internet Protocol (DoIP)*. It consists of four parts: Document ISO 13400-1 [5] gives general information and use cases and documents ISO 13400-2 [6], ISO 13400-3 [7], ISO 13400-4 [8] that describe services from Physical to Transport layer of TCP/IP model. Several of these standards are under active development and the aim is to provide unified method for accessing automotive in-vehicle networks for diagnostic purposes (including remote diagnostics).

3.2.2 Standards for guaranteed response time

Standard Ethernet is non-deterministic protocol that does not provide any guarantees in terms of response times which is critical for real-time communication in automotive domain. There are multiple standards dealing with this problem and we provide short overview for reference. However this work does not deal with the communication according to these standards.

Ethernet Audio Video Bridging (AVB) was primarily defined to guarantee the upper bound of delay for multimedia oriented traffic streams. It comprises three standards (IEEE 802.1AS [9], IEEE 802.1Qav [10] and IEEE 802.1Qat [11]) that deal with timing and synchronization, stream reservation and resource reservation respectively. The guaranteed maximum latency for high priority traffic is limited to 2 ms.

Time-triggered Ethernet (TTEthernet), developed by TTTech Computertechnik AG [12], provides enhancements to enable deterministic time-triggered communication using *TDMA (Time Division Multiple Access)* scheme. System-wide time synchronization is accomplished by protocol

control frames. TTEthernet is standardized as SAE AS6802 and defines three types of traffic according to timing requirements.

4 Related work

Authors in [13] present an approach for deriving upper bound for end-to-end delay in a double star switched Ethernet network. They are specifically oriented on Ethernet in automotive area with accent on the comparison of theoretically obtained bounds with experimental data.

The test topology reflects current in-vehicle network architecture where Electronic Control Units (ECUs) serving similar functions are grouped into the same domain. The system model is shown in the Figure 3. Four types of traffic according to their priority have been considered where the individual classes were divided into different 802.1Q VLANs and 802.1p priority levels. Hard real time messages are used in powertrain chassis and body domains, soft real time messages are used for diagnostics communications; video class represents the video stream and best-effort data are used for FTP transfer.

Development boards used in the experiment were equipped with Freescale MPC5567 processor running ArcCore[1] AUTOSAR 4.1 [14] implementation.

Measured end-to-end delays through the backbone network for the highest priority class ranged from 25 μs to 300 μs which was in accordance with the expected theoretical upper bound. The validity of the theoretical calculations have been confirmed and the experimentally acquired values indicate suitability of Ethernet for automotive use.

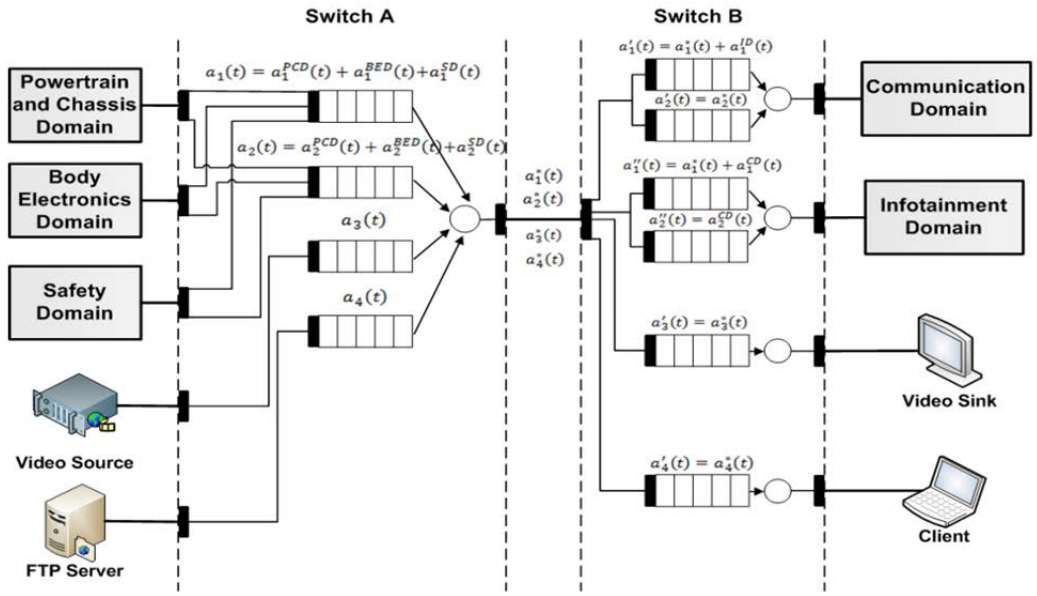


Figure 3. System model for experimental end-to-end delay measurement in [13].

5 Gateway implementation

The main goal of this paper is realization of an experimental CAN – Ethernet gateway that will serve as a base for further research regarding CAN to Ethernet communication. In order to provide the required flexibility a Linux-based Single Board Computer (SBC) with dual core ARM

[1] http://www.arccore.com

Cortex-A7 processor and 1 GB RAM has been chosen for implementation. The processor provides integrated CAN controller with freely available Linux driver[2].

Gateway application was implemented in C programming language and it provides encapsulation of the received CAN messages into UDP segments as well as decapsulation of UDP traffic to CAN.

6 Description of experiment

An experiment to evaluate the latency of the realized gateway has been carried out. The test topology consisted of a SBC acting as a CAN – Ethernet gateway, a CAN node based on Arduino development board with MCP2515 CAN controller and a Windows 7 PC connected to the SBC via gigabit Ethernet interface (see Figure 4).

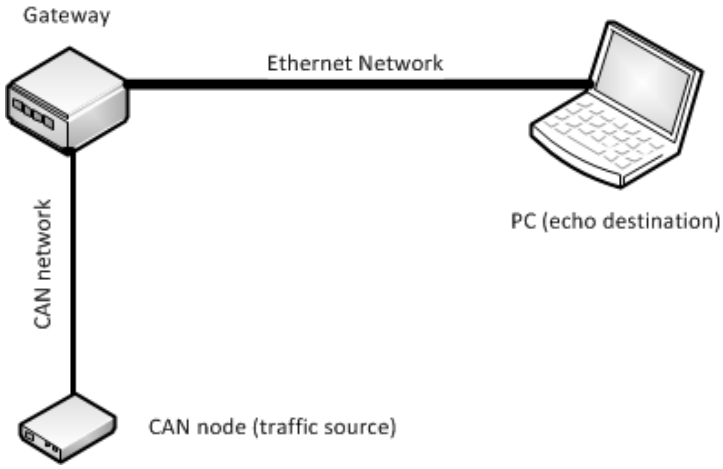


Figure 4. Test topology.

The aim of the evaluation was to determine the round trip time (RTT) of messages sent from source CAN node and relayed via UDP to PC that echoed the received data back. The CAN messages contained (maximum available) 8-byte payload and were periodically transmitted from CAN node with interval of 100 ms. The baud rate of CAN network during experiment was 500 Kbps. In addition to total RTT, the following partial delays were measured:

- Gateway outbound processing time (from CAN to Ethernet),
- RTT for UDP communication between Gateway and PC,
- Gateway inbound processing time (from Ethernet to CAN),.
- RTT for CAN communication (initial CAN message + echoed message).

The measurements on the gateway have been obtained with `clock_gettime()` function and measurements on the CAN node have been obtained with `micros()` function.

6.1 Results

The experiment consisted of five runs of 100 transmitted CAN messages to evaluate the impact of process planning overhead in CAN gateway OS by altering priority of the process providing

[2] `http://sourceforge.net/projects/can4linux/`

gateway functionality. The average processing and transmission times for each of five experiment runs are summarized in Table 1 (lower "nice" value means higher process priority).

Figure 5 presents the measured processing and transmission times visually in a stacked column charts representing the individual portions of the total round trip time (shown on the top of the columns).

As can be seen the largest share of the delay is introduced by Ethernet RTT. This could be the consequence of using traditional PC running relatively high number of tasks unrelated to the communication. Due to the limitations of running the gateway applications in user-space the overhead of CAN driver is included in CAN RTT which is the possible reason of its large value. The large value of Outbound processing time compared to the inbound processing time is left to more in-depth analysis. According to the results the effect of gateway process priority on the total round-trip time is relatively small (around 50 μs).

Table 1. Measured processing and transmission times.

	Nice 5	Nice -5	Nice -10	Nice -15	Nice -19
Outbound proc. [μs]	173,414	173,522	197,475	168,686	185,651
Eth. RTT [μs]	999,187	1039,822	958,610	971,063	989,669
Inbound proc. [μs]	9,826	10,346	11,998	10,229	11,136
CAN RTT [μs]	745,283	694,289	697,979	726,888	702,039
Total [μs]	1927,711	1917,979	1866,062	1876,866	1888,495

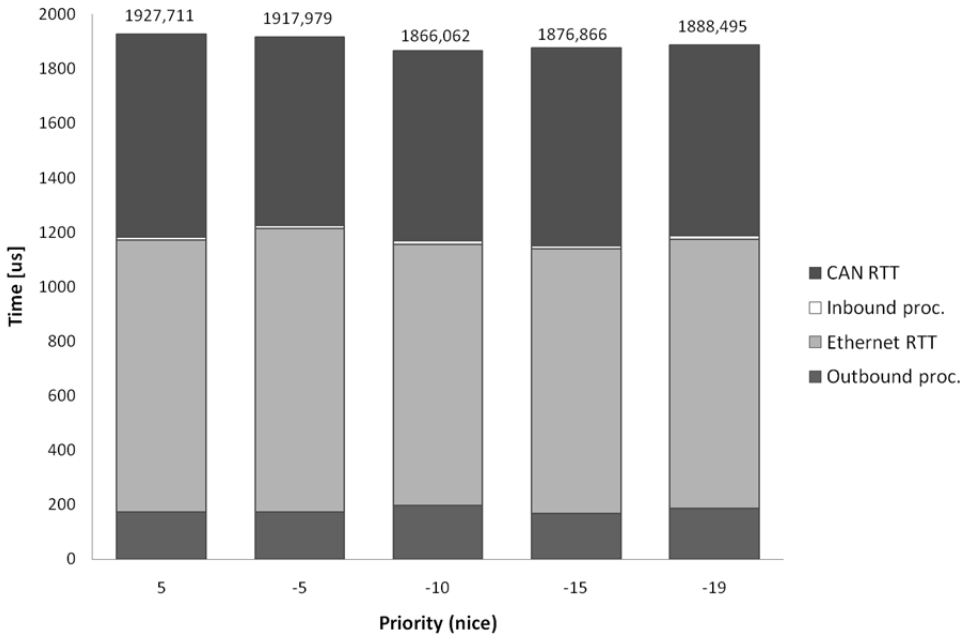


Figure 5. Portions of processing and transmission times.

7 Conclusion

In this paper, realization of experimental CAN-Ethernet gateway based on a single board computer with Linux operating system was presented. The main advantage over the solutions available on the market is the flexibility in terms of supported protocols and interfaces. The disadvantage is lower performance in comparison with dedicated devices due to the overhead of conventional OS.

An experimental evaluation of performance revealed that the total round-trip time of echo messages sent from CAN node to Ethernet node through the gateway is around 1900 μs with the largest share being the Ethernet communication time. The delays will be subject of future research. The gateway represents a base for further research of gateway strategies, namely in the area of automotive communication security which is also the contribution of this work.

Acknowledgement: The paper was prepared with support of research grant VEGA 1/0361/15 entitled "Security in distributed computer systems".

References

[1] Navet, N., Simonot-Lion, F.: In-vehicle communication networks – a historical perspective and review. Richard Zurawski, *Industrial Communication Technology Handbook, Second Edition*, CRC Press Taylor&Francis, (2013).

[2] Robert Bosch GmbH: *CAN Specification Version 2.0*. (1991).

[3] ISO 11898-1:2003: *Road Vehicles – Controller area network (CAN) – Part 1: Data link layer and physical signalling*. ISO/TC 22, (2003).

[4] IEEE Std 802.3-2012: *IEEE Standard for Ethernet*. IEEE, Inc., (2012).

[5] ISO 13400-1:2011: *Road Vehicles – Diagnostic communication over Internet Protocol (DoIP) – Part 1: General information and use case definition* ISO/TC 22, (2011).

[6] ISO 13400-2:2012: *Road Vehicles – Diagnostic communication over Internet Protocol (DoIP) – Part 2: Transport protocol and network layer services*. ISO/TC 22, (2012).

[7] ISO/CD 13400-3: *Diagnostic communication over Internet Protocol (DoIP) – Part 3: Wired vehicle interface based on IEEE 802.3*. ISO/TC 22, (under development).

[8] ISO/CD 13400-4: *Road Vehicles – Diagnostic communication over Internet Protocol (DoIP) – Part 4: Ethernet Diagnostic Connector*. ISO/TC 22, (under development).

[9] IEEE Std 802.1AS-2011: *IEEE Standard for Local and Metropolitan Area Networks – Timing and Synchronization for Time-Sensitive Applications in Bridged Local Area Networks.* IEEE, Inc., (2011).

[10] IEEE Std 802.1Qav-2009: *IEEE Standard for Local and Metropolitan Area Networks – Virtual Bridged Local Area Networks Amendment 12: Forwarding and Queuing Enhancements for Time-Sensitive Streams*. IEEE, Inc., (2009).

[11] IEEE Std 802.1Qat-2010: *IEEE Standard for Local and Metropolitan Area Networks – Virtual Bridged Local Area Networks Amendment 14: Stream Reservation Protocol (SRP).* IEEE, Inc., (2010).

[12] TTTech Computertechnik AG: TTEthernet: Deterministic Ethernet Network. [Online; accessed February 19, 2015]. Available at: https://www.tttech.com/technologies/ttethernet/

[13] Kostas, B., Ieroklis, S.: Experimental evaluation of end-to-end delay in switched Ethernet application in the automotive domain. In: *CARS: 2nd Workshop on Critical Automotive applications : Robustness & Safety*, Toulouse, France, (2013).

[14] AUTOSAR. [Online; accessed February 19, 2015]. Available at: http://www.autosar.org

Experimental Embedded System for Intelligent Actuators Control through the Internet

Ondrej PEREŠÍNI*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
ondrej.peresini@stuba.sk

Abstract. Intelligent households control is becoming trending topic not only in academic circles, but also in manufacturer opinions. Primary focus of this work is to propose a new solution of intelligent house actuators realization, which is less expensive, more robust and more secure against intrusion. The hearth of the system consists of the intelligent modules which are modular, autonomous, decentralized, cheap and easily extensible with support for encrypted and standardized network communication protocols. The proposed solution is opened and therefore ready for the future improvements and application in the field of the Internet of Things. Proof of concept was verified by the experimental prototype which is ready for direct household deployment.

1 Introduction

The worldwide industry is currently experiencing intelligent smart homes boom. Besides the recently built modern houses new projects arise to modernize existing and older buildings. This creates a demanding potential for new smart homes and increases the demands for advanced features that are easier to control and more suitable for home environment.

Currently there is no exact definition of intelligent smart home capabilities, however there are many common aspects shared between various implementations. Some smart homes are focused on HVAC (heating, ventilating, and air conditioning) thermal control and others may control household lighting. The most common problem with these models are lack of interoperability and compatibility amongst them. Our approach is focused on proposal and implementation of new communication protocol and smart home model which will allow flawless interoperability between various types of actuators and sensors. This ecosystem should be considered as a baseline for future improvements and application in the field of the Internet of Things.

* Doctoral degree study programme in field: Computer Engineering
Supervisor: Assoc. Professor Tibor Krajčovič, Institute of Computer Systems and Networks, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 404–411.

2 Limitations of current approaches

During analysis stage we have analyzed many academic approaches as well as some commercially available systems. As a result we were not able to discover system which will be sufficiently modular, safe and will meet all system specifications. However we have found two interesting yet differently specified academic publications related to our topic by *P. Jombík* [1] and *S. Kišák* [2].

Numerous examined commercially available systems for intelligent household control (*Sensaphone FGD-W600* [3], *LockState Connect LS-90i* [4] and *Ecobee EB-EMS-02* [5]) are linked with several disadvantages:

- Extremely high price – which is a result of this lucrative market and limited offer.
- Proprietary and closed implementations – as producers tend to guard their know-how and the majority of systems are closed and do not allow modification or extension of the functionality.
- Complex configuration and installation – a common problem of commercially available systems that require professionally trained and expensive assembling staff (this is another well-paid service which companies usually offer).
- Low level of security – where most of the equipment is protected only by access password, but does not have the additional security validations that may avoid system hacking.
- Various functional limitations – provided devices often do not provide any type of software or hardware update and users are, therefore, locked to provided functionality without the ability to extend this system.
- Centralized control – in case of one module failure the whole system is completely disabled and users are forced to perform expensive repairs or change the entire system.

2.1 Existing standards

Since most of the intelligent household systems are proprietary and closed, it is virtually impossible to operate devices from different vendors. Beside previously mentioned aspects the main problem lies in unavailability of open standard that would be sufficiently widespread and complex to cover whole field of intelligent household control. However there are two standards HVAC and DLNA (*Digital Living Network Alliance*) which can provide some of the intelligent household functionality, but both of them are dedicated to separate function fields and even combining both of them will not result in sufficient complexity that would not limit different hardware components. As a resolution of this crisis we have decided to design and implement our own standard that will be open source and will meet all established specifications for functionality.

3 System specifications

The main objective of the project was to design and implement a system of intelligent household management which will consist of several modules performing specified tasks. The system is composed of a central unit and numerous sensor modules, which may be combined with actuators. The central unit serves as a register of external modules and provides an intuitive user interface for easy configuration of the whole system. The proposed system consists of modules that are:

- *Modular* – which allows user to change or expand the functionality of individual modules.
- *Autonomous* – and therefore capable of independent action.
- *Decentralized* – system allows dynamic adding or removing of different modules without the need for a central unit or any manual configuration.

- *Standardized* – communication protocol between modules should be standardized and thus allowing future design and implementation of additional modules with different or extended functionality.
- *Cheap* – component design goal is to minimize financial cost mainly through usage of standardized components and existing communication network.

3.1 Network communication

One of the main pillars of the project is to incorporate already existing network infrastructure, which brings many advantages but also some disadvantages, whose impact can be minimized by appropriate system design. The modules communicate with the other elements through the standardized Ethernet (*IEEE 802.3*) or WiFi (*IEEE 802.11*) network. As embedded systems provide limited hardware resources to fully support extensive TCP (*Transmission Control Protocol*) we have decided to design and implement our own communication protocol based on a simple UDP (*User Datagram Protocol*) which we have expanded by the number of elements known from TCP. The result is a fast and reliable communication protocol with minimal memory requirements and ability to de-duplicate, retransmit and recover from errors.

