

**PROCEEDINGS IN
INFORMATICS AND INFORMATION TECHNOLOGIES**

STUDENT RESEARCH CONFERENCE 2017

MICHAL KOMPAN AND PAVOL NÁVRAT (EDS.)

KEYNOTE BY CESARE PAUTASSO

2017 | iit
src

Proceedings in
Informatics and Information Technologies

IIT.SRC 2017
Student Research Conference

Michal Kompan and Pavol Návrát
(Eds.)

IIT.SRC 2017: Student Research Conference

13th Student Research Conference
in Informatics and Information Technologies
Bratislava, April 27, 2017
Proceedings



IEEE

Czechoslovakia Section



Slovakia Chapter



**SLOVAK UNIVERSITY OF
TECHNOLOGY IN BRATISLAVA
FACULTY OF INFORMATICS
AND INFORMATION TECHNOLOGIES**

Proceedings in
Informatics and Information Technologies

IIT.SRC 2017
Student Research Conference

Editors

Michal Kompan, Pavol Návrat
Faculty of Informatics and Information Technologies
Slovak University of Technology in Bratislava
Ilkovičova 2
842 16 Bratislava, Slovakia

© 2017 Contributions are printed as delivered by authors
without substantial modifications

Papers included in proceedings upon the authors request.

Visit IIT.SRC on the Web: <http://iit-src.stuba.sk/2017/>

Executive Editors: Katarína Mršková
Copy Editor: Peter Gašpar
FIIT STU

Cover Designer: Peter Kaminský

FIIT STU
Ilkovičova 2, Bratislava, Slovakia

Published by Slovak University of Technology in Bratislava in SPEKTRUM STU
2017

ISBN 978-80-227-4727-1

Preface

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 an active participation of students in research. This holds not only for students of the 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 the master’s and even the bachelor study.

Universities of technology have a long tradition of students participating in a research work where they have to apply their theoretical knowledge. Outcomes of such endeavours have usually been presented at various students’ competitions or exhibitions. Our university, too has a long tradition in organizing such events. However, in 2005 we decided to upgrade the framing of the event by transforming into a Student Research Conference covering topics of Informatics and Information Technologies (IIT.SRC). Participants are students of all three levels of studies, i.e. they recruit from both the undergraduate and postgraduate studies. The conference has adopted a process of reviewing as is usual at any other scientific conference. The accepted papers are available to all participants in a form of printed preprints (Proceedings). The aim of reviewing is not so much in achieving a certain acceptance rate, which is currently rather common in our community, but in assisting the student authors to amend their papers in the process of revision by providing helpful reviews. Of course, filtering out papers with a too low quality happens, too.

This is only part of our effort to support students in their (often first) steps in communicating their results to the research community at large. We also encourage students to improve their papers and attempt to publish them in international journals or conferences to make them available to their elder peers. Especially, we support the youngest generation of researchers (bachelor and master’s students) by special travel grants that cover partially their travel expenses to conferences.

IIT.SRC 2017 attracted 75 student papers from which 66 were accepted as research papers (17 bachelor, 33 master, 16 doctoral) and 9 as papers to the innovative application and technologies track (7 bachelor, 2 master). The number of papers slightly varies each year. This year we have noticed a decrease in bachelor and master categories comparing to IIT.SRC 2016.

The research track of the IIT.SRC 2017 conference was organized in six sections presented in course of live discussion in two poster sessions:

- Intelligent Information Processing (13 papers),
- Computer Science and Artificial Intelligence (13 papers),
- Web Science and Engineering (12 papers),
- Software Engineering (6 papers),
- Computer Networks, Computer Systems and Security (15 papers),
- Computer Graphics, Multimedia and Computer Vision (7 papers).

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

The conference was opened by Professor Cesare Pautasso’s keynote titled Liquid Web Apps. Cesare Pautasso is an associate professor at the Faculty of Informatics at the University of Lugano, Switzerland. His research interests focus at building experimental systems to explore the intersection of model-driven software composition techniques, business process modelling languages, and autonomic/Cloud computing.

IIT.SRC 2017: Student Research Conference

Besides the 76 papers presented at the conference and included in these Proceedings several accompanying events were organized. This year we organized for the ninth 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 semesters long team project module. The competition has three stages. 13 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.

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 fifth time, we organized IIT.SRC Junior. It provides a room for presenting inventive high school student projects within the topics of the conference. Five 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.

This year for the third 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.

IIT.SRC 2017 was for the fifth time organized in the new FIIT building. We all benefited from well-disposed space, which supports lively discussions. IIT.SRC 2017 is the result of considerable effort by a number of people. It is our pleasure to express our thanks to:

- members of the IIT.SRC 2017 Programme Committee who devoted their effort to reviewing papers and selecting awards,
- members of the IIT.SRC 2017 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 and Peter Gašpar 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 2017

Michal Kompan and Pavol Návrat, jointly with Mária Bieliková

Conference Organisation



The 13th Student Research Conference in Informatics and Information Technologies (IIT.SRC), held on April 27, 2017 in Bratislava, was organised by the Slovak University of Technology (and, in particular, its Faculty of Informatics and Information Technologies) in Bratislava.

General Chair

Mária Bieliková (dean, Faculty of Informatics and Information Technologies, Slovak University of Technology in Bratislava)

Programme Chairs

Michal Kompan, Pavol Návrat

Programme Committee

| | | |
|---------------------|-----------------|------------------|
| Michal Barla | Eduard Kuric | Viera Rozinajová |
| Vanda Benešová | Peter Lacko | Karol Rástočný |
| Peter Drahoš | Ján Lang | Ivan Srba |
| Anna Bou Ezzeddine | Ján Laštinec | Petr Šaloun |
| Ladislav Hudec | Marián Lekavý | Jakub Ševcech |
| Daniela Chudá | Mária Lucká | Jakub Šimko |
| Katarína Jelemenská | Dominik Macko | Marián Šimko |
| Peter Kapec | Alena Martonová | Peter Trúchly |
| Ivan Kotuliak | Peter Pišteľ | Jozef Tvarožek |
| Tibor Krajčovič | Ivan Polášek | Valentino Vranič |
| Tomáš Kramár | | |

Organising Committee

| | |
|---------------------------|-----------------------|
| Zuzana Marušincová, Chair | Ľubica Palatinusová |
| Zuzana Horniaková | Viera Rozinajová |
| Zuzana Macková | Branislav Steinmüller |
| Katarína Mršková | Roman Stovíček |
| Michal Kompan | |

all from FIIT STU in Bratislava, Slovakia

Accompanying Events Chairs

- *IIT.SRC Junior*: Jakub Šimko
- *TP CUP*: Mária Bieliková
- *FIITapixel Exhibition*: Ján Lang
- *Girl's Day*: Petra Kotuliaková

Sponsors

- Exponea
- Magix
- Microsoft Slovakia
- Soimco
- Vnet

Supporting Professional Societies and Foundations

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

Application for Evaluation of Common Scanpath Algorithms

Michal MELÚCH*

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

Abstract. Recent growth of interest in studying eye movements has sparked the development of new methods for common scanpath retrieval. These can be then used for modelling the behaviour of a much larger user base. Each of the analysis methods was developed in order to solve a specific problem and their measures differ with respect to their target characteristic. Our demo paper is focused on development of a web application suitable for comparing methods of a common (trending) scanpath retrieval. Primary benefit of our work consists of providing a unique way of evaluating and visualizing scanpath data in a web browser.

1 Scanpath analysis

Eye movements and visual stimuli are closely related and together, they represent an essential part of human-computer interaction. Eye tracking and gaze visualization are often discussed as means of measuring and possibly improving the level of user experience. There are numerous approaches to evaluating data collected by eye tracking devices. In this paper, we demonstrate our application based on an approach originating from the Scanpath theory [5].

This approach examines durations and targets of individual gaze fixations of the observed user along with saccades - fast transitions between the fixations. Using this information, we can model user's overall visual attention. In order to perform exact measurements, the data is usually formatted into a sequence containing information about fixations and their durations called a *scanpath*.

Since both order and the spatial information about fixations are desired, scanpaths in our application are formatted as sequences of letters describing the target area of the current fixation (also called an area of interest – AOI) and are associated with numbers which indicate their duration:

A[150], B[200], C[300], A[100]

In the example above, we can identify a single scanpath consisting of four subsequent fixations targeted on three different areas (regions) of interest labelled by letters A-C in a total duration of 750ms. Such formatting of eye movement sequences allows us to apply various scanpath analysis algorithms and to directly compare or visualize their results.

The application demo proposed in this paper is focused on the problem of common (or trending) scanpath identification which is a process of combining several scanpaths into one. Such scanpath can be then used for detecting user experience issues or transcoding the visual stimuli (namely, websites) to make them more accessible [2].