In the main design we have specified various types of communication messages that are divided into Broadcast and Unicast communication. Whilst the *Hello* message is broadcasted through the whole network segment, all other messages including *Identify*, *Request*, *Post*, *Configure* and *Acknowledgment* are directly exchanged between only two nodes. Periodically broadcasted message *Hello* builds up communication between the central unit and all external modules. Upon receiving the *Hello* packet follows the unicast message *Identify* containing all information about module identification and configuration. Afterwards all previously stored sensor readings and unsent data are transmitted via *Post* messages. Central unit could solicit readings update or revoke data transmission through *Request* message. Configuration changes are committed by central unit which sends *Configuration* messages to desired modules. All messages need to be confirmed via *Acknowledgment* message or retransmitted due to packet loss or corruption.

All messages consist of a header, encrypted message body and completion segment. In order to ensure correct messages order all packets include fields with sequence number, checksum and time stamp. Control logic using these fields then ensures message integrity as in the case of TCP.

Table 1. Specification of the communication messages.

Field	Identification Header		Data			Completion
Description	Message type	Sequence number	Data size	Padding	Checksum	Time stamp
Size (B)	1	4	1	8x – 2	1	4
Encrypted	No		Yes			No

3.2 Security measures

Security system is divided into several layers that provide advanced security against intrusion. The transport layer performs drop of duplicated, replicated and delayed messages that could be maliciously created by attackers. To ensure confidentiality symmetric XTEA encryption is used together with dynamic password generation which makes it virtually impossible to perform brute-force attack on used encryption algorithm. XTEA uses bock size of 64 bits (8 B) and therefore all messages must be stuffed with padding data of size (8x – 2 B) to effectively align encrypted data.

The application layer of user interface provides access from Intranet as well as from Internet and therefore advanced security features must be provided. Limited resources of embedded systems do not allow usage of asymmetric cryptography well known from HTTPS (*Secure Hypertext Transfer Protocol*) and therefore we used standard HTTP interface with user authentication hidden in BASE64 encoding. This approach does not provide sufficient level

of security and thus we incorporated 2-stage GSM (*Global System for Mobile Communications*) authentication. One-time verification codes are transferred via SMS messages. Successful security breach may be achieved only by compromising both independent communication channels.

The operation layer is secured through Watchdog timer which is able to restore system upon unpredictable operation which may lead to system hangout and failure.

4 Solution proposal

For implementation of a central unit and external modules we used Arduino platform which provides a wide developer base with many open source projects. This platform also provides various development kits suitable for the prototype realization. Whilst external modules do not need fast processor and should be power efficient we decided to choose popular *Arduino UNO*[1] kit with *Atmel ATmega 328* [6] processor together with *Ethernet*[2] [7] or *Wireless shield*[3] [8]. Central unit needs more processing power as it collects sensors data and provides user interface and therefore we have chosen *Arduino Mega 2560*[4] based on *Atmel ATmega 2560* [9] processor.

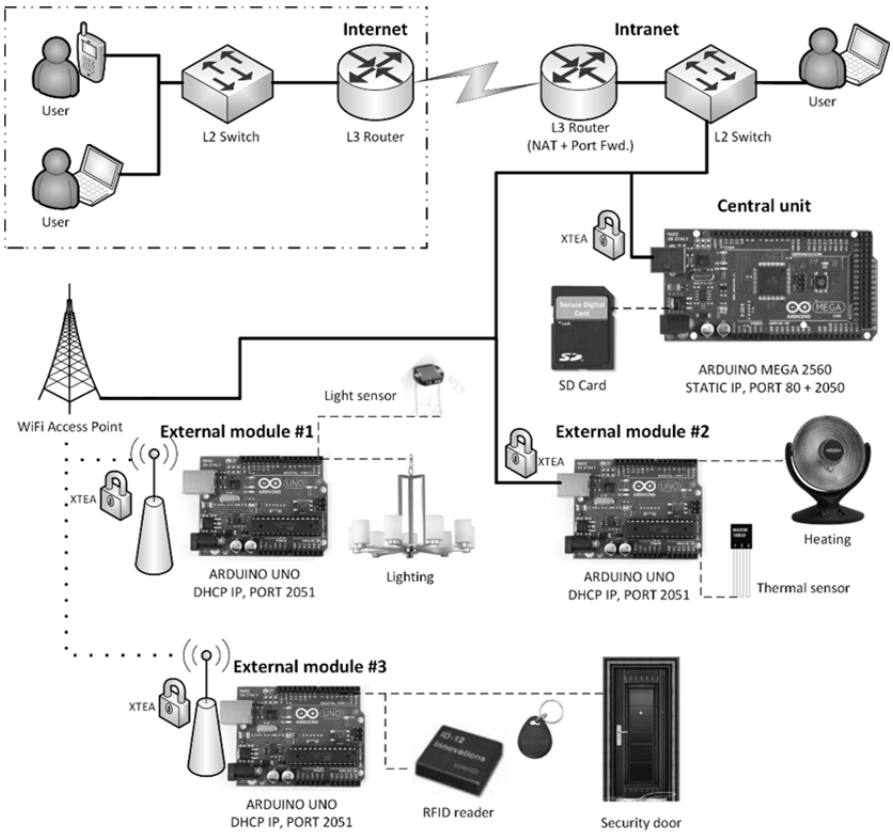


Figure 1. Block diagram of the proposed solution.

[1] http://arduino.cc/en/Main/ArduinoBoardUno
[2] http://arduino.cc/en/Main/ArduinoEthernetShield
[3] https://www.sparkfun.com/products/9367
[4] http://arduino.cc/en/Main/ArduinoBoardMega2560

In the design we have used three types of external modules as exemplary demonstration of project capabilities. Our selection of the sensors and actuators should cover basic requirements on intelligent houses. Adding more sensors and actuators is suitable area for future project expansion.

Each of the module collects data about the environment through selected sensor and interferes with the surrounding environment via dedicated actuator. For monitoring of environment we have used light sensor (*PERKIN ELMER FW300* [10]), digital temperature sensor (*DS18B20* [11]) and RFID card reader (*ID Innovations ID-12* [12]) which will be mounted inside front door and grant access to authorized users only.

Actuators can directly regulate environmental characteristics that are linked with chosen sensors. The lighting control is based on a LED (*Light-emitting diode*) with direct PWM (*Pulse-width modulation*) regulation, the door lock is linked with electronic release mechanism connected to the opening relay and room temperature is regulated via stepper motor, which may control heating valves in a different areas.

5 Hardware realization

All hardware modules are based on the development platform of Arduino which allows to use stackable headers shields to extend functionality of this compact system. Every module have 2 shields stacked on the top of the mainboard. Whilst the first shield provides network connectivity (Ethernet or WiFi) the second shield that is designed and created by us allows to connect specified types of sensors and actuators. On the top of the central unit is located GSM module with external communication antenna. All WiFi adapters have only internal antenna which is sufficient for our use case and may by extended by adding external antenna connected through the provided connector. Only two of the modules are connected through Ethernet and these modules are powered via POE (*Power over Ethernet*). Remaining modules with wireless connectivity must rely on batteries or power adapters.

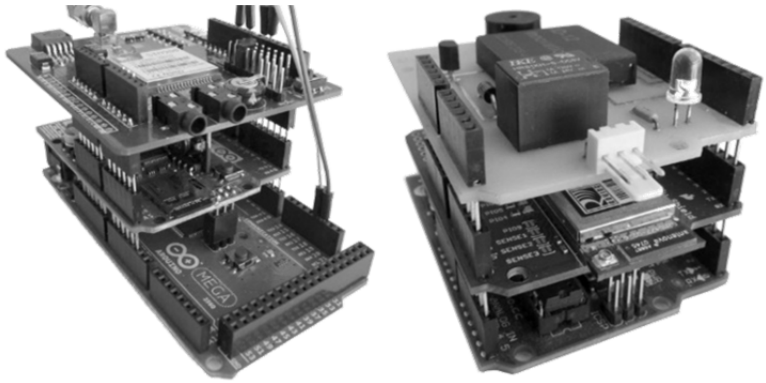


Figure 2. Prototypes of Central unit and External unit with RFID sensor shield.

6 Software realization

The most important parts of the proposed system include sensor value readings and associated actuators control. In case of power failure all measured values are safely stored into built-in EEPROM memory and SD Card storage where we created our own dynamic data structures to preserve all data. Network communication performs not only sending and receiving packets, but also packets analysis, processing and data encryption. In addition to the proposed protocol we have used standardized NTP (*Network Time Protocol*) and DHCP (*Dynamic Host Configuration Protocol*) messages as well.

The central unit is in comparison to the external modules quite complex due to additional processing of GSM shield[5] [13] communication, generating user interface and data handling. Readings form all modules including their settings, sensors readings and error messages are stored at the memory card. The central unit also provides a user interface that is based on HTTP web protocol. This interface form enables interoperability across a wide variety of different devices.

In the software part of the implementation we had modified and adapted some of the existing libraries to meet our requirements:

- *Ethernet communication library* – added new communication protocol.
- *WiFly interface* – adopted to cooperate well with the new protocol.
- *SoftwareSerial* – tweaked to operate with *ID Innovations ID-12* RFID reader.
- *MemoryFree* – modified to identify memory allocation leaks.
- *WebServer* – hardened security and tweaked to support faster data transfers.
- *GSMshield* – reduced vast bloat-ware to minimalize memory footprint.

Our specifically implemented software parts contain:

- Network protocol implementation – based on proposed protocol solution.
- XTEA encryption library – developed to enhance system security via symmetric encryption.
- Different sensors readings interpretation – Serial, OneWire and Analog.
- Various methods of actuator control – PWM, Relay switching and Stepper motor control.
- EEPROM – read / write operations to store configurations.
- SD Card functions – allowing sensor data retention and graphic representation.
- GSM communication – to send user SMS messages with One-Time-Passwords.
- Unique user interface – compatible with HTTP protocol and accessible via various devices.
- WOL (*Wake on LAN*) and Watchdog reset – as advanced features supporting power saving.

6.1 Implementation challenges

During implementation process we encountered a number of problems, which were mainly based on various embedded hardware limitations. The biggest problem was the lack of RAM, where we were forced to optimize our code and libraries several times to fit the whole functionality into the available memory of capacity only about 1800 bytes. RAM size would not allow us to utilize graphical user interface and therefore we outsourced graph rendering to the online graphical framework known as *Xively* (or previously *COSM* or *Pachube*). Perhaps the most problematic area was network communication implementation, which is prone to flooding overload and limitation of maximum number of active TCP sessions. We had to carefully tune whole network communication.

Meanwhile we discovered the strange behavior of WiFly module that does not comply with the specifications [8] and we had to modify the logic of establishing network communication. Another minor issue was the lack of control mechanism handling expiration of DHCP IP address lease and therefore we have implemented our own DHCP mechanism. In addition, digital temperature sensor has long processing delay of about 1 second, which caused data retransmission as opposite node presumed loss of connection. We have solved this problem by automatic sensor readings buffering which may slightly increase energy consumption but provides much faster response times. Since the Arduino platform does not have real time clock circuit we have used NTP time synchronization sent at regular intervals.

[5] http://imall.iteadstudio.com/im120417009.html

7 System verification

We had established thorough system verification of the final product prototype not only in terms of the available functionality, but also for long-term deployment in real homes. Test cases were divided into positive and negative verifications. Positive test cases were designed to verify defined system specifications and overall product capabilities. We have successfully verified network communication process together with designed transport protocol and its behavior. Network topology represented standard home local network where all external modules had dynamic IP addresses and Central unit was assigned with static IP address due to the static port mappings used for system accessing from the outside internet.

Separate parts of test cases were dedicated to XTEA packet encryption and accuracy of data storage on both SD Card and embedded EEPROM memory. We have established some function tests as well and verified proper functionality of all hardware sensors, actuators and developed Shields. The most time consuming testing was verifying proper output and commands of the user interface because every sensor and actuator has its own configuration page and control settings.

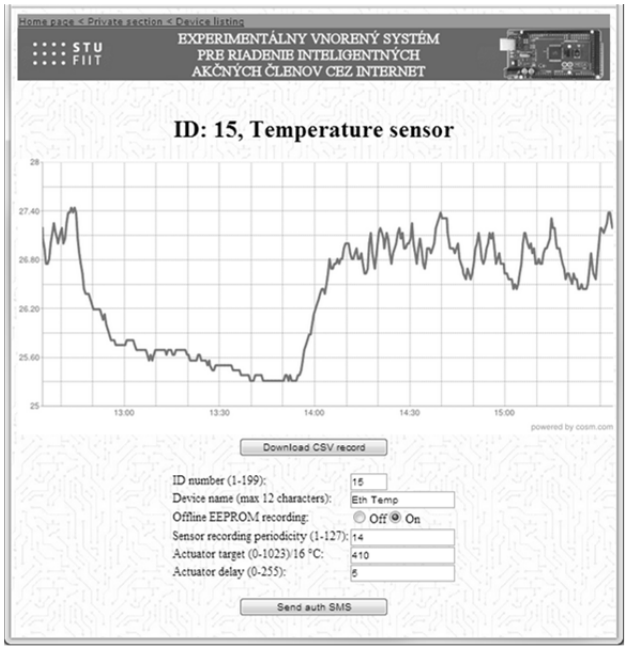


Figure 3. User interface page of Temperature module settings.

Negative verification testing consisted of many attempts to disrupt the system in any way. We have simulated security network attack by sending some invalid and fabricated packets. In first attempt we tried to create invalid requests to gain access to sensors or actuators. This attempt was unsuccessful thanks to XTEA encryption, however in case of rapid packet generation we could experience DOS (*Denial of service*) failure. This is a common server problem and can be avoided with a Firewall.

Another part of the test was central unit failure simulation and random disconnection of the Internet connectivity and SD Card. The result of the test proved that autonomous external modules can automatically disconnect from Central unit and start to record measured data locally to EEPROM memory. In addition to the described tests we have done numerous sub-tests and subsequently repaired all discovered problems. System prototype is therefore sufficiently robust and suitable for direct application in homes.

8 Conclusions and further work

During final product development we have analyzed, designed and implemented fully functional household management system. This system has in comparison to commercial solutions several benefits including modularity, autonomy and decentralization of external modules. The advantage is in the simple scalability of all modules and the system is therefore practically unlimited in this direction. During system design we have focused on the minimum cost and energy consumption. The indisputable advantage of the system is the use of existing network infrastructure, so there is no need to create a new type of network. In the design we have also paid attention to adequate security of the whole system against unwanted intrusion what is not a strong case of commercial solutions.

The proposed system is currently in the prototype stage but can be immediately deployed in the home environment. The modularity of the system allows for future expansion of hardware capabilities or software control. Provided sensors and actuators may be used as good examples to show basic household functionality. This system can be expanded in the future to further improve user experience via applications primarily directed to different mobile phone platforms.

Acknowledgement: This work was partially supported by the Grant no. 1/0616/14 of the Slovak VEGA Grant Agency.

References

[1] Jombík, P.: *Embedded system for remote monitoring and management of the household*. Diploma thesis, FIIT STU, Bratislava, 2011.

[2] Kišák, S.: *Experimental embedded system for remote monitoring and management.* Diploma thesis, FIIT STU, Bratislava, 2011.

[3] Sensaphone: *Web600 Web-Based Monitoring System*. [Online; accessed January 20th, 2015]. Available at: http://www.sensaphone.com/sensaphone_web600.php

[4] Lockstate: *LS-90i Internet Thermostat*. [Online; accessed January 20th, 2015]. Available at: http://www.lockstateconnect.com/ProductDetails.asp?ProductCode=LS-90i

[5] Ecobee: *The ecobee Smart Thermostat*. [Online; accessed January 20th, 2015]. Available at: http://www.ecobee.com

[6] Atmel: *ATmega 328 DATASHEET*. [Online; accessed January 25th, 2015]. Available at: http://www.atmel.com/Images/doc8161.pdf

[7] WIZnet: *W5100 DATASHEET*. [Online; accessed January 26th, 2015]. Available at: http://www.wiznet.co.kr/UpLoad_Files/ReferenceFiles/W5100_Datasheet_v1.2.2.pdf

[8] Roving Networks: *WIFLY GSX DATASHEET*. [Online; accessed January 27th, 2015]. Available at: http://www.sparkfun.com/datasheets/Wireless/WiFi/WiFlyGSX-um2.pdf

[9] Atmel: *ATmega 2560 DATASHEET*. [Online; accessed January 25th, 2015]. Available at: http://www.atmel.com/images/doc2549.pdf

[10] SOS: *Fotorezistor PERKIN ELMER FW300*. [Online; accessed February 8th, 2015]. Available at: www.sos.sk/?str=371&artnum=28919&name=perkin-elmer-fw300-95503513

[11] Maxim: *Digital Thermometer DS18B20 DATASHEET*. [Online; accessed February 8th, 2015]. Available at: http://datasheets.maxim-ic.com/en/ds/DS18B20.pdf

[12] SparkFun: *ID series DATASHEET*. [Online; accessed February 9th, 2015]. Available at: http://www.sparkfun.com/datasheets/Sensors/ID-12-Datasheet.pdf

[13] SIMCom: *GSM/GPRS Module – SIM900*. [Online; accessed February 2nd, 2015]. Available at: http://wm.sim.com/producten.aspx?id=1019

TP CUP 2015

TP Cup – The Best Student Team Competition Showcase at IIT.SRC 2015

Mária BIELIKOVÁ*

Slovak University of Technology
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
maria.bielikova@stuba.sk

Abstract. Best student team competition TP-CUP is organized 7th time this year. The competition is aimed at excellence in development of information technologies solutions within two semester long team project module in master degree programmes. This year 15 student teams presented in IIT.SRC 2015 showcase their projects. Key concepts of their projects are included in the following sections of the proceedings.

1 Background of the Competition

Team projects play an important role in the education of engineers. Team projects have a long tradition in informatics and information technologies study programmes at our university. Module firstly named *Team project* was introduced in the academic year 1997/1998 in software engineering and in subsequent years it was adopted as compulsory module for all master degree students. Its intake is each year 25-30 teams of 5-7 students in all study programmes. The main objective is to give students a hands-on experience with different aspects of working in team on a relative large task.

In designing a team project as a part of a curriculum, we considered several aspects or different alternatives to particular issues such as team formation, team communication methods, team assessment, problem assignment, development process and team supervision. Our experience with such projects is that a satisfiable solution (in terms of the team project objectives, i.e. experience with different aspects of working in team on a large problem) requires time longer than one term, so we designed our team project as two semester module. Supervisors who are available (either academic staff or an industry partner) determine problems being solved. Teams consist of 5-7 students. They are created under our active control. Our criteria aim at balancing differing specific knowledge of team members and different experience in various team roles. We also respect the students' preferences to some extent (a student can specify one student to become a member of the same team).