2 Common scanpath identification

There are several approaches to the problem of the common scanpath identification, all of which propose a slightly different solution for solving this task. Our work is based on the current state-of-art algorithms for finding common scanpaths such as Scanpath Trend Analysis [2], Dotplots [4], eMINE [1] and String-edit algorithms [3]. However, the expected outcome is generally the same regardless of the algorithm used – a single scanpath representing a path the most commonly followed by multiple users [2]. Even if the methods of analysis differ significantly, there are some shared qualitative features which are reviewed in following sections.

2.1 Considering fixation durations

Some of the techniques (for instance, the Dotplots [4] and eMINE [1] algorithm) tend to ignore fixation durations while analysing scanpaths. However, it is widely accepted that fixation duration is associated with the depth of processing and the ease or difficulty of information processing [3].

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



Figure 1. An example of the dataset hierarchy where a user-created dataset (*template_sta*) has currently assigned one eye tracking task containing scanpath data (*my_task*).

2.2 Considering visual elements positions

Existing techniques usually do not consider the positions of visual elements defined on visual stimuli (e.g., webpages). However, eye movement lengths are shorter between close visual elements in comparison with the visual elements which are distant from each other [3].

2.3 Reductionism

There is a number of techniques using an approach known as reductionist. We refer to the reductionism as an oversimplification of multiple scanpaths with the loss of some important information – e.g., intolerance of small deviations within individual scanpaths [3].

3 Evaluation of common scanpaths

Our application intends to provide real-time comparison of selected techniques for identifying common scanpaths through a web interface. The individual scanpath data and their relationships are stored in a database and organized in a way that reflects their relative hierarchy (see Figure 1).

The user can choose an algorithm for a common scanpath computation and see the results immediately when the computation is finished. The list of supported algorithms now contains Scanpath Trend Analysis [2]; we plan to add String-edit methods [3], and at least one other advanced approach – currently the Dotplots [4] and eMINE [1] algorithms are under consideration. The algorithms will also have a brief description containing their key characteristics and/or weaknesses available.

Total time of computation and the final result might vary according to the selected algorithm. However, the user can see all the individual scanpaths together with the computed similarity metrics during the computation and he or she will also be able to filter or sort them according to his or her needs (see Figure 2).

| ID | Eye fixations (ms) | Scanpath length | Most similar to |
|----|---|-----------------|-----------------|
| 01 | A ¹⁵⁰ - B ²⁵⁰ - D ⁸¹⁷ - C ²⁵⁰ - B ²⁵⁰ - C ³⁵⁰ - | 6 | 04 |
| 02 | A ¹⁵⁰ - B ⁷⁵⁰ - C ³⁰⁰ - | 3 | 03 |

Figure 2. An overview of the individual scanpaths.

Concept of our application is innovative in the way it automatically handles various aspects of common scanpath data evaluation. The user needs only to provide data previously exported from an eye tracking software along with the defined areas of interest. System then processes the data and lets user choose preferred common scanpath calculation method. The user will be able to immediately validate the results based on the scanpaths visualisation.

4 Conclusions

Our application serves as a proof-of-concept web platform for common scanpath evaluation and it offers many possibilities of further improvement. Currently, the application supports calculation of a common scanpath using STA algorithm and provides basic similarity calculations between the scanpaths.

In our future work, we plan to add other common scanpath calculation algorithms, improve their logic and cluster scanpaths based on their characteristics. We also intend to test our application on the real eye tracking data from the studies carried out in faculty User eXperience and Interaction Research Centre.

Acknowledgement: This contribution was created with kind support of ČSOB Foundation and is partial result of the project University Science Park of STU Bratislava, ITMS 26240220084, co-funded by the ERDF.

References

- [1] Eraslan, S., Yesilada, Y., Harper, S.: Identifying patterns in eyetracking scanpaths in terms of visual elements of web pages. In: *Proc. of 14th Int. Conf. Web Eng. - ICWE 2014*, Springer International Publishing, (2014), pp. 163–180.
- [2] Eraslan, S., Yesilada, Y., Harper, S.: Eye Tracking Scanpath Analysis on Web Pages : How Many Users? In: *Proc. of 9th Bienn. ACM Symp. Eye Track. Res. Appl. - ETRA '16*. (2016), pp. 103–110.
- [3] Eraslan, S., Yesilada, Y., Harper, S.: Eye Tracking Scanpath Analysis Techniques on Web Pages: A Survey , Evaluation and Comparison. *J. Eye Mov. Res.*, (2016), vol. 9, pp. 1–19.
- [4] Goldberg, J.H., Helfman, J.I.: Scanpath clustering and aggregation. In: *Proc. of 2010 Symp. Eye-Tracking Res. Appl. - ETRA '10*. (2010), pp. 227–234.
- [5] D. Noton, L. Stark: Scanpaths in saccadic eye movements while viewing and recognizing patterns. *Vision Res.*, (1971), vol. 11, 929–942.

A New Scheduling Algorithm for Real-time Networks

Veronika BÚCSIOVÁ*

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

Abstract. Time-sensitive networks (TSN) are deterministic networks designed for distributed real-time systems and data transfer based on periodic hard real-time tasks. In this paper, the importance of deterministic networks in data communication is explained and a new static scheduling algorithm for periodic communication in TSN is proposed. Existing solutions are focused on minimizing jitter, thus increasing the determinism of real-time communication. However, our algorithm not only minimizes the jitter, but it also optimizes the latency of all communication pairs. Depending on the network topology, physical parameters of the network and frame requirements within the specified communication pairs, the proposed algorithm identifies and solves potential conflicts in the network. When the scheduling is completed, the algorithm generates a schedule for periodic real-time communication between all network nodes, if a schedule that can meet all the requirements of all communication pairs exists.

1 Introduction

Nowadays, the number of different types of devices connected to the internet and different networks is increasing at a fast pace. These new devices open up new opportunities and new user cases, which require different properties from the network that standard Ethernet could not provide. Some devices require transmission of time-sensitive data, which are valid for a limited time. Time criticality is usually used in real-time systems where it is not only important to get the data, but get it at the right time [1]. In case of a printer that works through a LAN network, there is no need for predictability. It is of course convenient to get the data quickly, but it is not a network requirement. In multiplayer-games on the other hand, where players depend on actions from other players, the requirements of fast data transmission are much higher. These requirements are even higher when video and audio are streamed with high quality. Failing to get the data in real-time may result in

skipping frames and overall lower quality. For safety-critical systems, it is even more important to meet the timing requirements. Imagine a space station where losing data during transmission could be life endangering. These real-time systems need data prioritization to make sure that the important data get through.

Real-time data communication differs from regular network communication in means of critical data streams. Time synchronization for network nodes and hosts should be usually less than 1 μ s. A software for resource reservation of critical data streams should be present (i.e. buffers and schedulers in network nodes and bandwidth on links), via configuration, management, and/or protocol action. Software and hardware to ensure extraordinarily low frame loss ratios and a guaranteed end-to-end latency for a reserved flow should be added [2].

In network communication, there are more types of frames. The most common ones are called Best Effort (BE) and they do not guarantee the delivery of packets. The delivery depends on a variety of factors such as the congestion of a network [3]. In contrast Time-triggered (TT) frames guarantee data delivery. The data is commonly obtained periodically from sensors, so the communication of TT frames can also be periodical [4]. This characteristic of periodicity means that TT frames can be scheduled. The predictability of such a scheduled traffic would mean that all data get delivered exactly when it is needed.

In a larger network with a great number of nodes scheduled traffic can be a problem, because of conflicts between nodes and their data transmission. To solve these possible conflicts inside the network and to ensure the delivery of the data, an algorithm is needed for scheduling.

2 Related work

2.1 Time-sensitive networking

Time-sensitive networking (TSN) proposes that different types of data can have different priorities

* Bachelor degree study programme in field: Computer Engineering
Supervisor: Lukáš Kohútka, Institute of Electronics and Photonics, Faculty of Electrical Engineering and Information Technology STU in Bratislava

and these should be distinguished. In a network with diverse packet priorities it is important to recognize these priorities in every node in the network and interact with them accordingly. This was the aim of IEEE 802.1Q standard and substandard IEEE 802.1p, that allows us to define levels of prioritization of data frames [5]. There are eight levels of priority available. The way the traffic is treated when assigned to a level of priority is undefined and left for implementation, however IEEE has made some board recommendations listed in Table 1 [6]. These levels of priority are an advantage if we have a system with a wide variety of data types, but in a system where all the packets would have the same level of priority (e.g. same type of machines in a factory), this becomes pointless.

Table 1. Levels of priorities as recommended by IEEE [7].

| Priority | Traffic types |
|-------------|------------------------------------|
| 0(lowest) | Background |
| 1 (default) | Best Effort |
| 2 | Excellent Effort |
| 3 | Critical Applications |
| 4 | Video, < 100 ms latency and jitter |
| 5 | Voice, < 10 ms latency and jitter |
| 6 | Internetwork Control |
| 7(highest) | Network Control |

2.2 Time aware shaping

Prioritization alone cannot always guarantee that the data is received on time. A situation where a device is sending critical data but this data has to be delayed can occur. If there is already a lower priority data being transmitted, it has to be finished before anything else can be transmitted. Time aware shaping solves this problem by using a guarding band, as shown in Figure 1. The guard band time must cover the time of transmission for the maximum size frame. This will work even for the worst case, but is not an ideal solution as some frames may have to wait until the end of the protected transmission, even though they could be transmitted more efficiently [8].

2.3 Credit based shaping.

Guard bands cannot be used for all data either, because treating all data special treats no data special. In some cases, the continuous stream of data is important e.g. video streaming. To provide such a continuous stream credit-based shaping can be used, so data can be transmitted more evenly. Credit-based shaper allows data to be transmitted if there are positive or zero credits left. Credits are earned while there is no transmission and consumed when data is being transmitted [9]. This leads to fairness in the data transmission, but a stream may have the highest

priority, as long as it does not have credits, it is not transmitted.

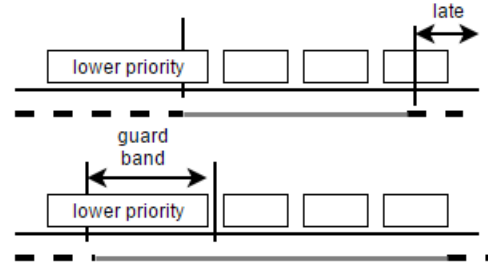


Figure 1. Importance of a guard band [8].

2.4 SMT Solver

Satisfiability modulo theory (SMT) is used in informatics and mathematics for finding solutions by making decisions. In SMT solving, there are constraints over individual variables ranging over integer, reals, bit-vectors, datatypes, and arrays. These constraints can involve theory operations, equality, and inequality [10]. In terms of scheduling, this could mean a possible formula for a network that could determine whether scheduled transmission can be possible taking into consideration all parameters and nodes. The SMT Solver is capable of solving a problem fast, but offers a really high level of abstraction and finding a formula is onerous.

3 Proposed algorithm

In the previous solutions, we could either determine if a schedule is possible for the given network configuration or make a schedule that is not effective for larger networks. Our proposed algorithm determines if a solution in terms of a schedule is possible and if yes, then designs a schedule. In addition to that, the proposed algorithm is trying to create a schedule as effective as possible in terms of the overall communication latency.

The functionality of the algorithm is illustrated in Figure 2. Inputs of the algorithm are the topology of the network, physical parameters of the network (such as latency towards routers, speed of the cables, latency of the devices and length of the cables) and for every communication pair the source, destination, period and size of data. The algorithm then creates a weighted graph and calculates the lowest latency path for each communication pair. Then the algorithm chooses the communication line with the most possible conflicts.

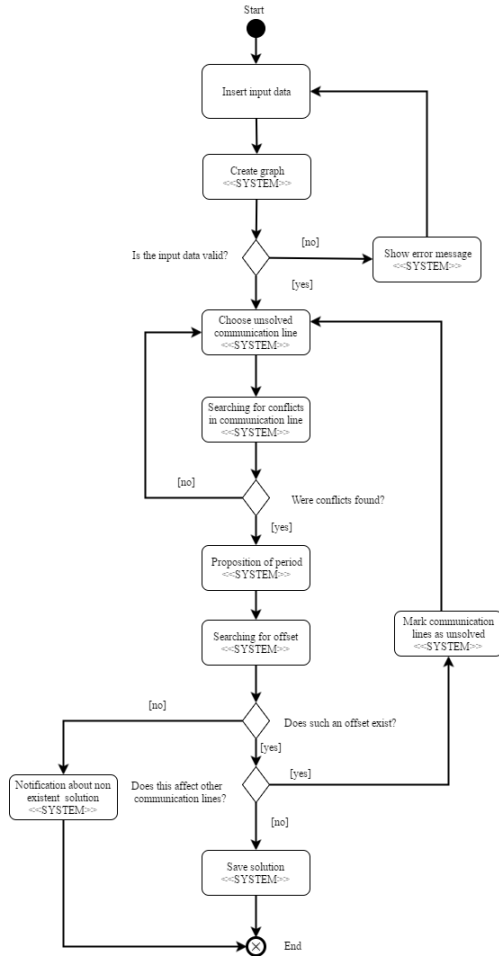


Figure 2. Activity diagram of the proposed algorithm.

If conflicts were found on the communication line, then the proposed algorithm calculates the period based on the lowest common multiple of the periods of the communication pairs. The next step is finding the offset for individual packets in a way that they do not block each other on the communication line. The offset is calculated from the throughput and other physical parameters of the network. If such an offset does not exist a solution is not possible for the given configuration. If such an offset exists and it affects other communication lines and these were already solved, then the algorithm marks them as unsolved and will calculate them later again. This way if a solution is possible, the algorithm finds it and saves the schedule. To make sure, that the proposed algorithm will not infinitely loop, a maximum of possible iterations is given on the number a communication line could be marked as unsolved.

The risk is that if the number is too low, then even if a solution is possible, the algorithm will not find it.

3.1 Schedule example

In the topology in Figure 3, we have two routers and four end point with three communication pairs altogether. As seen below a conflict takes place on the communication line C_1_2 when more communication pairs try to transfer data at the same time.

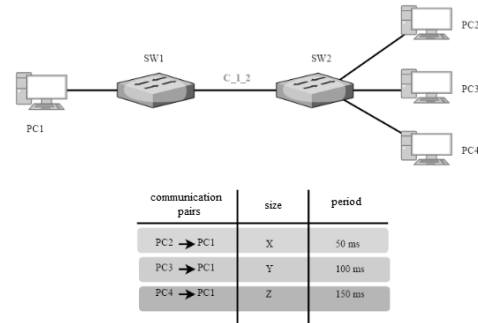


Figure 3. Simple example topology.

After inspecting the period of each communication pair (50ms, 100ms and 150ms) we can easily sketch up where the conflicts are bound to happen, as shown in Figure 4. The communication pairs want to send different sized frames at the same time which leads to multiple conflicts on the line. This can be solved by calculating the lowest common multiply which is 300 ms in this case and after this time frame, the communication would repeat itself periodically every 300ms. Now our algorithm calculates the amount of time that it would take for each communication pair to send a packet. If it is not possible to schedule the packets periodically into the 300 ms then no solution exists.

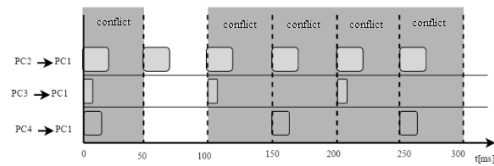


Figure 4. Illustration of conflicts.

4 Future work

We want to implement this algorithm into a free program that allows users to find scheduling solutions for their network. In the foreseeable future, this would mean a simple GUI for navigation, but could develop into a more user-friendly one. For example, the network topology could be added in a graphic drag-and-drop way.

5 Conclusion

In this paper, a new scheduling algorithm for communication in distributed real-time systems is presented. In comparison to the existing algorithms, the proposed one tries to lower the overall latency and takes into account the physical parameters of the network.

Acknowledgement: This work was supported in part by the Slovak Research and Development Agency under grant APVV-15-0254 and by the Slovak Republic under grant VEGA 1/0905/17.

References

- [1] H H. Kopetz.: Real-time systems: design principles for distributed embedded applications. Springer Verlag, (2011)
- [2] BOF Status: Deterministic Networking. [Online; accessed September 19, 2014]. Available at: <https://www.iab.org/wp-content/IAB-uploads/2013/01/tsn-nfinn-Deterministic-Networking-BOF-0914-v1.pdf>
- [3] M. Stamp: Information Security: Principles and Practice, John Wiley & Sons., (2011).
- [4] R. Obermaisser: Event-Triggered and Time-Triggered Control Paradigms, Springer Science & Business Media, (2006).
- [5] Institute of Electrical and Electronics Engineer, Inc: IEEE. Ieee p802.1qbv/d2.1. Standard (draft), (2014)
- [6] Institute of Electrical and Electronics Engineer, Inc.: IEEE 802.1 Standard, Available at: <http://www.ieee802.org/1/>
- [7] Institute of Electrical and Electronics Engineer, Inc.: IEEE. Ieee std 802.1q-2005. Standard, (2006)
- [8] M. D. Johas Teener: IEEE 802 Time-Sensitive Networking: Extending Beyond AVB. Available at: https://standards.ieee.org/events/automotive/08_Teener_TSN.pdf
- [9] K. Chan, B. Bensaou, and D. Tsang. Credit-based fair queueing (cbfq). ELECTRONICS LETTERS, 33(7):584–585, (1997).
- [10] L. de Moura: SMT Solvers: Theory and Implementation, Oregon (2008) Available at: <https://leodemoura.github.io/files/oregon08.pdf>

Methods of Automatic Brain Tumor Segmentation in 3-Dimensional MRI Data

Róbert BIRKUS*

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

Abstract. The segmentation of anomalies in anatomical organs from a medical data is a very important and challenging task. The manual segmentation requires high concentration and expertise of physicians and it is very time-consuming. In radiotherapy, an imprecise segmentation of tumor could have a devastating effect. Therefore, a precise automatic segmentation is highly required. In this paper, we present our two methods for automatic brain tumor segmentation in MRI data. The first method is based on deterministic algorithms of computer vision and image processing. The second method is based on deep learning. It is classifying individual superpixels according to calculated features.