We let teams bid for problems proposed by supervisors. A competition between teams is established and students have opportunity to exercise writing and presenting the bid. The students bid with their knowledge, skills and achievements related to the selected problem, and with

* Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 415–416.

a preliminary sketch of solution based on the open question-answer session with a customer (mostly a supervisor).

Although the quality of the final result is an important measure of a success of a team, we markedly concentrate on the process applied. Through the years of providing Team project module we adopted the development process with at least two iterations. Seven years ago we have started with agile developments methods. First four years several teams employed agile development methods each year. From academic year 2012/2013 all teams follow agile development according selected agile methodology. This year all teams work according the Scrum methodology utilized for educational context.

The amount of freedom and supervision should be balanced in order to create a true learning experience for students. To simulate the reality, students should have a considerable amount of freedom. On the other hand, since students usually have no or just little project experience, some amount of supervision, monitoring and guidance is needed to ensure sufficient progress and a successful result. In order to reach balance between freedom of students and supervision we specify in advance certain requirements on the content of documentation to be produced. Students have to prepare and follow a detailed project plan. We prescribe certain parts of the project plan, such as list of activities, milestones, dependencies, and responsibilities according to established team process. Students are free to define the activities that are necessary to successful accomplishing of the project. We accompany the Team project by lectures on project management, teamwork, and quality assurance.

2 Stages of the Best Student Team Competition

In order to emphasize excellence of the students' teams we established the Best Team Competition called TP Cup in academic year 2008/2009. The competition is aimed at excellence in development information technologies solution within our two semester long team project module in master degree programmes.

The competition has several stages. It starts with an application in the middle of autumn semester.

- First stage finishes by the end of autumn semester when the teams submit an interim report. We filter out teams which do not fulfil basic criteria on quality of work performed.
- Second stage culminates in the middle of spring semester when student teams submit key concepts of their projects in the form of two page report published in the IIT.SRC proceedings and present their projects in the TP Cup showcase organized as a part of our student research conference IIT.SRC. This year we started with 18 teams with accepted applications; 15 teams advanced to the second stage. These 15 teams presented their projects in form of showcase at the IIT.SRC 2015.
- Third stage presents finalizing the projects. It ends by our grand finals where board of judges consisting experts from industry selects the winning team which lands the challenge cup – "*Best FIIT Student Team of the Year*".

More information about TP Cup can be found on the Web:
`http://www.fiit.stuba.sk/tp-cup/`

ProjectAid
A Web Portal to Support Growth of IT

Daniel ANDRÁŠIK*, Bálint CSÓKA*, Erik GRMAN*,
Eduard KUBINEC†, Radoslav NEMEC†, Vladimír OGURČÁK†

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
team13everyone@gmail.com

ProjectAid is an initiative to create a web portal to support research and projects in the sector of information technology. Its main purpose is to create a communication and donation channel between the desired organisations and their sponsors. With the use of this portal, contributors are allowed to see and follow the development of projects, initiations, conferences and anything else enrolled for support, transparently. The donated amount from the system is assigned to these causes individually and every registered participant is notified about the use of their own donation, if it was placed non-anonymously. Moreover, the application simplifies publicizing projects or any other causes that need raising certain amount of money.

The contributions can be added one-time only or can be set as recurring for a specific period of time. The concept of crowdfunding relies heavily on the supporting community [1], so since the application is collecting funds for the Faculty of informatics and information technologies of Slovak University of Technology in Bratislava, the most important target audience are the alumni of this faculty. The user base is not limited to these individuals, other users can also register and participate. Legal persons, like companies, organisations, etc., may also create an account and use the functions of the portal as well as individuals.

As the main principle of this project is to find supporters for specific causes or projects at the Faculty of informatics and information technologies, at first the analysis of identical portals was arranged. Due to their similarity in functionality, Kickstarter[1] and Indiegogo[2] were selected for analysis. They both share the same idea: they were made to raise funds for a project until their predefined goal is not reached, they are designed user-friendly and easy to use. On the other hand, they lack the ability to fully anonymise contributions, they do not allow supporters to donate the organisation itself and not the project, and do not track the evolution of the project or cause after it has been successfully funded.

Above mentioned crowdfunding platforms utilize multiple funding models. These funding models can be categorized as “return rule” or “direct donation”. Return rule is all-or-nothing funding model. This means that when the particular amount of funds is not raised until specific

---

* Master degree study programme in field: Software Engineering

† Master degree study programme in field: Information Systems
Supervisor: Karol Rástočný, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

1 https://www.kickstarter.com/hello

2 https://www.indiegogo.com/about/our-story

date, all funds are returned to the respective contributors. On the other hand, the second funding model "direct donation" provides more flexibility, because this allows project creators to keep the resources even if the desired fundraising goal was not reached [2]. We decided to employ the direct donation method in our solution, since we expect wide range of various projects and causes to be added to the system, therefore flexible funding is highly appropriate.

Based on our team's analysis of the current state of similar systems and clarified needs of the customer, the main functions of the portal were defined. It was equally important to enable potential supporters to log in, manage their profile, add a contribution by their name or anonymously, to enable researchers, colleagues and employees of the faculty to register and add their causes for donation and to create a system, which can manage the payments both automatically and manually.

One of the main functions of the portal is to allow specific users from the faculty to create a new project. Users can add basic information about their project or cause, or more advanced information like the project webpage, video and logo. They can also stylise description and abstract of the project with rich text function. A project is an entity in the system, which can be added by a registered user of the faculty, it can be either a collaborative or individual work, which is planned and done to achieve a particular aim, a list of items, that need to be bought or supplied, activities, that need financial support or needs of the organisations or employees of the faculty for a specific cause.

The supervisor of the project can set special user permissions to other members of the project team. With sufficient permissions, users can edit project information, or add a new status update to the project, which represents the progress of work on the project. Possibility of adding statuses to projects ensures that project information are always up-to-date and participants, such as supporters, are always kept informed.

After making a donation, the supporter can see the used and the left amount of the resources in the system. The used amounts are always presented to the supporter with the particular project, which used the amount. A list of the donated amounts so far can be easily retrieved and reviewed by this user. If the contribution was made anonymously, it is not shown to any user in their donation list. The registered supporter also has the ability to manage his own profile and change personal information as he needs.

While creating the system, special emphasis is put on transparency of all processes. Contributors, who will support projects non-anonymously, are able to track their contribution and see for which cause it was used. This helps to ensure that all raised funds will be used in a proper way. In order to make resource flow fully transparent, contributors will get notifications about usage of their money and about progress of funded projects via email. Moreover, statistics will be available that will present information like amount of raised funds or number of funded projects in an understandable and user friendly way.

The application is built with the growing user base of mobile and tablet users in mind, based on modern technologies like .NET, HTML5, jQuery, BootStrap, to make it responsive in every situation on every device. Since it is available and optimized to almost all of the modern web-enabled devices, it gives the ability to register to anyone with a browser, and support the growth of IT by receiving and processing donations.

References

[1] Greenberg, M., Hui, S., Gerber, E.: Understanding the role of community in crowdfunding work. In: *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, ACM, (2014), pp. 62–74.

[2] Gerber, E., Muller, M., Wash, R.: Crowdfunding: an emerging field of research. In: *CHI '14 Extended Abstracts on Human Factors in Computing Systems*, ACM, (2014), pp. 1093–1098.

3D UML

Hana BARANOVIČOVÁ[†], Miroslav KUDLÁČ*, Francisc JURAS*, Lukáš MARKOVIČ[†], Martin MELIS*, Michal VALOVIČ[†], Andrej ŽELEZŇÁK*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
team4.fiit1415@gmail.com

In the world of technology, there are lot of complex and hard to understand software systems. Simply, because of the difficulty of large systems, software engineering relies on auxiliary languages, that serve to describe software systems through different layers. Probably the most widely used language is UML, which is popular for its expressive power and simplicity.

UML is a language, which can be used to describe software systems on each level of abstraction. However, each described system is unique and special, which means, that result diagrams could be too complex and less readable, then originally intended. Commonly used CASE tools are working with UML in two dimensional space. But there is a possibility to extend UML into three dimensional space, with preserved original language definition and diagrams metamodels. This UML could be used for simplified views on multi-layer systems or complex diagrams divided by other logical relations into layers. Our team focused on three dimensional activity diagram, which is used to describe workflows of activities and actions. In this paper we describe the main concept of our implementation.

Activity diagrams basically include start node, end node, activity nodes and control nodes like fork, join, merge and decision. These nodes are connected by flow lines. Structured activity represents a special node, which is built using another nodes and flows, but apparently acts separately. The links between these nodes are described by the activity diagram metamodel [1]. Our concept of three dimensional activity diagram contains all of these basic building blocks from the classic two dimensional activity diagram and place them into the three dimensional space. For our solution, a new metamodel of three dimensional activity diagram was created, which respects all relations from the classic two dimensional activity diagram metamodel, but is enriched by fragment characteristics from sequence diagrams, due to constructions, that cannot be covered by structured activities.

The actual prototype is implemented using C++ language, while for the graphic part, Ogre3D graphical framework was used. Existing system also contains other types of diagrams, which were implemented in previous parts [2]. Our prototype part contains scene, with variable number of layers. These layers can be selected by user using a selection manager and selected building blocks can be inserted into them by using a mouse pointer. Each node can be connected with a different node and also on a different layer. Connection is created after selecting two inserted blocks appropriately based on the metamodel.

---

[*] Master degree study programme in field: Information Systems

[†] Master degree study programme in field: Software Engineering
Supervisor: Dr. Ivan Polášek, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 419–420.

Architecture of our prototype part is based on the model-view-controller design pattern. In spite of our three dimensional UML representation, there is a need to keep diagram information consistent, to ensure cooperation with existing CASE tools, like Enterprise Architect, or Rational Software Architect. Because of this feature, we need to keep two types of information – diagram structure and appropriate three dimensional representation. Classic diagram information will be stored in Model class, but because of three dimensional representation we introduced a special model-model-view-controller-controller pattern. Data for displaying diagram in three dimensional space is then stored in second – 3d model. Because of this model separation, also two controllers were used, one for model modifications, and other for 3d model modifications. Patter contains only one view class. This view class is responsible for the preparation of the diagram data, which are presented on graphical user interface.

Pattern, which was used in our prototype, presented in Figure 1, provides a very flexible and comfort way, on how to work in the three dimensional space. Identical or similar concept of architecture can be also used in the future work on this CASE tool, that consist of implementing other types of diagrams, like use the case diagram, communication diagram or component diagram. Another challenge is a proper XMI serialization of each diagram, which will ensure the compatibility with the mentioned CASE tools.

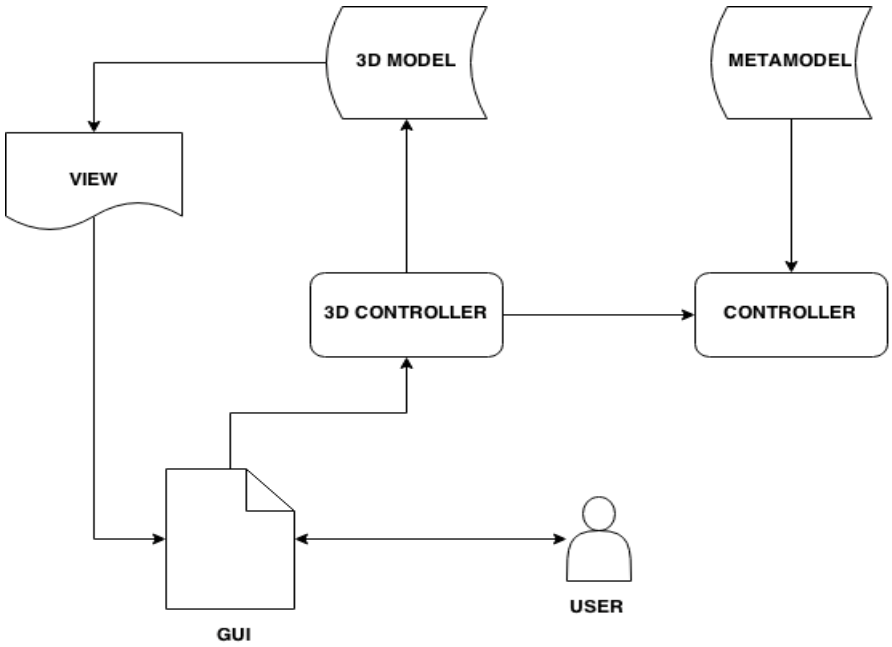


Figure 1. Architecture of three dimensional UML tool.

References

[1] Bock, C., Cook, S., Rivett, P., Rutt, T., Seidewitz, E., Selic, B., Tolbert, D.: *OMG Unified Modeling Language TM (OMG UML) Version 2.5.* [Online; accessed September 05, 2013]. Available at: http://www.omg.org/spec/UML/2.5/Beta2/PDF/

[2] Škoda, M.: Three-dimensional Visualization of UML Diagrams. In: IIT.SRC 2014: Student research conference. *10th Students research conference in informatics and information technologies*, Bratislava (2014), pp. 597–598.

Appmonitor: Replacing Traditional Logging Mechanisms in Software Applications Monitoring

Barbora BROCKOVÁ[†], Matej ČAJA*, Martin ČÍŽ*, Peter KYSEĽ[†],
Adam LIESKOVSKÝ[†], Daniel UDERMAN*, Michal VISKUP*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
gappers-tp-fiit@googlegroups.com

Application monitoring is an important part of the project development lifecycle. Application logs are the primary source of the information on the application performance. They often contain valuable information on the application instability. While various mature solutions for logging exist, the processing and analysis of their output still remains a challenge.

Alternative approach is to process the application events on the fly. Several solutions based on this approach already exist. Most notable are NewRelic and KISSMetrics. NewRelic focuses on monitoring application performance, while KISSMetrics monitors user activity. NewRelic produces data valuable for developers and operations staff, while KISSMetrics data is valuable for people in marketing.

Nonetheless, none of these tools does replace the traditional logging mechanisms. We present a tool that processes the user-defined events from the monitored applications and presents the result of the processing through the web-browser interface (see Figure 1). We aim to completely replace the existing logging mechanisms and thus relieve the development teams of the tedious and unpopular checking of the application logs.

To achieve this, Appmonitor offers the following:

- monitoring of both generic and user-defined application events,
- automatic deduplication of events,
- statistical data kept for each event,
- visualisation of statistical data.

We are facing two major problems. Large enterprise applications produce megabytes or even gigabytes of log entries in minutes or hours. Thus data stream to process in real time is intense. We adress this issue by modular architecture of our solution. Such architecture results into great vertical and horizontal scallability of our solution.

---

* Master degree study programme in field: Information Systems

† Master degree study programme in field: Software Engineering
Supervisor: Jakub Ševcech, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 421–422.

The second issue comes up after the data is processed. In effort to maintain short response time of our application, we experiment with two solutions: OLAP [1] and Elasticsearch [2]. This however still remains a challenge.

We intend to provide extensive analytics and visualizations upon gathered statistical data. Combined with short response time and ability to process large amount of data in real time our solution has the potential to replace the traditional logging mechanisms. What is more, with all the statistical data our solution offers additional insight to user reading through the events and thus helps him make more qualified assumptions.

Future work will thus focus on utilizing the gathered statistical data and further enhancing the data storage to maintain short response times.

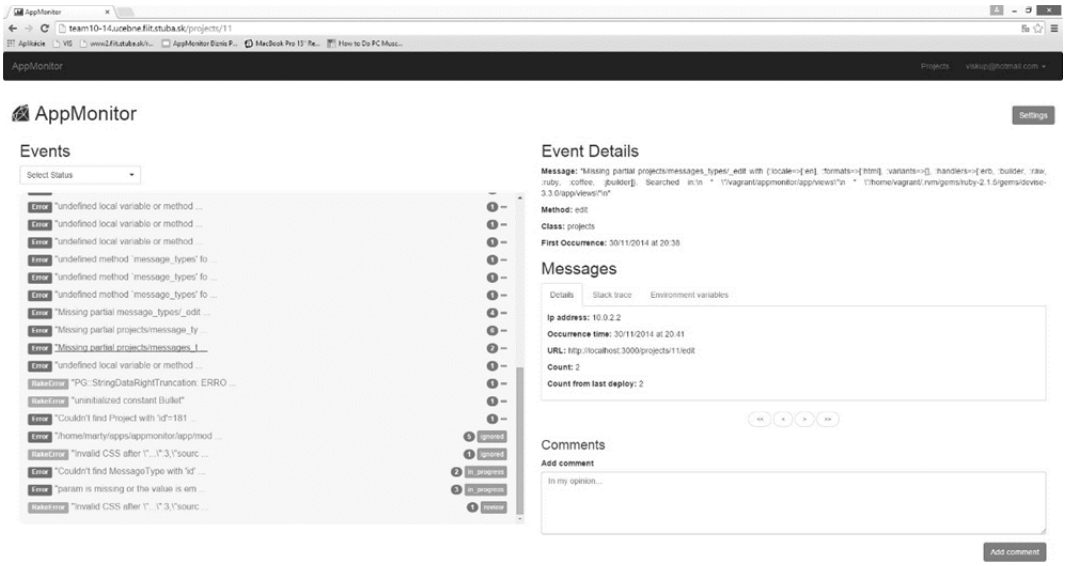


Figure 1. Appmonitor dashboard showing events and details.

References

[1] Niemi, T., Nummenmaa, J., Thanisch, P.: Constructing OLAP cubes based on queries. In: *Proceedings of the 4th ACM international workshop on Data warehousing and OLAP*, ACM Press, (2001), pp. 9–15.

[2] Kononenko, O., Baysal, O., Holmes R., Michael W. Godfrey : *Mining modern repositories with elasticsearch*. In: *Proceedings of the 11th Working Conference on Mining Software Repositories* (MSR 2014), ACM Press, (2014), pp. 328–331.

Make Games a Better Fun

Rudolf BRISUDA[*], Tomáš DRUTAROVSKÝ[†], Andrej LENČUCHA[*], Lukáš PODOLÁČEK[*], Stanislav RYKALSKÝ[†], Michal SUSKO[†]

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
tp1415.01@gmail.com

Mobile games are usually fun just for a specific group of players. However, players often do not have much time to play. Whether they travel on the bus, wait at the doctor etc., they need to start the game and play immediately. The most games force players to lengthy game characters selection, read tutorials and test what is appropriate for them. They waste time here and the game still do not have to be fun. On the other side, players often do not entertain the game and it uninstall after the first experience. Our solution aims to help making games catchy right from the start with personalization. Game personalization is a way of making games experience more powerful for players. Key factor is adapt games according to player personality, information from mobile devices and other player preferences.

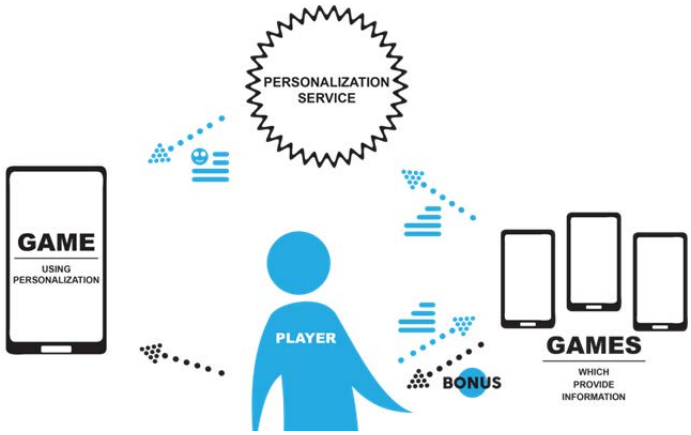


Figure 1. Fundamental principle of the proposed solution.

The fundamental principle of the proposed solution is illustrated by a diagram in the Figure 1. There are four elements making the personalization to work: an end user (player), games collecting information via questionnaires, personalization service and a personalized game. Players provide their data while playing their favorite games, which send the information to the personalization

[*] Master degree study programme in field: Software Engineering
[†] Master degree study programme in field: Information Systems
Supervisor: Peter Krátky, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

service. In return, player is rewarded in the form of a bonus within the game. Personalization service processes the information received from games and creates players' profiles in the database. When the player starts a game that uses personalization service, the game requires the player's profile from the personalization service.

To build the database of user models we use the power of crowdsourcing. We provide an API that could be easily integrated into any Android application as a monetization alternative to ads. The API shows questionnaires with basic questions about player's gaming preferences and various improvements for the completion of these questionnaires are offered in return. Of course, the developer earns money too. The personalization service processes the questionnaires, updates the user model or creates a new entry for the player identified with a unique ID across all the games with the API implemented. We provide multiple sets of questionnaires asking for the Big Five personality (10 items questionnaire split into two sets [1]), Bartle's player style [2], favorite color, gender, age and so on. Another (immediate) way of building user model is to obtain data from a mobile phone, such as games that the player played in the past to get favorite game genres. The API provides then a simple way of obtaining the user model right at the start of the game so the adaptation could start.

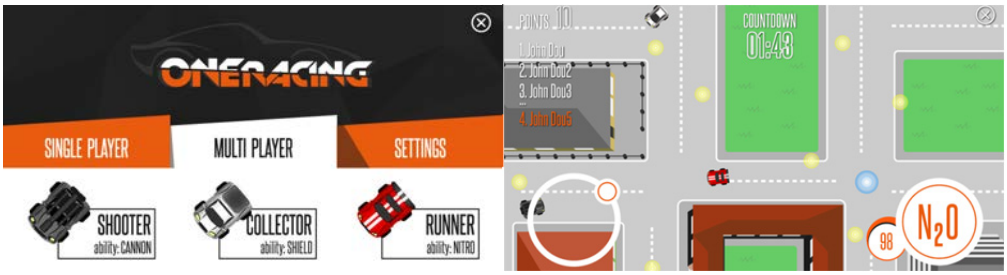


Figure 2. OneRacing – game screenshots.

To verify the proposed personalization service, we designed an adaptive game, OneRacing, (see Figure 2), with an ambition to attract players right from the start. The principle of the game is to control cars around the map and undertake specific tasks. The basis of such a game is a module that estimates the profile of the player thanks to the information from the mobile device. According to the personality profile, the nature of the game specific for the player could be adjusted and thus it could increase its attractiveness. The different nature of the player game is confronted in multiplayer mode. The game includes 3 types of characters (shooter, collector and runner). They are recommended based on favorite game genres. The game provides also several views for four Bartle's player types. Killers see a rank list according to successful games, explorers have possibility to unlock powerful upgrades, achievers reaches badges and socialisers are ranked by number of successful guild quests. Another purpose of OneRacing is mentioned verification of the services. It would include the possibility of completing a questionnaire and would contain mechanisms for retrieving data from the user.

References

[1] Bartle, R.: Hearts, clubs, diamonds, spades: Players who suit MUDs., The Journal of Virtual Environments, 1(1), 1996, Available online: http://mud.co.uk/richard/hcds.htm (Last Access: 19.05.2015)

[2] Rammstedt, B., Oliver, P.: Measuring personality in one minute or less: A 10-item short version of the big five inventory in English and German, Journal of Research in Personality, 41(1):203–212, February 2007.

Happen: Event Based Social Network

Michal CIHÁK[*], Michal GAJDOŠ[*], Lukáš MASÁR[*], Pavol MICHÁLEK[*],
Vladimír OSVALD[†], Matúš PIKULIAK[†], Tomáš SÝKORA[*]

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
tp12calendar@gmail.com

We live in rush times, in which it is often very difficult to organize a meeting for several people in suitable time and place. However, we live in the age of the modern technologies, which should help us to live our lives in a comfortable way. Expectations of nowadays users of IT services and products are continually increasing as fast as the technologies evolves. Success of an application is not only in its incomparable functionality but also in the users' experience itself. Together with enjoyment of usage these are the main factors that affect whether the application becomes popular and successful.

There are many tools that provide various options of time management. In our opinion there is one big problem about such applications, which is their design. We find it unattractive and not really user-friendly in many cases.

Happen is the right solution of the problems connected with time management. This innovative web application, allows users to organize meetings, using the modern web technologies. The main advantage of our application is not only its functionality that covers and even extends existing solutions but also contemporary and modern design that support intuitive use of the entire application.

We looked at the problem of time management and event organizing from different point of view than existing solutions do. The base of whole application is event which may be any action user decides to share. We focus on life cycle of event, as it starts with suggestions of times and places, and goes on with refinement. Its creation is very simple and quick and it is not even necessary to know and set time of the event, neither the place. The only attribute that has to be set is name of the event. When the event is created, invited users may add suggestions of time and meeting-place and vote for suggestions they like.

There are also some advanced options, which may be configured by creator of the event like restrictions of voting, adding new participants or suggestions about time and place.

Happen displays time suggestions together with already planned events to make it easier to decide about the suggestion. It may take a while and many iterations until few people make agreement on time and place of the meeting. We simplify this process making it intuitive with no need of long texting.

There is automatic mechanism that evaluates all the votes. Application offers the list of the best options for the time and the place of the meeting based on the users' votes. After that creator of the event may choose one option from that list or choose any other suggestions that has been

[*] Master degree study programme in field: Software Engineering
[†] Master degree study programme in field: Information Systems
Supervisor: Lukáš Turský, Institute of Applied Informatics, Faculty of Informatics and Information Technologies STU in sBratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 425–426.

added to the event by the invited users and close the event, so that all the invited users will be notified about the final time and the place.

Though, the main purpose of the Happen is not the complex time management of individuals. For suchlike it is often sufficient to use simpler calendar or some of many time and task management applications which tries to increase your productivity by saving the time. Of course, Happen can also help in such ways. But the idea we try to fulfill is to help people synchronize and link their time among themselves. To be able arrange a meeting by simply clicking the proper time on their profile or create easily an event for many people without even knowing the right time or place, so the people can decide together. To be able to see what is happening around yourself, your friends or your favorite spots and to easily join all these events together with your friends.

Thanks to the well-designed core of the application, it is easy to plug extending modules, which can serve the functionality in many different environments, such as reservations systems in restaurants, or internal needs of companies like assigning shifts.

We are working on three versions of Happen, which are all suited for needs of specific type of event and customer.

Free version – this version is mainly focused on common users. The version is available for anyone. It includes a user's profile page, through which user browses through his events as well as the events and the calendars of other users. This version supports creating and managing events as it is described above.

Business version – in business version we aim for potential customers, who would be interested in getting closer to their customers by sharing and suggesting their own special events. It will still be about creating the events, but these events will serve as offers or invitations to the other events. For example, this could be an offer for bulk discount in restaurant for users that want to meet up. In this version we are able to provide functionality covering reservation system for restaurants, accommodation companies and many others.

Enterprise version - deals with time management and is tailor-made for specific needs. It contains calendar which is supposed to be used for managing human resources. For example, it may be used as a system for assigning shifts. Thanks to the core of the application which is used for all versions, the records can be saved in the profile page of the employee who is user of our application. After that he will be able to see his shifts in his personal calendar alongside with any other events in his profile. Last but not least, enterprise version can be also adapted and deployed at the customers' office.

Automated Exploration of Citation Data for the Purpose of Assessment Factors Acquisition

Martin CIBULA[1], Marek ČERNÁK[2], Vladimír DEMČÁK[2]
Ján KRIVÝ[2], Jozef MELKO[2], Martin ŠUSTEK[1*]

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
fiitteam15@gmail.com

Science is one of the many areas in the contemporary life that requires processing of large amount of information. Moreover, if this processing is carried out in a heterogeneous environment, the process becomes even more complicated. It is essential to obtain the important information and to get a general overview of the papers. Existent citation indexes include bibliographical data which function as a source of feedback on the publications.

The process of obtaining data from multiple sources is a lengthy and not always entirely successful task. From the perspective of the authors of the individual works it is important to get as punctual and up-to-date data regarding the success of their papers as possible. Moreover, they need to dedicate the time they spend searching for feedback to more important activities.

We created a system for retrieval of the information from various sources, that is able to analyze and interconnect them. This allows us to make the obtaining of feedback much easier for the researchers and save them significant amount of valuable time. System also offers database of unified bibliographic entries that contains only accurate and up-to-date results, which are retrieved in real time. As there is no such system at this time, or at least not in service, our system represents a unique proposal.
Key system properties:

- *Data collection* – system's most important property is to facilitate the retrieval of feedback to the user's published works. Our system identifies and obtains the sources from external systems (primarily WOS and SCOPUS), analyzes them, saves the obtained data and analyzes the reactions to selected publications,

- *Automatization* – after the record of the entry to obtain feedback the system saves the information about the requesting part and from that moment on it updates the feedback regularly. This process allows the feedback on the works of various institutions to accumulate in our system, which enables the institutions to get the feedback later on directly from our database,

---

* Master study programme in field: [1]Software Engineering, [2]Information Systems
Supervisor: Dr. Nadežda Andrejčíková, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 427–428.

– *User interface and API* – our next aim was to offer the user an interface in which the request could be entered easily. Since the user may not be able to control the communication data exchange format (JSON), we decided to implement a web interface with simple interactive elements. For any other system, our system will provide RESTfull API to communicate,

– *Import of the researchers' profiles from external systems* – the possibility to import the profiles of researchers from already existing academic systems simplifies the access to rankings and in case of demand from the users' side provides various metrics to evaluate their success based on the works published.

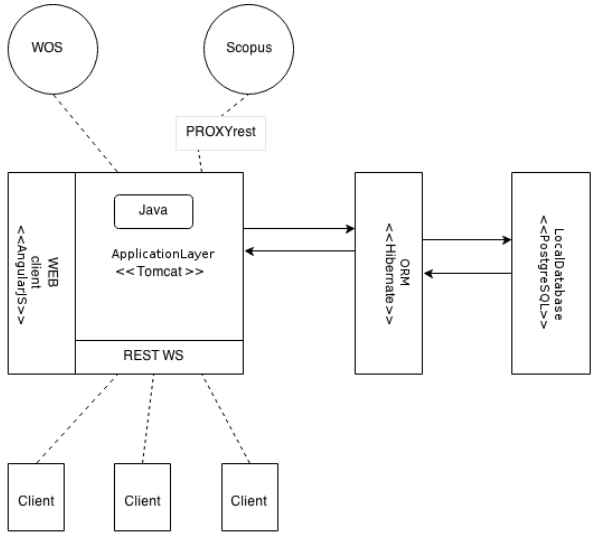


Figure 1. System architecture design.

As Figure 1 demonstrates, system is implemented as a JAVA web application providing the REST API for communication with clients. By providing this interface we facilitate a simple expansion of the system to other clients or a simple provision of the system's services to other systems. Flow of request processing and record parsing is driven by centralized configuration and can be therefore simply modified on-demand. System is also equipped with a minimalistic user interface which can be, if appropriate, modified into widget, that can be integrated into the client-side system.

APIs of particular citation indexes are used for data retrieving. The retrieved data is consequently saved and stored in local database. This makes it possible to ensure immediate access to this data by the following requests. The database also stores catalogue of institutions with access to system. Informations about works of authors published under these institutions are being updated on regular basis.

It will be possible to extend the system based on the incoming users' requests by constructing our own database. This way, the system has a potential to be extended by other features. Using its own local database, the system is able to access the documents faster in comparison to requesting the citation indexes by adoption of new requests. System also provides direct access to publishable parts of authors' works in form of fulltexts.

Web Interface for Distributed Platform for Volunteer Computing – BOINC

Pavol ČURILLA[1], Patrik GALLIK[1], Martin KAŠŠAY[1]
Matej KLOSKA[2], Juraj KOCHJAR[2], Roman ROŠTÁR[2*]

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
tp07-1415@googlegroups.com

Nowadays, in the era of vast amount of online data, researchers are more likely to face the problem of processing outputs of their work. Despite the growth of computing power, there is still a lack of resources for handling and processing data currently used in all areas of research. Following the idea of scaling out – distributing challenging tasks to more participants, University of Berkeley created an open source platform for distributed computing – BOINC [1, 3]. The aim is to share, mostly unused, computing power of common devices to deliver a meaningful research results by collecting partial results of BOINC's project participants. The more users participate the sooner and quicker are the valuable results delivered.

As mentioned, the power of distributed computing system is made by the community of its participants. The more people you want to involve working on your project, the more requirements are related with usability of work distribution system you provide to them and the harder it is to motivate the end users to participate. Default BOINC installation brings a basic web interface, which may not be suitable for every end user. The aim of our system is both to spread the word about the BOINC platform and involve people to participate in the faculty projects by creating web interface easier to use in the context of user experience. We want to contribute to BOINC open source project in the form of brand new accessible web interface for BOINC applications administration.

The analysis of the default version of web interface points out several user experience issues, namely:

- difficult sign up process,
- insufficient analytical interface for project administrators,
- no interface for task submission,
- missing notifications for key computation events such as no volunteers for computing, all task completed etc.

Currently, only few users can submit tasks [2] for computing to the BOINC server because a direct SSH connection with proper credentials is required and the task submission process overall is rather

* Master study programme in field: [1]Software Engineering, [2]Information Systems
Supervisor: Dr. Peter Lacko, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 429–430.

difficult for an inexperienced user and demanding in terms of required skills. We want to transform the whole process of computation task setup to an easy, zero-time consuming procedure in the form of a user-friendly web interface. The easier task submission the more tasks will be available for computation and project will be more attractive for volunteers.

Figure 1 shows system architecture with respect to the new interface which is completely redesigned and developed with RESTful API in mind. BOINC PHP back-end core is replaced with modular Ruby on Rails REST API code. Entities in the picture describe system core technologies from the existing functionality point of view and the new one.

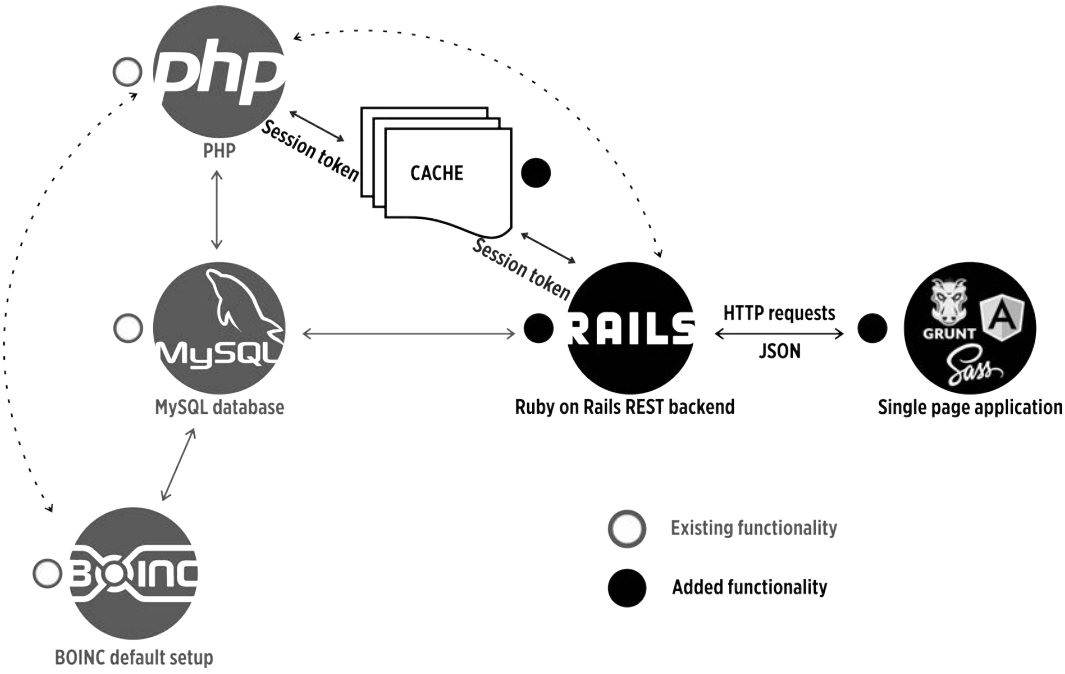


Figure 1. Architecture overview of the proposed solution.

We implemented functionality for a quick scan of the computing tasks for security threats, such as memory leaks or malicious activity attempts hidden in the submitted tasks. This feature protects contributors from malicious software, possibly distributed using BOINC infrastructure.

Over time we would like to deliver a newer, cleaner and more usable version of web interface which will have an extended analytical features for project administrators that will deliver the usage statistic data in an easy to understand form, thereby helping BOINC project administrators to understand and monitor utilization of their projects.

References

[1] Anderson, D.P.: Boinc: A system for public-resource computing and storage. In: *Grid Computing, 2004. Proceedings. Fifth IEEE/ACM International Workshop on*, IEEE, 2004, pp. 4–10.

[2] Ujhelyi, M., Lacko, P., Paulovic, A.: Task scheduling in distributed volunteer computing systems. In: *Intelligent Systems and Informatics (SISY), 2014 IEEE 12th International Symposium on*, IEEE, 2014, pp. 111–114.

[3] University of California, B. `https://boinc.berkeley.edu`, [Online; accessed February 19th, 2015].