1 Introduction

Nowadays there are various modalities of medical imaging [5] such as MRI, CT, and PET, which help physicians diagnose various diseases of patients without the need for surgery. However, these modalities produce a huge amount of medical data which is needed to be analyzed or segmented by the physicians. These processes require high expertise and concentration of the physicians and because of the amount of data, it is also very time-consuming. Here can be the different techniques of image processing, computer vision and machine learning very useful. The automatic or semi-automatic analysis and segmentation of the medical data help physicians to establish the diagnosis of patients better and quicker with less effort.

2 State-of-the-art

The current state-of-the-art methods are mostly built on machine learning techniques such as neural networks and decision trees. However, we believe that even a method based on deterministic algorithms can

achieve state-of-the-art results if it is complex enough.

Dvořák et al. [2] proposed a brain tumor segmentation method based on Convolutional Neural Networks (CNN) which focuses on strong intensity correlation of nearby pixels. To solve this problem they used local structure prediction approach (local image patches, local label patches). They were experimenting with different parameters of this approach and they achieved in whole tumor segmentation an average dice score of 0.83.

Zhao et al. [7] proposed a brain tumor segmentation method by integrating a Fully Convolutional Neural Network (FCNN) and Conditional Random Fields (CRF), instead of using the CRF as a post-processing step of FCNN. The CRF could be formulated as a Recurrent Neural Network (RNN), which making possible to use it integrated with FCNN and creating a deep network together. Their model is trained in 3 steps and they are using slice-by-slice segmentation which is faster than patch-by-patch. In the first step, they use image patches to train an FCNN. In the second step, they use image slices of axial view to train the CRF-RNN with parameters of the FCNN fixed. In the final step, they use image slices to fine-tune the whole network. In whole tumor segmentation they achieved average dice score of 0.87.

3 Related work

Thresholding is the simplest image segmentation method. It is a conversion of a grayscale image into a binary image based on some defined threshold value.

Region Growing (RG) is a simple region-based segmentation method. It initials some seed pixels, then it is examining the neighboring pixels and determines whether the pixel neighbors should be added to the region.

* Master study programme in field: Information Systems

Supervisor: Assoc. Professor Vanda Benešová, Institute of Computer Engineering and Applied Informatics, Faculty of Informatics and Information Technologies STU in Bratislava

BRATS 2015 dataset, *menze2015brats*, *kistler2013dataset* is a unique dataset of multi-contrast magnetic resonance (MR) scans of low and high-grade glioma patients. The BRATS dataset also contains scans generated using tumor image simulation software.

The *Microsoft Cognitive Toolkit* (CNTK) is a unified open-source deep-learning toolkit by Microsoft Research. It can be used as a library for C++ or Python programs, or as a standalone machine learning tool through its own model description language (BrainScript).

Neural Networks (NN) are a computational approach which loosely modeling the working of a biological brain. The neural networks contain nodes which are similar to neurons and connection between these nodes which are similar to the synapses. Neural networks can be taught to solve different classification problems based on the input data and the expected output.

Simple Linear Iterative Clustering (SLIC) [1] is a clustering segmentation method based on the k-means algorithm. We use its parallel version on the graphical processing unit called gSLICr [4].

Superpixel is a group of connected pixels with similar properties. It is the product of the *SLIC* algorithm.

4 Proposed methods

In this section, we are introducing our two proposed methods for automatic brain tumor segmentation in MRI data. The first method we built on different deterministic algorithms of computer vision and image processing. The second method is based on deep learning and we classify superpixels according to their calculated features.

4.1 Method 1 - using deterministic algorithms

As we already mentioned the first proposed method is based on deterministic algorithms, concretely on optimal thresholding and region growing algorithm. In Figure 1 we can see the flowchart of the method. As the input, we used the FLAIR sequence of the MRI dataset because the whole tumor is visible on this sequence the most. In the first step, we applied a slice-wise optimal thresholding. The optimal threshold value we determine according to our own histogram analyses algorithm.

In Figure 2 we can see the histogram of a slice. If we ignore the black background we get a bimodal histogram with two peaks. The first (bigger) peak is representing the healthy cells of the brain (healthy peak) and the second (smaller) peak is representing tumorous cells of the brain. The optimal threshold

value lays between these two peaks. We determine it by checking the inclination angle from the top of the healthy peak to the bottom until it falls below 10° .

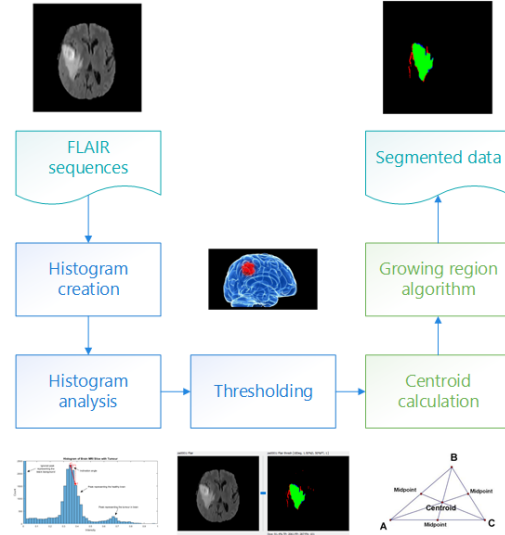


Figure 1. Flowchart of method 1 (first three steps processed in 2D, last two steps processed in 3D).

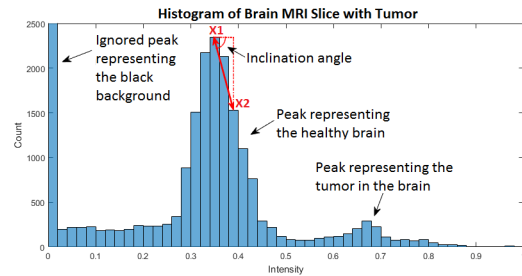


Figure 2. Histogram of an MRI slice with descriptions.

After the application of optimal thresholding we use the region growing algorithm to the output of the thresholding to remove separated false positive regions from the segmentation to improve the results. We process the data in 3D. In the first step, we determine the seed of the region growing algorithm by calculating the centroid of the thresholded data. If the centroid does not hit a voxel belonging to the thresholded part, we look for the nearest voxel that belongs to it. This centroid is the seed for the region growing algorithm. We let the region grow only in thresholded voxels.

4.2 Method 2 - using deep learning

For the second method based on deep learning, we use The Microsoft Cognitive Toolkit (CNTK). We use the CNTK for classifying superpixels calculated from the slices of the MRI volumes. For the superpixelisation of slices, we use a publicly available implementation called gSLICr [4]. We teach the neural networks to classify the superpixels according to the following features of the superpixel:

1. Intensity histogram of the superpixel
2. Intensity histogram of the superpixel and its neighboring superpixels
3. Shannon entropy (Equation 1) of the intensity histogram from point 2
4. Euclidean distance between the superpixel centroid and the brain centroid (Equation 2)
5. XY, XZ and YZ angles between the superpixel centroid and the brain centroid

The equation of the Shannon entropy:

$$H(X) = - \sum_{i=1}^n P(x_i) \log_p P(x_i) \quad (1)$$

The equation of the Euclidean distance in 3D:

$$dist = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \quad (2)$$

We selected most of these features because Tamajka et al. [6] achieved good classification results with similar supervoxel features in an automatic brain segmentation method. Intensity is the key information which an MRI can give us. Therefore, we selected an intensity histogram of the superpixel as one of the features. The context is also very important in tasks like detection, segmentation, and classification. Therefore, we also selected features based on the neighboring superpixels. The entropy should tell the network about the similarity between the current superpixel and its neighbors. We also selected features based on spatial position because we did not want to rely on intensity only.

5 Evaluation

In Figure 3 we can see the evaluated segmentation results of the first method based on deterministic algorithms using the dice score. We can see that on some data we achieved state-of-the-art results and even better. However, on the other hand, there are data where the results are not that good or even horrific. This method is too data specific and therefore the average dice score is only 0.7235.

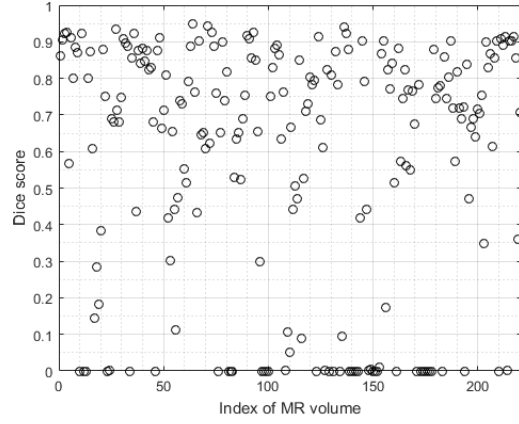


Figure 3. Evaluation of method 1 segmentation results (220 patients).