Vehicle Tracking System for Cargo Transportation

Jakub DLUHÝ, Ľubor FERENCZY, Maroš KRAJČOVIČ, Tomáš PETRÍK, Michal VALIČEK*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova2, 842 16 Bratislava, Slovakia
Fleetteam7@gmail.com

There are many practical reasons why it is very popular to use vehicle-tracking systems in companies. One of the most significant is the process automation of paperwork required by law, for example driving log. On the other hand such systems are able to monitor vehicle maintenance status and save money for repairing and optimizing their operation [1]. There are plenty of variables to observe in company fleet with the aim to save resources and enhance affectivity, which create demand of consumers. That is why many concurrent systems are available on market, from basic ones suitable even for individuals to complex ones. Majority of them are focused on passenger cars or have general purpose without specific target features. One of those systems has been developed at Slovak University of Technology and has been presented in IIT.SRC 2014 conference [2]. We have decided to create such system for specific field – cargo transportation.

We have analysed 5 fully deployed and commercially available solutions and made interviews with dispatching operators, in order to design a robust system with all features needed in this area, fully gaining benefits from installation of tracking hardware into fleet vehicles. We have implemented basic, must-have features, such as:

- real-time tracking of vehicles,
- driving-log preparation in respect of actual law,
- gaining and logging actual and service data from vehicle CAN bus,
- enabling communication between dispatching center and driver or driver group,
- tasks – routes assignment to drivers and vehicles by operators.

Those features are supporting all basic and necessary routines needed in the cargo transportation. We have also identified more specific features making the system even more reliable, such as:

- supporting sensor network with defined, universal interface to connect any type of sensor and logging subsystem of periodic and aperiodic events,
- merged system of traffic events notification gaining data from more sources, including own company-specific database of events, points-of interest, etc.,
- driving style analysis to identify safe, effective, economically-driving employees, without penalization of the ones, who drive more often in city, fully loaded or in traffic jams [3],

* Master degree study programme in field: Computer Engineering
Supervisor: Dr. Peter Pištek, Institute of Computer Systems and Networks, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 431–432.

- to provide a tool for operator to insert marks to driver´s map in order to help him get to the destination in the real time.

Those are functional requirements on the system with good chance to compete in the difficult market field. Last, but not the least important are the non-functional aspects of system, such as high reliability and independent operation, user-friendly interface to help drivers stay focused on their work, straight-forward interface for dispatching operators with clear and all important data to help them make decision fast and effective according to actual situation, modularity and scalability of system in order to handle company-specific needs.

To stay platform-independent, have implemented the device placed in a vehicles on generic Android tablet, thus in future the client is free in decision of choosing specific product, starting in portfolio of single-board computers hidden in a vehicle, continuing through Android smartphones or tablets and finally Funtoro devices, which are certified to be safely used in drivers cabin.

Key features of proposed system

During analysis we have not found deployed and commercially offered product with these features. Key attributes of proposed tracking system are universal sensor network, unified warning system and advanced driving analysis.

Even more it is possible for driver to access navigation – Google Maps, since it is crucial for him to be informed about actual situation on the roads. That is why we have brought unified system to warn drivers about significant events on the roads ahead. Our server is collecting data from more sources and serves them through tablet or smartphone directly to the driver in order to help him to choose the best path to the destination.

The sensor network is created by wired interconnection of microcontrollers based on AVR architecture. Each microcontroller is operating one or more close sensors and communicates by unified protocol, defining the type of measured physical variable and its value. Each trail can be equipped with basic set of sensors checking environment of transported cargo. There is always a need to register shocks (measured by accelerometers) and all events when the trail doors were open and it was possible to access cargo (touch sensor attached to trailer loading doors). The data are recorded on server and it is possible to print a protocol of the transportation environment of any route, anytime by operator. In case other specific cargo is transported, it is possible to add any other analogue or digital sensor to network and log temperature, humidity, slope, etc.

For driving style analysis, we have used outputs of last research in field of applied statistics. We have chosen variables with the highest impact on driving economy and vehicle depreciation. Our statistic model is considering green band driving, harsh acceleration, harsh breaking and distance driven on cruise control according to long-term collected data by Scania [3]. The statistical model output is useful in rewarding drivers in order to optimize operating expenses.

Our goal is to implement tracking system supporting special needs of cargo transportation business. With above mentioned features and key attributes, especially unified traffic information, sensor network and driving analysis, the system could be unique product on the market field of tracking systems for cargo transportation companies.

References

[1] Gadri, R.Ch., Chavan, A., Sonawane, R., Kamble S.: Land Vehicle Tracking Application on Android Platform. In: International Journal of Engineering Research and Applications (2012),vol. 2, issue 3, pp. 1978–1982.

[2] Onder, J., Polák, M., Trávniček, T., Urbán, D., Zemaník, L.: Application for Funtoro Platform. In Bieliková, M., eds.: IIT.SRC 2014, Bratislava, April 29, 2014, pp. 625–626.

[3] Viswanathan, A.: Data driven analysis of usage and driving parameters that affect fuel consumption of heavy vehicles. Master's thesis, (2013).

Crowdex: A Crowd Work Management System

Dušan CYMOREK[1], Peter GAŠPAR[1], Vladimír ĽALÍK[2]
Michal POLKO[1], Miroslav ŠAFÁRIK[2], Slavomír ŠÁRIK[1]
Štefan ŠMIHLA[1*]

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
fiit.tp05@gmail.com

Finding enough and the right people to perform small, short-term and easy tasks – process referred to as crowdsourcing – is often difficult and time-consuming. Yet, students at universities, their teachers and researchers face the problem of testing their hypotheses on adequate number of people. Companies often need to carry out small tasks which they do not need to contract, yet they can still offer some kind of a reward to those completing them. Without the support of information systems, crowdsourcing often results in calling the closest friends, acquaintances or finding strangers online on social networks or internet forums and asking them to carry out the specific tasks. This, however, brings up some risks. To mention only a few, it is uncertain whether enough people can be found and more importantly – whether the people found are truly competent to the tasks specified. Another risk is motivation without which the quality of the results might be questionable.

Crowdsourcing and crowd work industry have become more and more popular over the last 10 years with a large growth potential in the future. There are several crowdsourcing platforms to this date, each of which specializes in a different area and has its own pros and cons. The most popular one is Amazon Mechanical Turk, which provides marketplace mostly for Human Intelligence Tasks and information retrieval. Despite its popularity, Mechanical Turk does not include user/author rating or discussion about the tasks without the use of third-party browser plugin or forum. Another popular system is Crowdflower data collection and enhancement platform focused primarily on business customers, which can be quite pricey for students. Other systems include domain specific systems like TaskRabbit, which provides help with cleaning, assembling furniture or moving, or Gigwalk specializing in retail execution.

Our goal is to make crowdsourcing easier, more reliable and less time-consuming for students, researchers or companies. Therefore we have developed our own crowdsourcing system from scratch – Crowdex – which we describe in this paper. Our system addresses some of the aforementioned issues, mainly the lack of user rating and feedback and availability of such systems to students. To keep the system as simple as possible, we have defined a lifecycle of the task, dividing it into four different steps: (1) Creating a task, (2) Obtaining participants, (3) Task realization and (4) Task evaluation.

Firstly, a user creates a new task simply by giving it a concise name and description and by setting up its parameters afterwards. These include user requirements (e.g. age or specific skills),

---

* Master study programme in field: [1]Information Systems, [2]Software Engineering
Supervisor: Dr. Michal Kompan, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 433–434.

specifying the start and finish dates, uploading a file attachment (e.g. application to be tested) or even creating custom questionnaires which gives the author ability to ask several types of questions including selectable, multiple-choice or fillable inputs. To make the results easier to evaluate, our system provides an option to view them in a form of graphs and aggregated statistics. If desired, results can be downloaded as a CSV file. After the task has been published by its author, it is made available to the other users of the system and the second step (obtaining participants) may follow.

To leave the decision of who to accept to complete the task to its creator, other users are only able to apply for it. Eventually, the creator decides whether to accept or reject the applicant. To make the decision easier, Crowdex offers several ways how to learn more information about potential participants. The most significant are user profiles which in addition to the basic user information also provide their reputation and skills. Skills characterize user knowledge and abilities which can be used to tell if the task is suitable for user. In addition to standard Crowdex account we also support login via Academic Information System of the Slovak Technical University. In this way, retrieval of user's skills is not only automatic but reliable too.

Once the task realization step begins (determined by the start date or by meeting custom requirements), approved participants are expected to complete the task by following the author's instructions. After the task completion users are evaluated and rewarded by the author. Our goal was to make a reward system as transparent as possible therefore the author has to decide which results he will accept or reject. Shall the author accept the results, he has an option to reward the user with Credit – our own virtual currency – which can be redeemed for various items (vouchers or promotional items such as pens, key rings, etc.) in a built-in e-shop. We believe that with this system, most of the users will be satisfied with the obtained reward and will be motivated to regularly use our system and participate in various tasks.

To make the features described above widely available and to keep the system platform-independent, we have decided to develop it as a web application powered by Ruby on Rails framework. We have also focused a lot of our attention to the graphical user interface – we have used well-known frontend frameworks like Bootstrap and jQuery and implemented some of the Google Material design principles. This enabled us to bring responsive and flexible design to the most of users and their devices.

User profile information along with the data of user's activity could be further used to implement a recommendation system which could not only recommend potential participants to the task authors but also recommend specific tasks to the users themselves. To extend support for multiple task types, we plan to enhance Crowdex by implementation of eye-tracking technologies. With the support of eye-tracking, not only results, but also the process of solving the task can be evaluated which is not used in any of today's crowdsourcing systems.

The main contribution of our work is the new innovative approach in crowdsourcing with a great emphasis on user's motivation to participate. We believe this is the key for successful and popular system, enjoyed by both participants and task authors. During the development of our system we have taken the openness and ease of use of Amazon Mechanical Turk and precise and reliable data collection of Crowdflower. We have then further improved it with built-in user rating and feedback and custom e-shop for redeeming rewards. In the future we plan to incorporate recommendation system and eye-tracking technologies, neither of which have been used in crowdsourcing systems before.

References

[1] Bederson, B.B., Quinn, A.J.: Web Workers Unite! Addressing Challenges of Online Laborers. *Human Factors*, 2011, pp. 97–105.

[2] Kittur, A., Nickerson, J., Bernstein, M.: The Future of Crowd Work. *Proc. CSCW '13*, 2013, pp. 1–17.

DriVR – Vehicle Controlled with Help of Virtual Reality Head-Mounted Display

Martin HABOVŠTIAK*, Erik MATEJOV*, Tomáš MORVAY*, Albert PRÁGAI*, Peter PROTUŠ*, Miroslav SIRO*, Natália TRYBULOVÁ*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
teamtechnologic@googlegroups.com

The concept of virtual reality has been known for several decades, despite the fact that public began to be more interested in this technology after the year 1990. Many years gone by since then and technology advanced by leaps and bounds. Nowadays, virtual reality is a rapidly growing technology which utilises the ever-increasing power of computers to simulate real-world and imaginary environments and situations with a high degree of realism and interactiveness [1]. Simulated presence in real world environment has wide range of utilization and it is important to focus on this area.

The aim of this project is to create remote controlled model of vehicle operated with help of virtual reality device. Our proposed approach, which represents images of the real world as 3D environment in virtual reality device, has many advantages. These advantages predetermine this approach to be successful in various areas.

Many police departments and military sections might use remotely controlled vehicles to increase operators' safeness at bomb or mine disposal. Vehicles can be also involved in searching for people stuck in ruins or unreachable places. Moreover, driving this vehicle in virtual reality has great potential in entertainment industry. These days scientists at Carnegie Mellon University work on a project with an attempt to land on the Moon using remotely controlled vehicle and the Oculus Rift[1].

Presented solution can enhance operator's spatial awareness, simplify control of vehicle and provide 3D image of inaccessible space for operator as if he was actually there. Figure 1 roughly describes our solution.

The *vehicle* is equipped with two moveable cameras which are used to capture stereoscopic image required for binocular vision. The images from both cameras have to be synchronized. The vehicle receives control commands from base station. Control commands are processed by single-board computer Raspberry Pi which determines the direction of vehicle and cameras. The vehicle also transmits compressed video data and optionally sensory data. WiFi is used for wireless communication in both directions.

User wears the Oculus Rift *virtual reality head-mounted display*, which displays processed 3D image from both cameras. By moving his head the user controls the position of both cameras

---

* Master degree study programme in field: Computer Engineering
Supervisor: Roman Broniš, Institute of Computer Systems and Networks, Faculty of Informatics and Information Technologies STU in Bratislava

[1] Robot will beam live Moon pictures to Oculus users, http://www.bbc.com/news/technology-29704953

simultaneously. The direction of vehicle is controlled with gamepad. We have used standard XBOX 360 gamepad.

The *base station* is the central point of the system. It processes and evaluates all the data. It transforms signals from control devices to commands for vehicle to move. It receives and processes the video data from cameras. Base station renders split-screen stereo image with distortion correction for each eye, what is required by Oculus Rift[2].

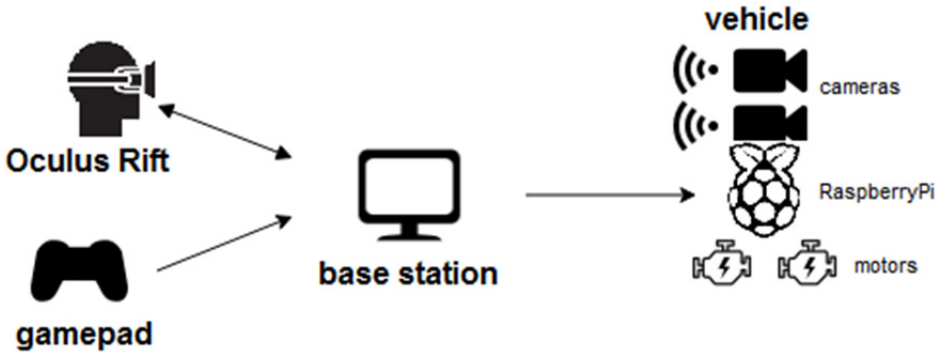


Figure 1. Main components of DriVR.

The project aims to provide easier control of vehicle and enhance operators spatial awareness. In addition to a number of previously mentioned areas of application, contact with new and perspective technologies is also an important contribution of the project. For example Oculus Rift device, which is used, has not been officially released on the market yet. Moreover the connection of a virtual reality with material world is very interesting. Finally, a lot of students interested in this area of education, will have available documentations about project and they will be able to extend this project to other functions, such as adding augmented reality features.

There is similar solution presented by Jonas Lauener in IRE (Intuitive Rift Explorer) project[3]. Basic idea of controlled vehicle is similar to ours, however there are multiple differences. IRE project uses older Oculus Rift DK1, while our project works with newer DK2 version. Sensor data monitoring environment are also additionally implemented in our approach, but main difference is in the method of video transmission – IRE project uses Wireless HDMI standard, where our project uses standard WiFi transmission.

Acknowledgement: This work was partially supported by Tatra banka foundation under the contract No. 2014et005.

References

[1] van Wyk, E., de Villiers, R.: Virtual Reality Training Applications for the Mining Industry. In: *Proceedings of the 6th International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa*, (2009)

[2] Oculus Developer Guide, SDK Version 0.4, http://static.oculus.com/sdk-downloads/documents/Oculus_Developer_Guide_0.4.3.pdf

[3] Project IRE, http://www.jolau.ch/category/ire/

AdHunter – Outdoor Crowdsourcing of Visual Smog

Jana EGRIOVÁ[*], Alexander FERENČÍK[†], Richard FILIPČÍK[†], Tomáš MELICHER[†], Juraj SLAVÍČEK[†], Jaroslav ZIGO[†]

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
tp-tim-14@googlegroups.com

There is no regulation of advertising on the streets in Slovakia, which leads to spreading of visual smog. Visual smog is a pollution through a high concentration of advertisements in many forms, e.g. billboards, megaboards, citylights, etc. The visual smog is making the environment less attractive and unaesthetic. The society is subconsciously blind to the visual smog and that is the main reason, why this problem is ignored by the authorities and it is getting worse.

Opening the public discussion will make society to realise the scale of this problem. The complete mapping of the smog will help to acquire enough data to make detailed statistics. This statistics can be used to show the actual state of visual smog to authorities so they can begin to solve the problem globally.

Our project is based on an idea that we can make collecting all of these data easier. Crowdsourcing [1, 2] offers a suitable solution. The idea is to outsource task to a crowd of workers – public in our case. Through AdHunter mobile application (see Figure 1) user captures advertisements and provides more details about its owner, type and other information he can find out. Mobile application is suited for users who are not willing spend a lot of time finding extra data. On the other hand, our application is also prepared to be used by advanced users. People provide us information which they are able and willing to find out.

AdHunter web application (Figure 1) can be accessed from the user's computer. It displays collected data on the map of Slovakia. Through the web application user uploads new advertisements using drag and drop technique. Registered user has much more possibilities to interact with map. User is able to filter uploaded advertisements by himself, correct information or fix the position.

Once the data are collected, application processes the data. Using computer vision metrics, automatic recognition of the same advertisements reported by multiple users improves the collected results. Web application displays detailed statistics such as number of legal and illegal advertisements, money loss because of illegal advertisements and concentration of visual smog heat map.

Some organisations are currently active in solving the visual smog problem but it is hard and time consuming to get all information and, most importantly, themselves will never be able to collect such an amount of data that can be gathered by crowd easily. People from these

[*] Master degree study programme in field: Software Engineering
[†] Master degree study programme in field: Information Systems
Supervisor: Dr. Jakub Šimko, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 437–438.

organisations, our early adopters supply AdHunter application with the very detailed information which leads to continuous expansion of type of data that can be tracked about advertisements.

AdHunter is unique of its kind although there are some similar projects based on crowdsourcing e.g. message for the mayor (`odkazprestarostu.sk`). Odkazprestarostu project's focus is on problems of public areas in cities, towns and villages. People can add a problem they find or know about and it pinpoints the specific issues to the mayor and the responsible authorities. The fact that this project is already successfully used in many places in Slovakia makes us confident that people will be interested in using our AdHunter application.

Our success with AdHunter project in ReStart Slovakia competition proves that people are concerned about problem of visual smog and also they are interested in getting involved in the solution.