The dice scores close to zero is caused by a wrong seed selection for the region growing algorithm. In case, when the seed is selected from a separated false positive region it has no chance to grow into the tumorous part. This problem could be solved by a more sophisticated seed selection method or by an interactive seed selection which would cause a semi-automatic segmentation instead of fully automatic segmentation. However, this is not a problem because usually a manual segmentation is done by two physicians in practice and the goal of the automatic or semi-automatic segmentation is to replace only one of these physicians, not both. One physician always has to be there because we cannot fully rely only on algorithms in medicine.

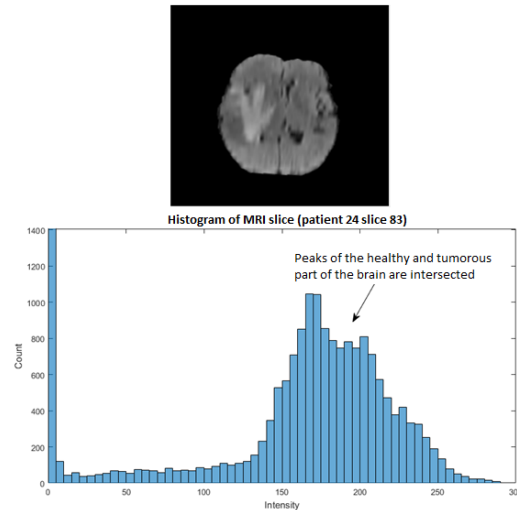


Figure 4. Example slice with its histogram where the tumor has similar intensities as the healthy part of the brain.

Another reason of low dice score on some data is the limitation of thresholding. In Figure 4 we can see an example slice with its histogram where the tumorous part of the brain has very similar intensities as the healthy parts of the brain. Therefore, a simple thresholding method cannot separate the tumor from the healthy brain accurately. A possible solution would be to identify this kind of data according to its histogram and then process it differently than just by a simple thresholding.

The results of our second method based on deep learning is limited by the quality of the calculated superpixels. Therefore, we also needed to evaluate the quality of the superpixels. For this purpose, we used the undersegmentation error metrics [3] (see Equation 3). In Table 1 we can see the average and the median undersegmentation error of the calculated superpixels.

$$UE = \frac{1}{N} \left[\sum_{S \in GT} \left(\sum_{P: P \cap S \neq \emptyset} \min(P_{in}, P_{out}) \right) \right] \quad (3)$$

Table 1. Undersegmentation error of superpixels.

| Average | Median |
|---------|--------|
| 0.0978 | 0.0510 |

Our proposed neural network is currently achieving 16.867% error rate in superpixel classification. However, we plan to improve the performance of our neural network by different network configurations and experiments with the input features. The segmentation results based on the classification of superpixels is not yet evaluated. However, based on the quality of superpixels and classification accuracy they look promising.

6 Conclusion and future work

We have presented our two automatic brain tumor segmentations methods in MRI data. One of them is based on deterministic algorithms of computer vision and another is based on deep learning. The first method based on deterministic algorithms was evaluated. The segmentation quality was very data specific. On some data, we achieved even state-of-the-art results, but on the other hand, some data achieved very poor results. However, we plan to handle the identified lacks of the method to improve the segmentation

quality. For the second method based on deep learning, we evaluated the quality of the superpixels and the classification accuracy of our trained model. We have not evaluated the segmentation results yet, but based on the quality of superpixels and the classification accuracy they look promising.

For the future, we consider appending our method based on deterministic algorithms with some interaction. We believe that a minimal interaction from the side of physicians can significantly improve the segmentation quality. Depending on the segmentation results of the second method based on deep learning we would like to experiment with the features of superpixels and the configurations of the neural networks. We also would like to try out sagittal and coronal slicing in addition to the axial slicing of the 3D MRI data. And it would be also interesting to try out supervoxels instead of superpixels in the classification process.

Acknowledgement: The authors would like to thank for contribution to the Research and Development Operational Programme for the project “University Science Park of STU Bratislava”, ITMS 26240220084, co-funded by the European Regional Development Fund.

References

- [1] Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Susstrunk, S.: SLIC superpixels compared to state-of-the-art superpixel methods. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 2012, vol. 34, no. 11, pp. 2274–2282.
- [2] Dvorak, P., Menze, B.: Structured prediction with convolutional neural networks for multimodal brain tumor segmentation. *Proceeding of the Multimodal Brain Tumor Image Segmentation Challenge*, 2015, pp. 13–24.
- [3] Neubert, P., Protzel, P.: Superpixel benchmark and comparison. In: *Proc. Forum Bildverarbeitung*, 2012, pp. 1–12.
- [4] Ren, C.Y., Prisacariu, V.A., Reid, I.D.: gSLICr: SLIC superpixels at over 250Hz. *arXiv preprint arXiv:1509.04232*, 2015.
- [5] Sala, E., Freeman, A.H., Lomas, D.J., et al.: Chap. 1. In: *Radiology for surgeons in clinical practice*. Springer Science & Business Media, 2008, pp. 1–14.
- [6] Tamajka, M., Benesova, W.: Automatic brain segmentation method based on supervoxels. In: *Systems, Signals and Image Processing (IWSSIP), 2016 International Conference on*, IEEE, 2016, pp. 1–4.
- [7] Zhao, X., Wu, Y., Song, G., Li, Z., Fan, Y., Zhang, Y.: Brain tumor segmentation using a fully convolutional neural network with conditional random fields, 2016.

Coreference Resolution in Text

Lukáš MIŠKOVSKÝ*

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

Abstract. Coreference resolution is an important sub-task of natural language processing, yet available systems often use shallow features and misjudge special cases as a result. We designed a rule based coreference system and constructed new rules to identify such cases. By exploring the impact of these rules, we can evaluate their usefulness on the overall system in contrast to other widely used rules in this field.

1 Introduction

Natural language processing is becoming increasingly more popular and needed because of the exponential growth of created information such as text. We can use it to automatically create meaningful summarizations, extract key information, translate texts in foreign languages and perform many other important tasks. However, a deeper understanding of the processed text is needed to provide accurate results. This is not a trivial task for computers.

Coreference resolution is a process in which we wish to identify a real world entity and find all of the other text references to such entity. The majority of coreferences is comprised by pronouns, acronyms and aliases. State of the art systems are using mostly string and distance related features combined with many types of dictionaries to create precise coreference identifying rules. The identification of coreferences is so complicated that it was even proposed as a more accurate form of Turing test [4].

Coreferences can be used as a basic input information to improve systems that are aimed at solving the high level tasks of natural language processing.

Coreference resolution, as well as other tasks in natural language processing, is highly dependent on the language used. In our work we analysed state of the art systems for English language in order to identify some of the special cases that are not correctly recognized as coreferences. Our goal is to create new

coreference identifying rules for English with the help of other sources of information. These are then independently evaluated and compared to other established rules, so we can conclude whether they can enhance the accuracy and precision of a coreference resolution system.

2 Related work

Coreference resolution consists of three main steps and an additional one. First step is preprocessing. Its input is a raw text and as an output we get tokens with information about them such as part-of-speech tags. In the second step called processing, all the available information used to coreference resolution is extracted by the system using dictionaries, basic rules or 3rd party knowledge repositories. Finally, there is a linking step in which algorithms decide whether given two candidates should be coreferently linked or not. The additional step is visualization of the results.

We analysed four main systems and their results are summarized up in Table 1.

Stanford system [3] is the oldest and most used. It uses rule sieves that are constructed from the ones with highest precision rates and low recall to some that have higher recall but are a lot less precise.

Berkeley system [2] is different because it is not aimed only at coreference resolution. Their researchers are trying to solve a few NLP subtasks at once, using outputs of these tasks for mutual improvement of other parts of this system.

Illinois coreference system [1] is focused on a few basic characteristics. Their goal is to prove that you can create a fairly good coreference system using only a few characteristics and simple weighing function that determines the existence of coreferent link between two candidates. Later, they joined it with another task [7], because of the same assumptions as in a previous system.

* Master study programme in field: Information Systems

Supervisor: Miroslav Blšták, Institute of Informatics, Information Systems and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

The last system is named CORT or coreference resolution toolkit [5]. The goal is to provide a set of tools that can be used to develop better coreference resolution systems. The main part of this system is highly customizable linking step using either pairing, ranking or tree ranking of candidates.

The correctness of a coreference system is measured by a few slightly different metrics. All of them are based on precision and recall, albeit calculated in different ways. For example MUC is really straightforward but if one would connect in his system all coreferent candidates into one entity he would get a fairly good results because of the flaws of this algorithm.

Table 1. Results of different coreference resolution systems.

| System | MUC | B^3 | CEAF (E) | Avg. |
|----------|-------|-------|----------|-------|
| stanford | 65.0 | 54.5 | 51.3 | 55.93 |
| berkeley | 71.24 | 58.71 | 55.18 | 61.71 |
| illinois | 72.22 | 60.50 | 56.37 | 63.03 |
| CORT | 72.17 | 59.58 | 55.67 | 62.47 |

3 Coreference resolution process

We noticed a few problematic cases that were not marked as coreferences in state of the art systems and came up with a new and modified rules to identify them. Such cases include alternative names of named entities that can not be identified without prior knowledge, appositional terms or location coreferences through a "there" adverb. Example sentences that contain mentioned cases are "We went on to *Paris* and stayed *there* eleven days", "*The Alaska Marine Highway* or *the Alaska Marine Highway System...*" as well as the sentence in Figure 2.

Our proposed method is focused only on coreferences inside of compound sentences. We decided for this step in order to model the special situations inside of the sentences more accurately and to refine our field of interest to a more constrained task.

The method can be decomposed to three main parts. First is a preprocessing with the use of CoreNLP suite. However, at the very beginning we have to break down the given text into sentences and choose some that have potential coreferences inside.

Second step is extraction of necessary information from the initial data stored in our database. For this we use tools such as WordNet, Google Knowledge graph, dictionaries, two special parsers and few other simple rules. The output is a list of candidates with extracted information.

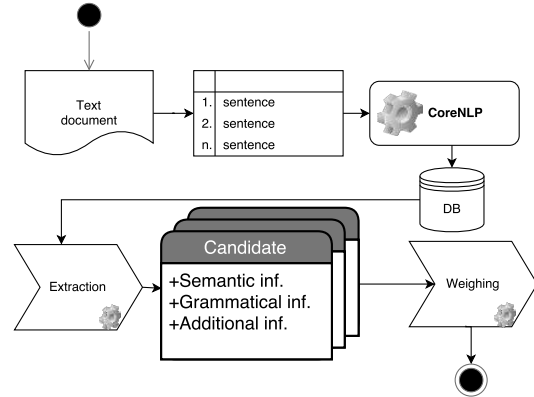


Figure 1. Graphical overview of the proposed method for our coreference resolution system.

The last step is weighing, in which we pair our potential candidates and decide whether the coreferent link exists based on our implemented rules. After pairing we have a representative candidate that was mentioned in the sentence first and a mention that was written later. The process is shown in Figure 1.

3.1 Base rules

The first set of rules is traditional and widely used. It relies mainly on string similarity and as a result we can achieve high precision rates but lower recall.

- *exact match* - true if the two candidates are the same phrases (excluding stop words).
- *partial match* - fulfilled if each token of a mention exists in a representative.
- *head noun match* - if the main nouns of both phrases are identical the rule is true.
- *abbreviation* - we implemented a simple algorithm to detect abbreviations.

We plan to utilize these rules to create a baseline system. With the established baseline, we can more effectively test new rules and determine their usefulness.

3.2 Modified rules

Rules that follow are the ones that we came up with to solve some special coreferent cases. We are testing and improving them to minimize the error rates.

- *relation 'or'* - with this rule we can detect coreferences connected with 'or' using simple dependencies and context.
- *'nsubj' relation* - two candidates connected with nominal subject relation are likely to be coreferent.
- *'appos' relation* - useful to detect apposition in a sentence, that is probably coreferent.

- *Knowledge Graph names* - we are able to connect well known names for named entities using Google's Knowledge Graph.
- *typo detection* - with Levenshtein distance we are able to detect coreferences between misspelled phrases that are likely to occur in real world texts.
- *here/there* - these adverbs are often coreferent with locations mentioned previously in the sentence.
- *hypernym chain* - the rule is true if head nouns of two phrases are connected with hypernym relation in WordNet [6] database.
- *synonyms* - same as above but with synonyms instead.

Relations are extracted from CoreNLP's basic dependencies. The annotator provides information about which tokens are connected by which dependencies. We then extract these dependencies to a candidate level. We do so by proclaiming that the relation exists if any token from the representative candidate is connected by the dependency with any other token from the mention candidate.

Google's Knowledge Graph provides REST API that returns the most similar entities for a text query. If the results have a matching score that is above our threshold and the score of each result is at least half of the previous then we take the first three results and save them as alternative names for a given entity.

Two of the rules are displayed in Figure 2 as an illustrative example. The rules are naturally overlapping, in this example appos and abbreviation specifically. This fact can be utilized in more complex ranking instead of currently implemented sequential scanning of available rules.

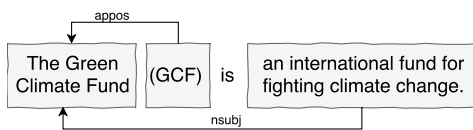


Figure 2. Example of a nsubj and appos rules.

4 System prototype

The system was implemented in Java according to our specification. It consists of different modules and it is capable of producing an output in standard CoNLL-2011/2012 format, intended for further testing or visualization.

Modules are used to solve subtasks such as parsing a parse tree to obtain candidates, merging relations between tokens to a higher level of relations between phrases or communicating with 3rd party knowledge sources.

5 Evaluation

During the implementation we tested our rules on a few random sentences copied from the web. After a significant part of our system was finished, we tried to obtain annotated dataset for more significant testing but every dataset used in previous works was licensed and unavailable.

We decided to put together a few sentences with approximately 40 coreferent links. This little sample is used as an indication of the system's potential. Further evaluation will be needed with bigger, hand-annotated dataset.

Since we had no format limitations from the dataset, we decided to use official implementation of various scoring algorithms [8]. For this purpose we created a parser to CoNLL-2011/2012 format and changed the output of our system to this format as well. Available metrics include MUC, B^3 , CEAF and BLANC. Each metric measures the correctness of coreferent links a little differently, in order to eliminate shortfalls of the other metrics.

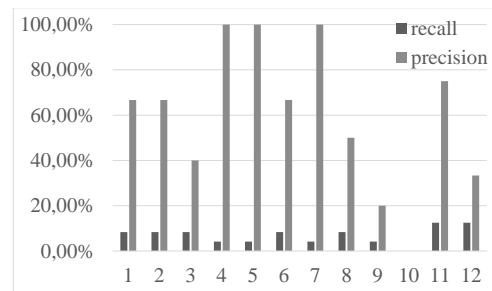


Figure 3. Comparison of suitability of different rules.

Overall we used twelve different rules and evaluated the usefulness of each one individually. Names of the rules are not important but the best one was nsubj (11th in Figure 3) with 12.5% recall and 75% precision and the worst is synonym match (10th in Figure 3) since it marks any two phrases that have head nouns in synonym relation as coreferent.

When we combined the rules and used all of them, the system achieved results summed up in Table 2. The last row contains the results of a state of the art system from Stanford (v3.6) evaluated on the same data sample. The results are promising, although we anticipate that they will change quite a bit after further testing on a larger dataset.

Table 2. Overall results of the initial testing.

| | MUC | B^3 | CEAF (M) | CEAF (E) |
|------------|-------|-------|-------------|-------------|
| precision | 48 | 51.49 | 58.13 | 54.28 |
| recall | 50 | 52.71 | 64.1 | 65.14 |
| F1 | 48.97 | 52.09 | 60.97 | 59.22 |
| CoreNLP F1 | 47.61 | 47.95 | 56.33 | 54.48 |

6 Conclusions

In our work, we successfully implemented a software prototype for coreference resolution. Initial testing shows promising results with rules tailored for special cases that other systems are unable to identify. The F1 measure of our system was slightly above 50% which is comparable to other research in this area and can be further improved by removing identified bad rules and refining the promising ones.

More thorough testing on bigger dataset will be needed in order to fully determine the usefulness of our approach.

7 Future work

In future work we plan to complete two main tasks. First and the most important is an annotating module to ease the creation of the annotated dataset. Afterwards we need to create mentioned dataset, so we can test our system more thoroughly.

Second task consists of testing and refining of the rules that we use in our system and possibly incorporating another source of information that can be used in these rules.

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

References

- [1] Bengtson, E., Roth, D.: Understanding the Value of Features for Coreference Resolution. In: *EMNLP*, 2008.
- [2] Durrett, G., Klein, D.: A Joint Model for Entity Analysis : Coreference , Typing , and Linking. *Transactions of the Association for Computational Linguistics*, 2014, vol. 2, pp. 477–490.
- [3] Lee, H., Peirsman, Y., Chang, A., Chambers, N., Surdeanu, M., Jurafsky, D.: Stanford’s Multi-pass Sieve Coreference Resolution System at the CoNLL-2011 Shared Task. In: *Proceedings of the Fifteenth Conference on Computational Natural Language Learning: Shared Task*. CONLL Shared Task ’11, Stroudsburg, PA, USA, Association for Computational Linguistics, 2011, pp. 28–34.
- [4] Levesque, H.J., Davis, E., Morgenstern, L.: The Winograd Schema Challenge. In: *Proceedings of the Thirteenth International Conference on Principles of Knowledge Representation and Reasoning*. KR’12, AAAI Press, 2012, pp. 552–561.
- [5] Martschat, S., Strube, M.: Latent Structures for Coreference Resolution. *Transactions of the Association for Computational Linguistics*, 2015, vol. 3, pp. 405–418.
- [6] Miller, G.A.: WordNet: A Lexical Database for English. *Commun. ACM*, 1995, vol. 38, no. 11, pp. 39–41.
- [7] Peng, H., Chang, K., Roth, D.: A Joint Framework for Coreference Resolution and Mention Head Detection. In: *CoNLL*, University of Illinois, Urbana-Champaign, Urbana, IL, 61801, ACL, 2015, p. 10.
- [8] Pradhan, S., Luo, X., Recasens, M., Hovy, E., Ng, V., Strube, M.: Scoring Coreference Partitions of Predicted Mentions: A Reference Implementation. 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. 30–35.