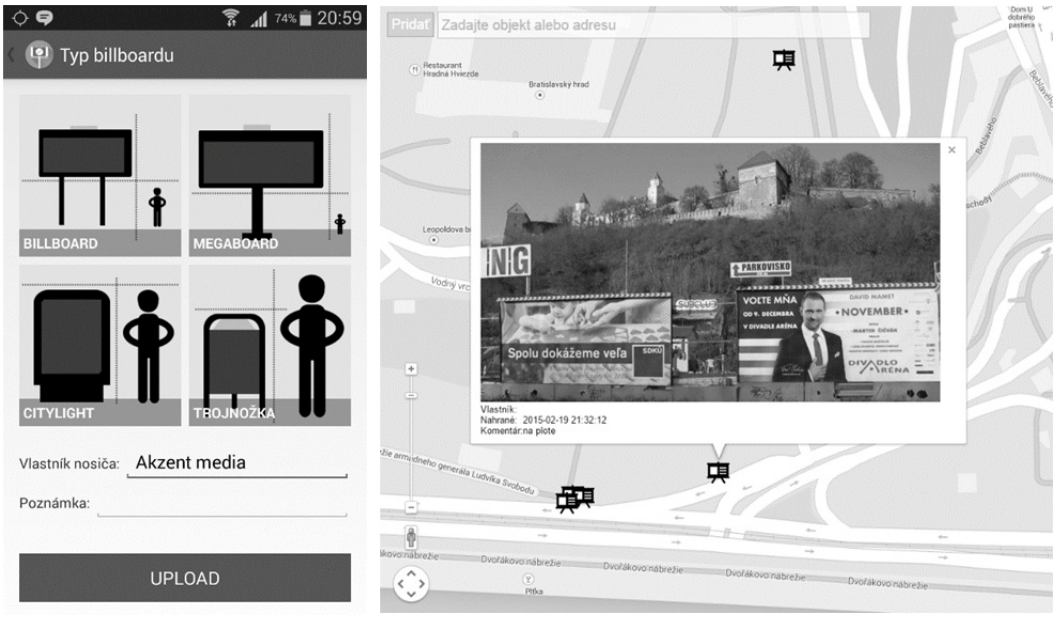


Figure 1. AdHunter mobile application (left) and web application (right).

References

[1] Seidel, C. E., Thapa, B. E., Plattfaut, R., & Niehaves, B. (2013, October). Selective crowdsourcing for open process innovation in the public sector: are expert citizens really willing to participate?. In *Proceedings of the 7th International Conference on Theory and Practice of Electronic Governance*, ACM. pp. 64–72.

[2] Chatzimilioudis, G., Konstantinidis, A., Laoudias, C., & Zeinalipour-Yazti, D. (2012). Crowdsourcing with smartphones. *Internet Computing, IEEE*, *16*(5), pp. 36–44.

Advanced Chess Tournament Manager and Pairing Engine

Igor JURÍK, Matej LISKOVEC, Jaroslav LOEBL, Martin TAMAJKA,
Peter TRUCHAN, Ľubomír VNENK*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
check-mates@googlegroups.com

Tournament of any game, no matter if it is the Olympics or FIDE chess tournament, is a very complex and complicated process. In chess there are steps that need to be followed to ensure transparency and fluency of a tournament: publish propositions, confirm player registrations, manage tournament and publish tournament results. Our goal is to help tournament organisers by simplifying the activities related to these steps. Using our product, they are able to perform all necessary actions in one place, whereby they do not even have to install any software – only thing they need is web-browser. On the other side, players want to see their results as soon as possible. With our product, they will see their results immediately, together with advanced statistics.

At the beginning of a tournament organisation, an organiser needs to create a proposition. It is a document containing information about the tournament (such as date, location, rules or other details). Problem is that there is no unified form. Some propositions are very messy and players tend not to notice the important information. Our product provides transparent and reasonable form of the propositions. At the moment, there is no system allowing organisers to create and upload the proposition at once. To publish the proposition, organiser needs to fill in a web-form with same basic information that the proposition already contains and also to attach created document. Once the proposition is published, players may pre-register in the tournament. Registration process is realized via email (published in the proposition). The organiser has to check each email manually and reply to each player, which is time consuming activity.

During the tournament, players need to be paired for each round. In pairing procedure FIDE-defined pairing algorithms are used [1]. Because (especially as far as large tournaments are concerned) it is barely impossible to pair players manually using appropriate algorithm, FIDE, world leading chess organisation, recommends using computer programs to perform pairing. At the moment, very few of supporting software products exist. In fact, just few tournaments are organised without Swiss-Manager, the most common software [2]. However, using Swiss-Manager is very complicated even for an experienced user. Pairing algorithm itself works fine, but the way of setting custom options is difficult, mainly because of chaotic organisation of settings dialogs. Our study (20 organisers, 67 players) shows, that about 50 % of organisers already tried another product. We believe there is a real need of new, better solution.

The goal of our product is to be useful for both organisers and players. As an organiser creates or manages a tournament, our product provides supportive tools to simplify the process.

---

* Master degree study programme in field: Information Systems
Supervisor: Dr. Marián Šimko, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 439–440.

Players, on the other side, are able to quickly filter tournaments they are interested in and pre-register to them just by one click. Each registered player has his own profile where he can see his past games, achievements and statistics. Player's profile may be connected to his/her official FIDE profile if this player is registered in FIDE. We believe that providing of such modern application helps chess organisers feel more comfortable with managing tournaments and players view statistics with ease. Currently most used product, Swiss-Manager, has also its own website for tournament results (http://www.chess-results.com/). There are stored results of each game of a particular tournament. It is difficult to find a result of a specific game or tournament without the direct link. Filtering tool used at the website is very difficult to use, even useless.

The aim of our solution goes beyond tournament organisation management. It involves several innovative features. Although it is not hard to find a database of past chess games, there are no statistical data of games at the moment. However, this data is very important for players. They may be used for personal training and preparation before a match. Analytical study of player's opponent gives a great advantage and allows to prepare special tactics. We plan to use our own data and third party data to provide the desired analytical analysis of players to make the pre-game preparation more interesting and helpful. To support the importance of such utility, our study showed that more than 80 % of players prepare rather extensively for their next opponent.

We perceive these problems as an opportunity to create an automatic and sophisticated software that allows to manage tournament on high level. Our product is unique because it provides tools for every stage of tournament management, simplifying every one of them. In contrast to Swiss-Manager, which is a desktop application, we offer much more intuitive web application with user friendly interface.

Figure 1. Application screenshot.

Our product has potential to become one of the most used tools as far as activities related to chess competitions are concerned. In the future, our plan is to deeply analyse collected tournament data. We want to develop advanced modelling techniques to model players according to their games and past activities, which helps us to better understand what makes a good chess player. These findings have potential to be used for educational purposes. In the near future, we will use computer vision for fully automatic recording of chess game. Also there are many other games similar to chess (e.g., scrabble) that may fit to our system.

References

[1] FIDE. 2015. FIDE Swiss Rules. Available at: http://www.fide.com/fide/handbook.html?id=18&view=category. Accessed on 2015-03-21.

[2] Herzog, H. 2015. Swiss Manager. Software tool homepage. Available at: http://swiss-manager.at. Accessed on 2015-03-21.

Beacon Indoor Localization

Filip MAZÁN*, Veronika OLEŠOVÁ*, Filip ŠOLTÉS†,
Michal KUČERA*, Daniel PRIBUL†, Jozef KARAS*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
tim02.tp2014@gmail.com

People have problems with orientation in a new environment. Despite the fact, that most buildings have navigation boards, people are not able to find their destination and feel lost. Our solution helps these people – we are developing a method for localization, which will be tested in our application. It offers various information about activities within the building.

The application is currently adapted for one building – our faculty (named Virtual FIIT), but can be easily transferred to any other – business centers, shopping centers or hospitals. The main group of users are fresh students, but can be also any student or employee. Thanks to previous teams of developers, the application offers lots of functionality including viewing maps, personal timetable, canteens in the surroundings or bus departures. Our goal is to maintain these features, improve them and expand them to fulfill the needs of the users even better.

The previous development team deployed the application using PhoneGap technology. Their chosen technologies and programming techniques caused a lot of problems and slowed down the development. Therefore we decided to rebuild this application from scratch. We chose Ionic framework, which is suitable for developing multi-platform smooth running mobile applications. It utilizes AngularJS and CSS extension Sass and makes the development more efficient. The insight of our new design can be seen in the Figure 1a).

The most promising feature of our application is the possibility of localization of the user within a building. GPS technology can be used in open areas, but the satellite signal is too weak to reach insides of the buildings. This can be solved by different approaches (e.g. using Wi-Fi signal), but they are not always sufficient [1, 2]. In our project, we use Bluetooth LE technology utilized by Bluetooth beacons – tiny transmitters of low energy signal located in the building. Existing solutions [1, 3] using this technology are not precise enough – the average error is overall bigger than 3 meters. This is caused due to the nature of 5 GHz signal on which Bluetooth operates – the signal is easily interfered with and its strength fluctuates in time. This is why we decided to research this area ourselves and come up with another solution.

At first, we tried to localize the user by using trilateration algorithms. This algorithm did not work well, since it is needed to precisely estimate the distance of the Bluetooth receiver from the beacon transmitter. We found out, that the signal fluctuations in the building are too strong to estimate the distance, yielding errors more than five meters. These high errors practically made the trilateration useless. The next approach we tried is to utilize machine learning algorithms to be able to estimate user's position based on the existing studies [1, 4]. We used simple multi-

---

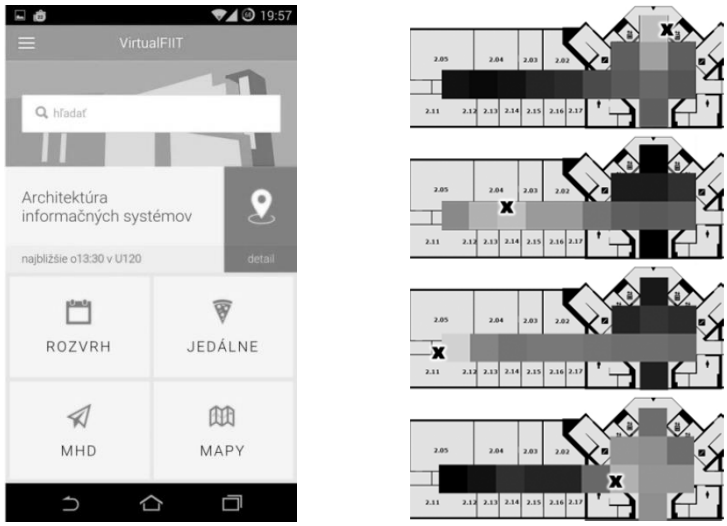
* Master degree study programme in field: Information Systems

† Master degree study programme in field: Software Engineering
Supervisor: Dr. Alena Kovárová, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 441–442.

layered neural network, which was trained on gathered data from 4 Bluetooth beacons located in one part of the building of our faculty. We gathered 100 samples of received signal strength (RSSI) from 14 locations 2 meters from each other. Received signal strength (RSSI) and positions of beacons are shown in a heat map in the Figure 1b). Since the gathered data correlates with transmitter locations, it can be used to train a neural network. We used backpropagation training algorithm on a feedforward neural network with 4 input, 7 hidden and 2 output neurons. We trained it on 105 samples, but have not had more samples available for proper validation. The achieved average error is 3 meters. Although this number may seem high, it is very promising. Next network improvements should lower the error significantly.

Our vision is to navigate the user to his destination based on his current location in the building. If the error would be small enough, it could serve the visually impaired people as an aid or could be used for automatic localization and navigation for moving machines.



a) *b)*

Figure 1. a) Main screen of the application, b) Four Bluetooth beacon measurements of their signal strength and their position marked by X on one floor of the building.

References

[1] Oksar, I.: A Bluetooth signal strength based indoor localization method. In: *Proc. of Int. Conf. on Systems, Signals and Image Processing (IWSSIP)*. IEEE. 2014. pp. 251–254.

[2] Li, X., Liu, X., Liu, K.: Guoguo: enabling fine-grained indoor localization via smartphone. In: *Proc. of the 11th annual int. conf. on Mobile systems, applications, and services*. 2013, pp. 235–248.

[3] Lu, T. et al.: Fusion of Wi-Fi and Bluetooth for indoor localization. In: *MLBS '11: Proc. of the 1st int. workshop on Mobile location-based service*. 2011, pp. 63–66.

[4] Ahmad, U. et al.: In-building Localization using Neural Networks. In: *Proc. of IEEE International Conf. on Engineering of Intelligent Systems*. IEEE. 2006. pp. 1–6.

Eyeblink – Evaluation of User's Eye Blinking Frequency

Pavol POLJAK*, Andrej PILIAR†, Lukáš PUTALA*,
Milan SMOLEŇÁK†, Dominik MELAGA†

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
weare11@googlegroups.com

Blinking is important for health of our eyes. Each blink renews the tear film which moisturizes and washes the eyes from dust and microorganisms [1]. Blinking frequency is influenced by currently performing activity. It is higher during conversation than during reading or at rest, at evening than during daytime, when ambient humidity decreases or when pollution of air increases. For eye's blinking is especially dangerous the work in front of computer display. Most people blink 10–15 times per minute. However the rate of spontaneous eye blinking while using computer reduces significantly (up to more than 50 % reduction). These people often suffer from dry eye having symptoms, as dryness, redness or burning. These symptoms then may indicate the Dry Eye Syndrome [2].

Purpose of this project is to help people with the Dry Eye Syndrome and at the same time connect the research between computer vision and medicine. The aim of the project is to develop application, which could eventually help people to prevent and cure the Dry Eye Syndrome.

In our project we measure the eye blink frequency using webcam, commonly accessible hardware and computer vision algorithms. During the recognition, webcam is pointed to the user's face to monitor his eyes.

Blink frequency can be also influenced by display brightness. The more light a monitor emits in an environment, the more pressured are the eyes. Therefore it is recommended to adjust the display brightness. Our application makes possible to automatically decrease the brightness of the display. Thirty videos were recorded in different light conditions while the screen brightness was adjusted to be comfortable and pleasant to the user. Within these videos an average brightness is calculated for each of them. These measures are further used to find the proper function which manages the monitor brightness settings. Input is the average brightness of the image captured from the webcam and the output is the value of comfortable brightness setting. The value can be shifted by the user settings to match his preferences.

We have developed a test application, which contains various types of notifications. We have divided them into two groups; the interactive and those that are non-interactive. In the group of the interactive notifications belong the blur notification and a notification that changes the brightness of a display.

---

* Master degree study programme in field: Software Engineering

† Master degree study programme in field: Information Systems
Supervisor: Andrej Fogelton, Institute of Applied Informatics, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 443–444.

Into the group of the non-interactive notifications belong the pop-up window and a notification that shortly flashes in colours of black and white. The most significant asset of the application is to verify the effectiveness of individual notifications without user's consciousness and to create the basic logic of the final application.

Our application detect the individual eye blinks and calculate the eye blink frequency. If the frequency is lower than the frequency of the healthy human, application notify the user in a noticeable but non-intrusive way. These are dismissed upon detecting an eye blink of the user, which help to not disturb his work. In this way we try to teach the user to blink sub-consciously and regularly to prevent and heal in a case of already having eye health issues, the Dry Eye Syndrome. Each user can review his statistics directly in the application in the form of a graph.

Our application is implemented in C++ using OpenCV library and MVC architecture. It consists of three main parts:

- *Model:* algorithm for face and eye blink detection,
- *View:* graphical user interface in Figure 1,
- *Controller:* business logic of application.

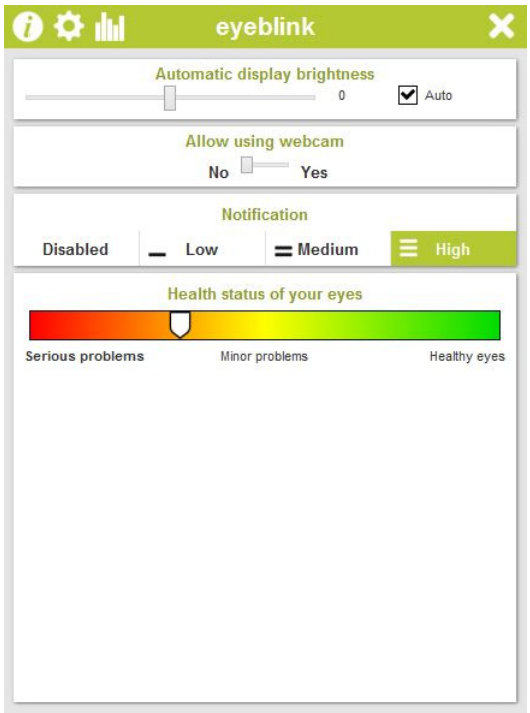


Figure 1. Eyeblink user interface.

References

[1] Acosta, M., Gallar, J., Belmonte, C. The Influence of Eye Solutions on Blinking and Ocular Comfort at Rest and During Work at Video Display Terminals. *Experimental Eye Research*. 1999, 68, 6, pp. 663–669. ISSN 0014-4835.

[2] Blehm, C., Vishnu, S., Khattak, A., Mitra, S., Yee, R.W. Computer Vision Syndrome: A Review. Survey of Ophthalmology, 2005, vol. 50, no. 3, pp. 253–262.

Collecting the Implicit and Explicit User Feedback on the Web by Means of Eye Tracking

Peter KIŠ[1], Matej KUCEK[1], Peter KUŠNÍR[1], Viktória LOVASOVÁ[1], Martin SVRČEK[1], Marek ŠULEK[1], Šimon VALÍČEK[2]*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova2, 842 16 Bratislava, Slovakia
tp-1415-16@centrum.sk

Nowadays, one of the most interesting methods for evaluation of the usability of software is eye tracking. The eye tracking is in fact monitoring where the user looks (recording the coordinates of the gaze).Besides user experience, eye tracking is also applicable in areas such as human computer interaction [1], product design, web design, etc. The employment of eye tracking requires two basic tools: an eye tracker and software for recording and processing the data. The existing eye tracking software solutions often have certain limitations; e.g., they lack the possibility to track multiple devices simultaneously or the eye tracking support in the context of dynamic web pages.

Our main objective is to use the eye tracker to support the users when using different applications or websites by identifying the problematic or uninteresting parts of the web pages. To achieve this goal we provide the means of communication between the existing eye tracking system and external (web) applications through the API we created. The main advantage of this API in collaboration with our own JavaScript library is the ability to respond to gaze events in the real time (e.g., by adjusting the content of the application).The API also enables:

- To add users to the system, create projects and eye tracking sessions.
- To collect the eye tracking data and calculate their aggregations (for example dwell time or the number of fixations for the given areas of interest).
- Based on the collected data and information, to perform various adaptations directly in the application, or generate various notifications (see Learning system in Figure 1).

Although the API is versatile and can be included to any system or a web page, we focus on the domain of learning. Thus, in the context of educational systems, we identify which parts of texts are read and which are not or which passages of text are problematic for students and which are not by analysing reading patterns from the collected data. This can lead to better feedback for the teacher and consequently to the improvement of the learning materials.

This concept is also interesting from the perspective of determining the style of reading (who reads carefully and who skips paragraphs) [3] and tracking user behaviour over a longer period of

* Master degree study programme in field: Information Systems[1]/Software Engineering[2]
Supervisor: Róbert Móro, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 445–446.

time. Currently, we know very little about the behaviour of the user and he or she uses the system in the context of longer sessions [2].

We also need to visualize the information to teachers and show them various statistics in an understandable form, because the data from the eye-tracker are quite complex. Therefore, we use this acquired data and process them into various statistics and charts that appear in the web application (see Web app in Figure 1).Such data processing has three basic meanings for us:

- It provides an understandable view of the data for people who are not researchers.
- It allows the researchers to obtain interesting charts that are important for them in the context of their work.
- It can assist teachers in developing textbooks by obtaining information about problematic areas.

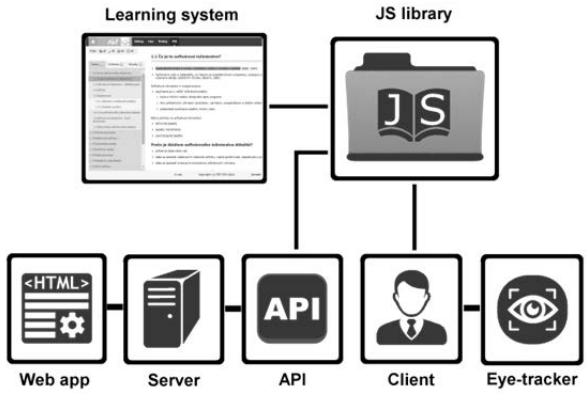


Figure 1. Basic structure of our system.

By solving these problems, we create full-featured tracking software. As mentioned the main benefit that we see is the ability to improve the work with the system, as well as supporting certain activities that are at the heart of the system (e.g., learning).However, it is important to note that our solution is not limited to the field of education. We developed the API and the JavaScript library so that it is portable to any (web) application. This allows the developers to benefit from the explicit and implicit feedback that can be collected by eye tracking in their own application, without the need to make any major adjustments.This would become especially interesting if the eye trackers became a part of the notebooks and other devices in the same way that the web cameras are nowadays, which can be a reality in few years.

References

[1] Jacob, R.J.K., Karn, K.S.: Eye Tracking in Human-Computer Interaction and Usability Research: Ready to deliver the promises. In: Hyona, Radach & Deubel, eds.: *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*. Elsevier Science BV,(2003), pp. 573–605.

[2] Jiang, J., He, D., Allan, J.: Searching, Browsing, and Clicking in a Search Session: Changes in User Behavior by Task and Over time. In: *Proc. of the 37th Int. ACM SIGIR Conf. on Research &Development in Information Retrieval*. ACM Press, (2014), pp. 607–616.

[3] Liu, Y., et al.: From Skimming to Reading: A Two-Stage Examination Model for Web Search. In: *Proc. of the 23rd ACM Int. Conf. on Inf. and Knowledge Management*. ACM Press, (2014), pp. 849–858.

IIT.SRC Junior 2015

High School Students at IIT.SRC Junior 2015

Jakub ŠIMKO and Mária BIELIKOVÁ[*]

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
{maria.bielikova,jakub.simko}@stuba.sk

Abstract. The IIT.SRC Junior track is a platform for talented high school students interested in informatics and information technologies to present their innovative ideas and projects to their senior colleagues – university students and staff. During poster sessions, works accepted to the IIT.SRC Junior track have been presented by their authors, who subsequently received feedback on their projects throughout discussions.

1 Seeking the talent

Seeking for the talented high school students is essential for maintaining quality of future IIT.SRC conference submissions as well as the life of the faculty. Therefore, we repeatedly started the IIT.SRC Junior track – a platform for high school students to present and discuss their innovative ideas and projects in the field of informatics and information technologies. Previous years of IIT.SRC showed to be promising since we managed to involve several talented high school students who recently became our students.

Student works accepted to this track have been presented by their authors during regular poster sessions. Here, the authors had the opportunity to receive valuable feedback from the faculty members as well as from their older colleagues. The authors had also the opportunity to view and discuss other works presented at the conference to gain experience and inspiration for their future projects.

2 IIT.SRC Junior 2015 Projects

This year, four submissions were selected. All of them presented as extended abstracts for more detailed explanation of proposed ideas and realized prototypes. The first project, authored by Michal Tomáš Buday and Jana Tomanová supports teachers during their work on student assignment point assessment. Their smartphone application transforms point gains from tests to gain percentages and marks.

In the second project, Dávid Majerčák and Jakub Kozák devised a video-chat which connects people from all around the world. It helps friends, to spend their time (virtually) together based on specific activities, like learning foreign languages and cultural exchange. The solution offers matching chat partners based on variety of parameters like country, language, sex and type of drink they like.

[*] Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

IIT.SRC 2015, Bratislava, April 23, 2015, pp. 449–450.

The third project was authored by Dávid Majerčák and Peter Majchrák and involved a prototype of an extension cord, which can be controlled over a cellular network or over the Internet using a website with a user-friendly interface. This extension cord also features a timer, power consumption meter, GSM and web switcher for each socket and also temperature and humidity sensor. The solution could be used to help people to make their household smarter and manage their devices wirelessly without any need of rebuilding the house

The fourth project was authored by Martin Pavelka, who created a website about web safety for common users. Author addresses the issue of articles on web safety being too difficult to understand by presenting simplified guidelines. Project focuses on preventing virus infiltrations and digital identity theft.

More information about the IIT.SRC Junior track can be found on the Web:
`http://junior.fiit.stuba.sk/`

Project Assessment Calculator: A Handy Tool for Teachers

Michal Tomáš BUDAY*, Jana TOMANOVÁ*

Saint Joseph Kalazans Grammar School
Piaristická 6, 949 01 Nitra
team.hodnotenie@gmail.com

Abstract. In this paper we will describe you how to improve the assessment of students' tests at our school. The main point is to create an application for smartphones and desktops which facilitate the work of teachers during correcting of tests. Instead of calculators, the application running on their smartphones and computers nearby could be used. Teachers put in just two values and get a mark, a percentage and a point scale.

1 Introduction

Our experience from the school is that when teachers are correcting tests they can easily make mistakes during counting percentages with calculator. And also it can be difficult to correct and count for couple of hours. But we can bring a little help for them. The application requires achieved points and total points of a test to give them everything necessary to rate students and also it gives them a point scale.

The first idea of our solution came last year and we implemented it within a school project as an application for desktops. It was running and after time we decided to create an application for smartphones because in this age lot of people use them more then computers. However, still there are older teachers who use computers or calculators so we would like to support each age group.

2 Our solution

Assessment calculator is a tool for teachers which can cut the time when determining grades for tests. By the first launching, the application shows basic instruction how to use the application. Afterwards, a list of school subjects is offered to a teacher to choose from. There are subjects like Slovak language, English, Math, Biology, etc. The teacher can find almost all of subjects taught at a secondary grammar school with predefined rules. In the next step there is a window with 4 boxes and a submit button. The first box is denoted as "Total points" that means you put in the maximal points students could achieve in the test. The second one "Achieved points" is for number of points which the student achieve in the test. The third one is for the percentage of success and the last one for a mark. The two last boxes teachers do not fill.

---

* IIT.SRC Junior Contribution
Mentor: Peter Krátky, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

Our application has been developed in several iterations. First, we made the basic application which counts just percentage (see Figure1a). This was a simple coloured app for a single school subject. Next we have added the function for mark determination from the percentage scale of the subject (see Figure 1b). Then we tried to add more subjects like languages, natural subjects and social subjects, everyone category or subject with its own percentage scale (see Figure 1c). We have tested our application with a two teachers in our school and so far we have got positive feedback for either functionality as well as usability.

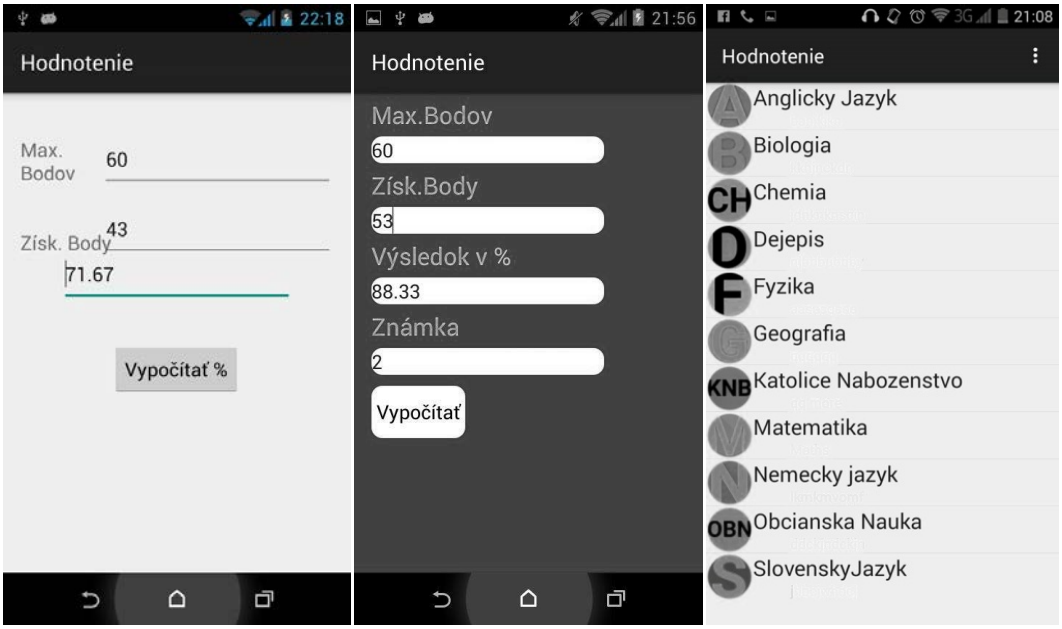


Figure 1. a) The basic application (left screenshot), b) The second version (middle screenshot), c) Subjects (right screenshot).

3 Conclusions

Application is fully functional and we are going to test them with more teachers at our school to get feedback about our work, errors or their insides and ideas on how to enhance it for better and comfortable usage. For the future we would like to add some extra functionality such as point scale from the test, history of assessment and activity to remember the subject which the teacher use mostly. So far our application is designed for the purposes of teaches at our school. It is a plan for the future to make the grade scales easily adjustable for teacher from other schools. We have also contacted more schools in our region to gather data and we will provide predefined profiles for teachers of those schools in our application.

Video-Chat as the Way How to Break the Ice

Dávid MAJERČÁK, Jakub KOZÁK*

Grammar school of J. M. Hurbana in Čadca
17. Novembra, 022 01 Čadca
david1majercak@gmail.com

Abstract. In this paper we will introduce our idea of video-chat which connects people from all around the world and also friends, help them spend their time based on specific activities, like learning foreign languages, getting to know with strangers who can offer a new point of view about a variety of cultures. Last but not least having fun while having a drink and playing a game together.

1 Introduction

The main goal is the connection between various people with the aim to speak foreign languages better, to relax and have fun at the same time. Our video-chat is similar to already existing RVCs, such as Omegle or Chatroulette. We do not just copy their idea to select a person randomly and connect them with another random person, though. We added a new idea and also new features to this kind of video-chat. Users can choose criteria in chat filter. This chat filter consists of possibilities like country, language, sex, type of drink to enjoy.

After a hard day or exhausting work everybody needs time to slow down, have a break and relax with a cup of red wine or a mug of German beer. The video-chat connects business with pleasure because users can enjoy their favorite drink while talking and playing with other users.

We all agree that learning foreign languages nowadays is more important than it used to be in the past. Students learn in school more than only 1 language and the skill to speak 2 or 3 languages has become a normal standard among young people. A lot of students are not able to speak these languages out of school with common people in the real world because the students do not have a chance to practice speaking. Not all students get opportunity to meet native speakers and have a conversation with them.

In our opinion, people who travel for weeks or months receive a better vocabulary and a larger view on a certain language than students who learn a foreign language in school for years. We cannot pay travel expenditures for everybody who wants to improve their language skills but we offer this way of learning. Two random people can use this video-chat to have a drink, play a game and talk.

* IIT.SRC Junior Contribution
Mentor: Lukáš Turský, Institute of Applied Informatics, Faculty of Informatics and Information Technologies STU in Bratislava

2 Proposal

We are currently at the state of development concerning web RTC random video-chat, using JavaScript libraries, such as socket.io, mongoose and express.js.

We are also planning to add some games developed with the use of JavaScript and its graphical libraries, especially three.js and D3.js. The game is based on dice, where each user rolls one and numbers are added and according to the result task to do will be chosen. There are tasks numbered from 2 till 12, which are randomly mixed and chosen at the beginning of the chat session.

There is also possibility of logging in using your facebook account and have you friends directly added to your profile and also contact book or you can just register with account made on website and try to find new friend in close neighborhood or people from other side of the world. It depends just on user which way he or she chooses.

When you meet a person you like to have a good drink and chat with, you can add a person to your contacts and call him or her everytime she or he is online and have a nice evening.

We know that some user are lazy or for their security they do not want to put their personal information to the web. For this reason, there is also choice for logging in using just username that every user can choose at the beginning of the session and with no need to create an account or inputting personal info.

Basic server runs using Node.JS platform and no-SQL database MongoDB. The database is used for users' profiles and also for interconnection between users. The users can choose from a list with whom they speak mostly and they can call them whenever they want. People can also choose a country from which is the person they want to be connected to. A user who is under the age which allows drinking alcohol in their country is aware of content that is only for adults.

3 Conclusion

We try to develop this video-chat to fulfil our idea that people can learn not only at school but also in their free time while having fun with foreigners. We mainly created this for users that are shy, because when a person drink a glass of good wine or mug of beer it is scientifically proven that a person is much more less shy.

This site was also created for users that are lazy to sign to a course of foreign language and they do not like style of learning in class, sometimes things that a man can never use. This type of person would love to try our idea.

We prefer this kind of learning because that is not only about memorizing words and facts by heart. All people can learn more if they are interested in the thing they are learning.

Smart Extension Cord – Make Your Residence Smarter

Dávid MAJERČÁK, Peter MAJCHRÁK*

Grammar school of J. M. Hurban in Čadca
17. Novembra, 022 01 Čadca
david1majercak@gmail.com

Abstract. In this paper we describe a prototype of an extension cord, which can be controlled over a cellular network or over the Internet using a website with a user-friendly interface. This extension cord also features a timer, power consumption meter, GSM and web switcher for each socket and also temperature and humidity sensor.

1 Introduction

Realize that you are far away from your home and you forgot to turn off some of your devices and it is really dangerous to leave it turned on or you are the type of person that often works from home and needs to turn many appliances on and off. Then our solution is exactly for you. We are working on a prototype of a Smart Extension Cord, which will help people to make their household smarter and manage their devices wirelessly without any need of rebuilding the house. Our solution is to make an extension cord, which is controlled over GSM, by calls or SMS, or over Wi-Fi by smartphone or computer through user-friendly web interface with many functions. These functions can be changed due to individual needs. Every user will have its own profile where he or she can edit functions which will be listed in main page for fast access.

2 Proposal

Our prototype will solve all problems mentioned in first paragraph in this document. Our prototype is able to switch on or off each plug separately over the web interface or by sending SMS or by a phone call. It is based on how many times the device ringed and according to that, the correct plug will be toggled. For example you will call and it will ring twice so second plug will be toggled. You can also check the humidity and temperature in the room where the extension cord is located or electric current and overall power consumption of electricity by using the web interface. There is also approximate price of used electricity during current year or month according to user-provided price of electricity in their region.

Now about the hardware. We have chosen to use STM32F103 microcontroller with the ARM Cortex-M3 architecture. It suits the purpose very well. For the Wi-Fi we are currently using ESP8266 module that handles all the protocols internally, so that we can offload the

* IIT.SRC Junior Contribution
Mentor: Maroš Ďuríček, Institute of Computer Systems and Networks, Faculty of Informatics and Information Technologies STU in Bratislava

microcontroller. GSM connectivity is done via a SIM900 module that is simple to interface to and does the job nicely. For hardware based user interface we use 128x64 pixel white OLED display with SSD1306 controller communicating with the main microcontroller over an I2C bus. The user interaction will be done with a rotary encoder, which makes everything very simple and it's also easy to interface to the microcontroller because of its hardware rotary encoder support.

There are multiple sensors being used. For measuring current we use ACS712 Hall effect based current measuring IC that outputs an analog signal that gets sampled by the microcontroller's built in 10bit ADC. We then use basic digital signal processing algorithms to get information about the current. The voltage sensing part is done very similarly but we first drop the voltage by using a resistor divider. We can the further process the samples to get other useful information about the power like power factor, real power, reactive power, etc. For sensing the humidity and temperature we use AM2301 sensor also known as DHT22. It communicates over a proprietary OneWire bus.

We also use Bluetooth to detect the presence of people. This feature can be disabled at any time but we think that it is very handy. It provides a way to save power by switching specific devices off when your Bluetooth enabled device, usually a smartphone, gets out of range and also switching then back on when you come back. This feature works nicely with room lighting.

On the server side the Node.js is used for receiving of data and also later for displaying the statistics of power usage and also statistics of humidity and temperature. Node.js is completely new platform for rapid building of web applications based on Google Chrome's JavaScript Runtime mainly built for real-time applications. It also features profile of each user, where there can be added more devices.

We also use JavaScript libraries for managing different tasks. For drawing graphs we use D3.js and vis.js. For MVC framework we use Angular.js. We decided to use MongoDB for our database because of its simplicity.

3 Conclusion

The project is still under heavy development. We managed to get working our microcontroller based board that is now able to receive calls and toggle the appliances. We have also managed to send data from it to the database through Wi-Fi.

We are looking forward to see our prototype be able to do all of the functionality mentioned above and to fully cooperate with the server.

Don’t Get Caught in the Network!

Martin PAVELKA*

Evanjelické lýceum
Vranovská 2, 851 02 Bratislava, Slovakia
martin.pavelka.ba@gmail.com

Abstract. In this paper we describe our idea of website about internet safety. Many free guidelines and manuals are available for users to download from the Internet. Often are these articles too difficult to understand for users and they cannot get advantage of the provided advice. Non-considering right user´s behaviour in the network may cause virus infiltration to the device, revealing of digital identity or stealing user´s files and accessing user’s financial resources. Our goal is creating a website about internet safety with simplified guidelines. These user-friendly articles will improve the user’s security anonymity and lessen vulnerability of his devices or his digital identity.

1 Introduction

Internet security is frequently mentioned topic in media nowadays. Web 2.0 provides users many services such as internet banking or remote access to other workstations. Attacks targeting end-users and their devices are more popular, because they are more vulnerable than highly-secured servers or datacentre infrastructures. Many users still do not know how to secure their computers and other electronic devices and they do not have the basic skills about internet behaviour. Our webpage about internet safety provides simple guidelines and articles that improve the user’s anonymity and security on the Internet.