Automatic Detection and Notification of Bumps on Roads to Drivers

Adam STRÁSKY*

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

Abstract. Based on current road conditions, it is necessary to focus on their improvement and modern technologies can be in help. Combining crowd sourcing ideas with smart phones, it is possible to map the current state of the roads, to find holes and bumps. Mapping road problems can prevent damage and often repairs of cars. Using our Android application users as well as competent persons will get information about roads' problems which need to be repaired. The main purpose of the application is to automatically detect bumps and show them to users and authorities. Whenever a car hits a bump, our algorithm evaluates it and the bump is marked into map in real time. Manual addition of bumps without necessity to hit it by car is also available in the application. The application shows notifications to drivers whenever there are approaching bumps in the road, thus making driving much safer.

1 Introduction

Bumps Detector is application that automatically detects bumps on roads with use of accelerometer. For added convenience, application will warn drivers for upcoming bumps, so they have enough time to react and avoid impact. Because of the internet data usage application will be available in offline mode. Bumps can be added manually, this way drivers do not need to make contact with bump, they can click on map and bump will be recorded. To see bumps anywhere on the map, users can visit webpage to get insight on the current state of roads. Application is based on the initial project realised in Detroit, where the system detected the bumps using accelerometer sensor. Accelerometers sensors were placed in taxi cars collecting data about the state of the roads, which were provided to the Municipality of the town in order to be able to solve this issue. Our intention is to move away from the idea of collecting data for the Municipality towards providing this data to the

drivers in order to harness these information and take advantage of them.

2 Our proposed system

Detection of bumps is based on obtaining data from accelerometer sensor. Based on change of acceleration while moving, device will evaluate the movement and when threshold is crossed it will be recorded. Geographic coordinates are added to each newly recorded value, new value will be compared with existing ones, followed by assessment whether the bump was created on the new coordinates. In that case bump will be evaluated as new one. In case when location of bump is within two meters, size of bump will be evaluated. If new size is greater than original size of bump, new size will be set. Based on changes in size of acceleration bumps are detected into three categories, big, medium and small [1].

Algorithm 1 Detect Bump.

```
1: procedure detectBump(detectedBump)
2:   for all getAllFindBumps do
3:     if (radius(detectedBump,bump) < 2) then
4:       if (SizeOfDetectedBump > bump) then
5:         updateSize(bump,SizeOfDetectedBump)
6:       else
7:         addNewBump(detectedBump)
```

Detected bumps are uploaded to the server. Bump sent from a device is compared with the database of the server containing the bumps from all devices. If bump in that place has not been detected yet, the new record is created. This record contains the location, the timestamp, the number and the intensity of concussion. In case that there has been created record on that location before, the amount will increase by 1.

Because of the internet data savings bumps are stored directly to the mobile device, this way bumps can be displayed on map without internet connection. The synchronisation is based on sending request by

* Bachelor study programme in field: Internet Technology
Supervisor: Dr. Tomáš Kováčik, Institute of Computer Engineering and Applied Informatics, Faculty of Informatics and Information Technologies STU in Bratislava

mobile devices in regular 5 minutes intervals. The device sends the primary key in the 14-km radius zone surrounding the actual position, which its local database contains together with its actual position. Afterwards, the server compares the received data with its database. The server generates the response regarding the missing data. Having received the response from the server, the device will update these records to the local database. The area of 14 kilometres is used because the coordinates are rounded to 4 decimal digits. By reducing the amounts of decimal places, it is possible to extent the area to 140 kilometres. On the other side, by increasing the number of decimal places, the area may be reduced to 1,4 km. If the different size of the area had been chosen, it would have been necessary to calculate the distance for each bumps separately, which would be compute-intensive.

Due to storage of bumps in offline mode it is suitable to use functions of offline map, which is secured through map service called Mapbox. User can directly select area which he wants to download or he can choose area of current position. Due to the recording system of bumps by accelerometer it is necessary for drivers to run into bumps. For this reason, alternative option to add bumps manually on map was designed. Adding bumps manually consists of choosing size of bump and marking its position on map.

In order to add the bump as quickly as possible, the simple widget is placed on the home screen. After pushing the button, the location of the device is registered automatically and therefore, the user does not have to turn on the application in order to add the bump. Considering safety of drivers, so they do not have to drive and use their device, voice service "OK, Google" was integrated. On simple voice command it will open application, or record bump in same way as widget mentioned above. When destination is specified, application is able to warn drivers for upcoming bumps. Detection of upcoming bumps is based on speed analysis and existing bumps on route. Speed analysis is done in two ways, because of inaccuracy caused by GPS signals. First method gets actual speed from GPS module every second and then average speed is calculated for specific time period. Second method is based on distance travelled for certain amount of time. From these two methods, higher average speed is chosen for further processing. Finding bumps on predefined route and analyzing distance from current position of device we can calculate approximate time of impact. Subtracting few seconds from impact time we can in advance warn drivers for possible collision.

The given algorithm shows basic principles of the function.

Algorithm 2 Nearest Collision.

```
1: procedure nearest_collision()
2: for all getAllFindBumps do
3:   if isBumpOnSelectedTrace(bump) then
4:     addCollision(bump)
5:   if (getSpeedGPS > getSpeedDistance) then
6:     bestSpeed ← getSpeedGPS
7:   else
8:     bestSpeed ← getSpeedDistance
9: nearestBump ← getNearest(position, collision)
10: return timeCollision(bestSpeed, nearestBump)
```

3 Conclusions

Despite attempts to display route this application cannot compete with fully navigational systems. Therefore, application can be used in background as additional service, while user can determine what degree of feedback will application offer. The offline maps used in non-commercial sphere are directly influenced by the size of the map, which can be downloaded. This fact may restrict the users during longer routes. The another drawback of automatical detection using accelerometer sensor is the human factor, when the user accidentally shakes the device. Therefore, it is highly recommended to use phone holder when driving. Despite the disadvantages, the application offers a helping hand to the drivers. With the further development, we intend to fix the shortcomings mentioned above and propose additional functionality which would be able to compete with navigation services.

References

- [1] Amick, R.Z., Patterson, J. A., Jorgensen, M. J.: Sensitivity of tri-axial accelerometers within mobile consumer electronic devices: A pilot study. *International Journal of Applied* 3, 2 (2013).

Proximity Based Access Control System with Encryption

Vladimír KUNŠTÁR*

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

Abstract. The goal of the paper is a description of the device, which uses proximity technology with encryption. The device authorizes and controls access to the building. The device is based on contactless technology, which works on 13.56 MHz with encryption support and on the microprocessor Atmel. This device consists of an operation unit, which can be connected to several devices. The operation unit can operate even in the case of the central server outage, using the data stored in the local database. These data is updated at a regular time intervals. The description of the device is based on an article that contains a detailed description of how the system works and where it can be most effectively utilized.

1 Introduction

Since we had previously paid attention to the development and creation of the hardware, we decided that description of the device which uses proximity technology with encryption is very interesting. We try to make wireless technology and the issue of transferring data in contactless technology understandable. In section II, we analyzed and described the available technology and encryption of transmitted data in contactless identification. We focused on the description of Mifare contactless cards due to their long history and good security. When we were designing the operation unit, we emphasized its security as well as overall reliability, flexibility and also its price. We used easily accessible and affordable parts. The design and communication of the operation unit is described in section III.

2 Related work

Nowadays great importance has been given to security technology and technology providing access control to the buildings. This equipment also belongs to devices using wireless technology such as proxy technology, Mifare, NFC, and so on. We will talk

about Mifare Ultralight C technology because of its prevalence and encryption support [4, 6, 7].

RFID was invented for contactless identification for a short distance. The operating principle is based on electromagnetic induction using amplitude modulation. A reading head sends a carrier wave at a certain frequency, for example 13.56 MHz, then RFID chip is charged, changes the amplitude and the reading head captures the reflected wave with varying intensity [5].

Mifare is based on a number of levels, which are specified in ISO / IEC 14443 Type A 13.56 MHz [1, 6]. The principle of the Mifare technology is similar to RFID. It operates at a frequency of 13.56 MHz and uses ASK modulation. The difference and probably the most significant improvement over the RFID is greater memory capacity and higher security of transferring data between a card reader and a contactless card. Security means to encrypt the whole communication and also the memory blocks with data. Mifare cards are used in banking applications and access control systems in buildings with a high degree of safety [2].