2 Related work

The topic about internet safety became a new part of computer sciences and its basis are taught pupils since elementary schools. Animated videos were formed to convince young children to become acquainted with internet safety – project ovce.sk [1]. These videos, in our opinion, give children only the experience about virtual identity. The project does not give practical computer advices to the children. Not only the children, but also adults often do not have the right information and knowledge about internet safety and other restrictions. During our servicing user’s computers we had the opportunity to work on computers with running two outdated antivirus systems, internet browser with pre-filled and pre-logged internet services such as Google and Facebook. Users may click on ad looking like menu, because they have many ad plugins installed. These are only a few things that may cause the user bad consequences and we need to prevent them by improving the “Internet literacy” of users.

---

* IIT.SRC Junior Contribution
Mentor: Ján Laštinec, Institute of Applied Informatics, Faculty of Informatics and Information Technologies STU in Bratislava

3 Organization of website

Our website consists of homepage and subpages. On the homepage is basic paragraph describing our project. Subpages INTERNET, COMPUTER, LOCAL NETWORK include the tips and tricks. Contact page provides the reader to email us via contact form.

3.1 Homepage

Homepage is the basic page providing useful information of usage other information posted in subpages, information about our project and main goals and also copyright information. On the homepage is also placed link for live download of the content of our webpage as a pdf file.

3.2 Topic

Each subpage has in the upper part a list of the topic. Reader can easily access the selected topic via hyperlink. Each topic has its name, motivating question and explanation with tips and tricks. Organization of topic is described in the Table 1 below.

Table 1. Organization of a topic on the subpage.

Title of the topic
Why to be interested in this topic? This section motives the reader for further reading the post. Some basic information about the risks and consequences are provided.
How to secure? How to prevent this situation? This section provides guidelines, explanations and more information about the topic. The reader will be able to understand the risks and will be able to prevent himself from getting to the situation described above.

3.3 Contact

On the contact page the contact form is provided. In the future we plan to set up a forum for discussion and or social network plugin.

4 Conclusions

In this article we presented our solution of improving internet security skills and manners of users that will help to lessen the possible attacks to their computers or losing their digital identity. We have created a website called “Don´t get caught in the network”, which provides readers many tips and tricks how to secure their devices and how to properly use some internet services[1]. Being able to react calmly in the network is the basic skill needed for secure browsing the internet and using Web 2.0 services. In the future many things will be renewed and updated, but if the user will once learn how to secure himself, he would not have problem to adapt in a new or renewed environment.

References

[1] Ovce.sk – Sheeplive.eu. [Online; accessed March 13, 2015]. Available at: `http://sheeplive.eu`

[2] Bednár, J.: *V sieti.* ENIGMA, Nitra, (2002).

[1] The website is available online at: `www.ulovenyvsieti.tk.`

Other Accompanying Events

Programming Contest at IIT.SRC 2015

Márius ŠAJGALÍK, Jozef TVAROŽEK, Peter TREBATICKÝ, Mária BIELIKOVÁ*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
{name.surname}@stuba.sk

Abstract. Programming contests have a long tradition at the Slovak University of Technology in Bratislava. The annual student research conference offers an open day without any lectures for all of our students, and we are looking for ways how to attract them to take part in the event. For the eighth year, we have refreshed the accompanying programming contest that is usually run in a traditional ACM-style way with novel types of problems and novel competition rules.

1 Background of the Contest

Programming contests have a long tradition at our university and the faculty. From the beginning in 1998 local contests were organized for our students in order to form teams to represent Slovak University of Technology in Bratislava at ACM International Collegiate Programming Contest (ICPC) Regional Contest in Central Europe region. Since 2002 our faculty has participated in organization of Czech Technical University Open, which is joint event where universities from Czech and Slovak Republic compete with the aim to select their best students for representation in ACM ICPC Regional Contest.

We prepare our students for this type of programming contest already before they enter their first year in university. We organize the ProFIIT programming contest for our prospective students since 2004. It consists of two rounds. In the correspondence round the contestants compete in solving several (around 10) programming problems. They are allowed to compete either on their own or in pairs. The best teams advance into onsite round organized at our faculty. They compete on their own in this round as they can gain bonus points into the admission process.

Previous three years, ProFIIT finals accompanied the IIT.SRC, where finalists could see potential future students exciting research opportunities awaiting them at our faculty. The evaluation system used in ProFIIT was also used in all past programming contests at IIT.SRC. This year, we moved to using our educational system Peoplia, which is used in several undergraduate courses to facilitate learning programming. It is used in the very first academic year in the course *Procedural Programming* and in the second year in courses *Data Structures and Algorithms* and *Analysis and Complexity of Algorithms*. Students can also choose an elective course *Development of Effective Algorithms and Programs* which further develops the algorithmic thinking in them and teaches them the more advanced techniques specifically usable in programming contests. Moreover, our undergraduate students selected for the research track have

* Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

more possibilities in algorithms training, mainly in research seminar, where they learn about some even more advanced algorithms.

2 Structure of Programming Contest at IIT.SRC

This year, we introduce new problem types into the programming contest. Small part of the problems retained the structure of the original contest, which resembles the structure of a traditional ACM-style context. Additionally, for the first time this year, we introduce database querying and error-checking problem types. The goal of this innovation is to motivate students to take part even if they are not very proficient in solving traditional algorithmic problems.

Contestants compete on their own onsite in our computer labs. They have 90 minutes to work on the problems. Each problem consists of: basic description of what should be solved, exact input format specification for the program (or schemas of database tables used) as well as the output format, along with sample inputs and the corresponding sample output.

For the traditional problem types, the task is to create a program in C++ that transforms test input, conforming to the described format but is unknown to contestant, into correct output according to the problem statement. This year however, new types of problems require new types of solution. Other problem types involve: filling in a missing part of a program, writing a database query in SQL, providing an input which exploits the error made in given program, or pointing out a mistake made in specification of a sample software system.

Contestants submit their solution through our educational system Peoplia, which compiles the code (if applicable), runs it against test input, evaluates the obtained output and informs the contestant of the result. The result is presented only as a simple statement, e.g. “Accepted”, “Wrong answer” or “Time limit exceeded”, which means the run time of submitted program exceeded the allotted time limit. Only the result “Accepted” means that a contestant successfully solved a problem.

The scoring system has been also innovated. Each problem is worth a certain number of points. After a successful solution, the contestant gains the points. The final ranking of contestants is primarily determined by the number of points gained and in case of a tie, by the sum of the times taken to solve each problem since the beginning of the contest. There is also a 10 minutes penalty for each submitted incorrect solution, but only for the eventually solved problems. This type of rank determination favours primarily contestants who solve more and tougher problems, but secondarily those who first solve easier problems and also those with lower number of incorrect submissions. The ability to quickly create correct solution is also very important apart from the ability to come up with a working idea. These skills are mainly trained by practice and learning that is where we help the students through activities mentioned here.

The contest is made more exciting for participants by the fact that during the last 30 minutes the preliminary results are not updated. This way, contestants cannot be sure about their final standing until the awards ceremony. The time interval of not displaying preliminary results was chosen in accordance with conference schedule, because there is another contest ending right before the second poster presentations in which the other conference attendants can tip the winner.

More information about our programming contests can be found on the Web:

- ACM programming contest – `http://www.fiit.stuba.sk/acm/`
- ProFIIT programming contest – `http://profiit.fiit.stuba.sk/`
- Peoplia educational system – `https://www.peoplia.org/fiit/`

FIITAPIXEL Exhibition at IIT.SRC 2015

Pavol NÁVRAT, Mária BIELIKOVÁ, Ján LANG*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
{name.surname}@stuba.sk

Abstract. FIITAPIXEL is an initiative of the Faculty of Informatics and Information Technologies that brings together its members (both students and staff) as well as its potential students and alumni in an effort to create, share and judge pictures. It is organized as an ongoing event, where anyone can contribute pictures to certain categories of photographs. The submitted photographs take part in a contest that is organized annually. Besides best photographs, also best photographers are announced based on their success with their photos. The contest has an expert panel of jurors who give their lists of best photos in each category. In parallel, visitors vote for any photo they like and their votes are counted to result in list of best photos according to popular voting. For the fifth time we organized at the IIT.SRC an exhibition of the best pictures this year contest.

1 FIITAPIXEL as an inspiration

FIITAPIXEL is an initiative of the Faculty of Informatics and Information Technologies to contribute in providing to its members, students and staff alike, an inspiring, creative, stimulating environment to study or to work in. Studying is mostly demanding and hard, and so is working at an institution which faces such a level of competition as is the case in the higher education sector in informatics and information technologies related fields in this region of Europe. From Budapest to Prague, from Vienna to Brno, in a relatively close proximity of Bratislava there several respected institutions with a similar scope of interest. Moreover, in the city itself, there are several other competing institutions.

We try to offer something that may make a little difference. By providing a platform and other forms of support, the Faculty creates an environment that allows expressing its members in a completely different way as it is usual in their professional work. Instead of writing programs or designing chips, they get a chance to express themselves by way of pictures. The language of pictures is intended as a language of artistic expression, even when respecting all the limitations given by the simple fact that these professionals in one (informatics related) field are complete amateurs in another (photography) and similar limitations apply when e.g. elements of journalism are involved.

* Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

2 FIITAPIXEL Organization

FIITAPIXEL started in 2009 and it has been organized ever since then. It takes place as a contest organized annually. The final results are usually announced and prizes awarded around the time of our student research conference IIT.SRC. Immediately after one year of the contest is closed, themes for the next one are published and the contest is open again. The contest is organized in two legs during one year that last approximately half a year each.

There are usually four themes open for each particular leg, but some of them may be adopted for the next period. For example, in academic year 2014/15 contest there were these four themes for the first leg (Summer and Autumn):

- *Colourful nature*
- *The place where I am right now*
- *City Lights*
- *Enchantment of tininess*

with *Lines, shapes, patterns* replacing the third one for the second leg (Winter and Spring).

Each participant can submit up to five pictures to each category both in the first and the second legs. These up to 40 pictures are published on the contest portal, where they are freely visible from anywhere in the world. Anyone can express her/his likes which are treated as votes for the particular picture. At the end of each period, votes are simply counted and the best dozen pictures are announced as winners, according to a popular vote, in each category.

There is also an expert jury formed by experts in visual arts which gives its opinion resulting in another set of lists of dozen winning photos in each category. Results of both opinions, expert and popular, are then used to determine a list of best photographers based on how their photos are placed in particular results.

In the 2014/2015 contest, we have had 807 pictures taken by 111 authors. They received nearly 1 400 votes from visitors. Pictures and wining photos are available on the contest portal: `http://foto.fiit.stuba.sk`.

3 IIT.SRC Exhibition

Annual evaluation of the best photographers of the FIITAPIXEL Contest takes place at the student research conference award ceremony. Moreover, we give conference participants the opportunity to enjoy an exhibition of the winning photos of each category in both legs, i.e. we exhibit two dozens of winning pictures, in 2015 in nice new building of the Faculty. IIT.SRC participants can cast their vote for the best photo during the conference. At the end of the day, winner of the participants'vote is announced and awarded.

FIITAPIXEL brings new dimension into our living space at the Faculty together with much inspiration for our activities. The selected best photos will decorate our environment in new building.

RoboCup Presentation at IIT.SRC 2015

Ivan KAPUSTÍK, Pavol NÁVRAT *

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 3, 842 16 Bratislava, Slovakia
{ivan.kapustik, pavol.navrat}@stuba.sk

Abstract. RoboCup is an attractive project theme with a free participation, designed to support education and research in artificial intelligence, robotics and information technologies. During the last few years, our students achieved some interesting results, which were presented during our student research conference.

1 Motivation

RoboCup is an international joint project to promote research in artificial intelligence, robotics and information technologies. It is an attempt to advance artificial intelligence and intelligent robotics study and research by providing a well-known and attractive problem where wide range of technologies can be integrated and examined. RoboCup chose to use soccer game as a central topic of research. The ultimate goal of the RoboCup project is to develop by 2050 a team of fully autonomous humanoid robots that can win against the current human world champion team in soccer.

In order for a robot team to actually perform a soccer game, various technologies must be incorporated, including design principles of autonomous agents, multi-agent collaboration, strategy acquisition, real-time reasoning, multi-level decision making, robotics and sensor-fusion. RoboCup is a task for a team of multiple fast-moving and skilled robots within a dynamic environment. It offers also a software platform for research on the software aspects. RoboCup is traditionally divided into four main fields: RoboCup Soccer – defined by the original domain of soccer, RoboCup Rescue – intended to do search and rescue in large scale disaster area, RoboCup Junior – aimed to child education and motivation and RoboCup @Home – oriented to provide various help not only at home. Continuous growth of interest in other robotic domains is shown by inventing new leagues. The RoboCup Logistics League is a sponsored league, focused on in-factory logistics applications.

From our point of view, the main goal of RoboCup is to promote research in areas of artificial intelligence and information technologies, especially in the area of multi-agent systems. This is a benefit for the students, making their studies more interesting and attractive. Students can meet with robotic soccer in courses like Artificial Intelligence, Team Project and others. Students are facing an interesting problem, which demands invention as well as use of modern artificial intelligence approaches. Teams of students have the possibility to directly compare their results in tournaments. This encourages the students to even higher effort and

* Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

motivates them for better results. More fundamentally, achieving progress requires tackling serious open research problems in artificial intelligence, such as planning of cooperation of multiple agents etc. That is why this area is of interest also for our doctorate students.

We have been organizing this tournament regularly for several years. Starting as a local event in 2000, it has grown to a regional contest under the official RoboCup authorization. Our Faculty organizes tournaments in the simulated category only, but we gradually include other categories. Our current contest event has three parts.

First part is a tournament of two-dimensional (2D) simulated player teams, where students try to make their own players win soccer game. 2D players are simple entities, ready to follow any possible action in their virtual environment. Students' main research is aimed to team tactics and autonomous player decision. It includes team formations and planning, player communication, use of a team coach and decision skill improvement. Methods here cover planning and player's action selection based on diverse sources – success evaluation of similar situation, teammate decision model and prediction of opponent behaviour. This contest part is currently more an exhibition of new approaches than a tournament, because interest of our students shifted to more complex three-dimensional (3D) robotic simulation.

Second and third parts of this tournament involve three-dimensional (3D) robotic simulation. These robots are true copies of their real master. They have limbs and joints. Primary students' task was to teach robots to reliably walk, turn, stand up and kick the ball. It was followed by design of a proper composition of these basic skills to achieve simple goals, like walking to the best game position or getting the ball. Then, the training support framework has been developed and test modules for robot learning were created. This academic year, students' effort is oriented to code refactoring and optimisation. Previous work of multiple student teams on one robot code made this code inefficient, entangled and too redundant. The code was difficult to read, change and maintain. This student team replaced more than 4000 lines of improper code with more than 1000 lines of new code and comments. New code is equipped with improved logging, testing and documentation. Robot can use stabilisation for any movement and its fall detection is enhanced. Training framework provides more types of training situation for improving new skills.

Any soccer player must be good with physical skills and must make good and fast decisions during the game. So the second part of our tournament contains skills match. Robots compete in speed and accuracy of given tasks. They can get a few points for "unusual" useful skills as well. Finally, third and most valued part of this tournament holds soccer contest, where both skills and decision making are verified in real-time game.

2 Results presentation

For this student conference we decided to hold an exhibition of new results achieved in 3D soccer simulation. Presentations were enhanced by show of robot skills performance. Our students improved some of old movement sequences and added few new ones. New skills included mainly fast and reliable movement to a chosen place and faster robot orientation. Improved player actions were also more attractive for audience.

The extension of the soccer game simulation to the third dimension shows the continuous progress in RoboCup and in our students' skills, too. Decision making of these robots is very complex and brings new challenge to everyone concerned. We hope that exhibition of robotic simulation will attract many present and future students and give them motivation for their study and research work.

More information about our annual tournament can be found on the web page `http://www.fiit.stuba.sk/robocup/`.

International Girls in ICT Day 2015

Petra KOTULIAKOVÁ

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
kotuliakova@fiit.stuba.sk

Abstract. The project Girl's Day, organized by Aj Ty v IT (You Too in IT in English), NGO, supports girls in choosing careers in informatics and technology. Girl's Day is an integrative action day when girls visit companies, informatics faculties, universities, with the focus on hands-on experiences. Opening session with an interesting program is held at the Faculty of Informatics and Information Technologies Slovak University of Technology in Bratislava, who is a key partner of AJ Ty v IT.

An international Girls in ICT Day was established in 2011 by International Telecommunication Union and is held every year on the fourth Thursday in April. For this year, the session is on 23rd of April. The Day celebrates girls' interests and strengths, and encourages them to choose a career in information and communication technologies.

Although women and men have the same choices of careers, there is still a mindset that distinguishes between typically male and female occupations. As a result, girls are under-represented in engineering and computer science. "Girls in ICT" Day encourages girls to let their career path be defined by personal interest and talent, not by stereotypes. It promotes an interest in technology, computer science and new communication media, offering girls opportunities to make contacts for their future career and to be inspired by women in leadership positions.

Girl's Day is an integrative action day when girls visit companies, informatics faculties, universities, with the focus on hands-on experiences.

The patronage of President of Slovakia Andrej Kiska under this year of Girl's Day is very important for reaching girls, their parents and teachers. He was personally present at the opening session and participated at the girl's workshop for a short time.

Last year, during the first Slovak session, we reached 500 girls and proposed 35 different programs to them. Feedback from girls and their teachers is very positive, for almost of them it is a unique possibility to visit these companies and being in contact with IT sector. Through this way, we can really influence their opinion and motivate them for choosing IT career.

After success in 2014 we opened more than 700 places on 62 different programs this year. One of program and opening session is directly at the Faculty of Informatics and Information Technologies Slovak University of Technology in Bratislava, who is a key partner of Aj Ty v IT initiative.

Girl's Day is a big opportunity for us to present information technology to young girls, their parents and teachers. Activities of AJ Ty v IT, NGO, brought after 3 years the rise of girl's students at the FIIT STU from 3 % to 10 % of all students.

More information about Girl's day can be found on the web pages www.girlsday.sk and www.ajtyvit.sk.

Author Index

Mária Bieliková (Ed.)

IIT.SRC 2015: Student Research Conference in
Informatics and Information Technologies

1st Edition, Published by
Slovak University of Technology in Bratislava

502 pages, Available at http://iit-src.stuba.sk/2015/
Print Nakladateľstvo STU Bratislava
2015

ISBN 978-80-227-4342-6

ISBN 978-80-227-4342-6