Mifare cards use 32-bit identification number that can be rewritten for certain types of cards. In addition to the identification number of the card it has size memory of 1K or 4K. The available memory is divided into blocks of 16 bytes and it is further organized into sectors [3].

Design and construction of contactless card defines the standard ISO / IEC 7810. This standard also defines the design of the other cards used in the world. Standard ISO / IEC 7810 was created in 1985 and it was created because of the outdated standard ISO / IEC 7813. Standard ISO / IEC 7810 regulates the aforementioned standard and it also comprises the card thickness and rounded corners of the card.

* Doctoral degree study programme in field: Applied Informatics
Supervisor: Assoc. Professor Tibor Krajčovič, Institute of Computer Engineering and Applied Informatics, Faculty of Informatics and Information Technologies STU in Bratislava

3 Operation unit and other systems

The goal was to design and create an operation unit, a program for the operation unit ("firmware"), utility program and database. The operation unit serves a variable number of reading heads and their corresponding number of locks. The operation unit could be also used to record arrivals and departures. The cards Mifare Ultralight C are used for authentication. Data encryption uses the 3DES algorithm.

The whole system consists of three parts. The first part is the operation unit Figure 1, which provides communication with the main database and other system components. The second part is a card reader Figure 2 and the last part is a lock control module Figure 3. The modules are connected to the control unit using RS485 bus. The block diagram of communication is shown in Figure 4.

Card reader

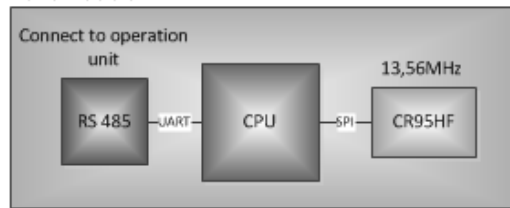


Figure 2. Card reader.

Operation unite

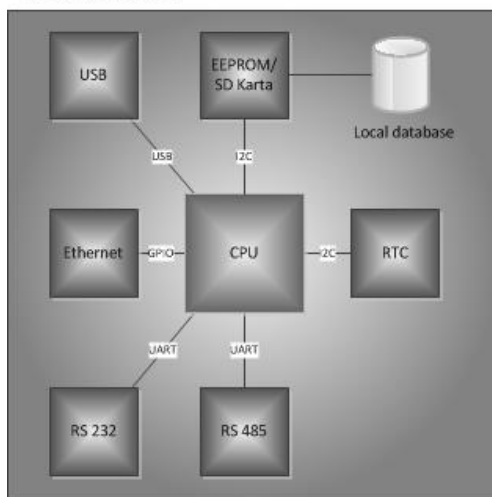


Figure 1. The operation unit.

Lock control module

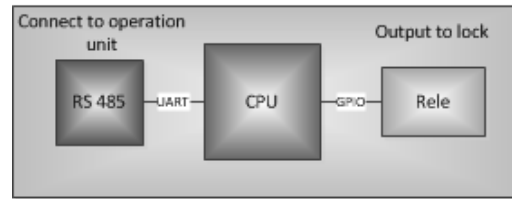


Figure 1. Lock control module.

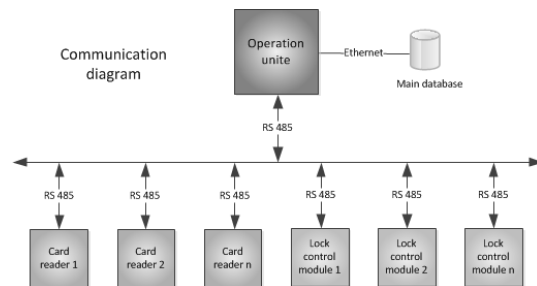


Figure 2. Block diagram.

The task of the operation unit is to mediate communication between devices that are connected to the operation unit using RS485 bus and the main database of the identifiers, which can be stored anywhere on an external server. The operation unit is connected to a network through the Ethernet port. The main database contains all data about access cards, individual approaches and all devices. The main database is a reflection of the local database stored in FLASH memory of the operation unit. Events from the system always enter into the local database first and then to the main database. Synchronization between the central and local databases are carried out either at regular intervals set by the user, or at a change in any of the databases, or when you push the button. The basic configuration of the device is possible via web browser.

The card reader provides communication between a contactless card and the operation unit. The card reader reads the serial number (UID) of the attached card and then it sends this UID into the operation unit via RS485 bus. The operation unit searches in the database record and record's private key and then authentication process starts.

The lock control module provides communication between the operation unit and controlled locks (or other devices). The control of this module depends on the outcome of the authentication. If authentication is successful, the operation unit

sends a signal to the lock control module and based on the received signal it chooses appropriate action.

The advantage of this device is secure change of the encryption key. This change takes place when the contactless card is added to the system. This is the reason why is unlikely breaking of the encryption key. Another advantage is, that this device has local database. The database provides data, when the central database of server goes down. All events are recorded to the local database, when this happens. Then, when the server is ready, both databases are automatically synchronized.

4 Implementation and results

The main part of the operation unit comprises a microprocessor Atmel ATSAM4E8C. The microprocessor is connected to the communication bus RS232, RS485, USB. It also contains signal elements such as a buzzer, LED diode, RTC module, four buttons, FLASH memory and power connectors Figure 5.

The main part of the card reader comprises Atmel Atmega8 and circuit CR95HF from STMicroelectronics. The microprocessor is connected to the aforementioned circuit. Which provides communication with contactless cards. The bus RS485, buzzer and addressing switch are also connected to the microprocessor. Addressing switch is used for setting the address required for identification in RS485 network. A slot for microSD card and connector for possible connection of keyboard and display via the I2C bus are also connected to the circuit Figure 6.

The main part of the lock control module is microprocessor Atmel Atmega8. Two relays are connected to the lock control module by optocouplers. These relays have the task of opening or closing the electromagnetic lock. The addressing switch recognizes and identifies the module in RS485 network. This addressing switch along with signaling LED, buzzer and slot for microSD card are also connected with microprocessor Figure 7.

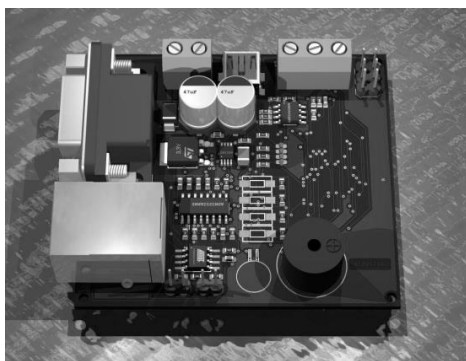


Figure 3. Operation unit.

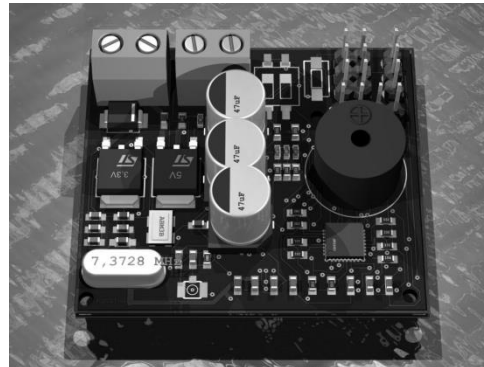


Figure 4. Card reader.

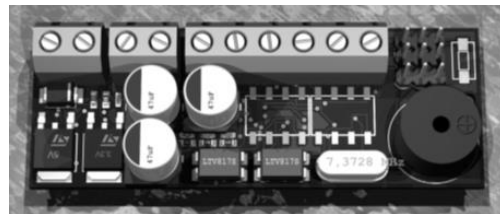


Figure 5. Lock control module.

As mentioned above, the modules communicate mutually with the operation unit and their number may be variable. The card reader and the lock control module are connected to the operation unit using RS485 bus. Once the contactless card is placed on the card reader, using the magnetic field, it reads the serial identification card number (UID), which is sent to the operation unit via RS485 bus. The operation unit receives the transmitted information, then searches for received UID in the database identifiers and reads its assigned authentication key. Authentication key is written to the card when the card is added to the system. When the record is found in the database identifier, the operation unit sends a command to the card reader to initiate authentication. The card reader transmits this command to the attached card and the authentication process begins. The contactless card generates a random 8 byte number RndB and encrypts it using the authentication key. This number is sent to the operation unit using the card reader. The operation unit decrypts this number. The operation unit generates its own random number RndA, then deterministic binary operation decrypts received number. The operation unit concatenates the binary number with RndA, encrypts and sends it to the card reader via RS485 bus. The card reader sends this operation to the placed card, the card decrypts the received data then it verifies the received RndB, performs deterministic binary operations on RndA, encrypts and sends to the operation unit using the card

Table 1. Packet structure for RS485 communication.

| Position | 0 | 1 | 2 | 3 | 4 | 5 -> LEN |
|----------|--------------|---------------|----------|-----|-----|----------|
| Size (B) | 1 | 1 | 1 | 1 | 1 | LEN |
| Name | PACKET_START | THIS_BOARD_ID | BOARD_ID | CMD | LEN | DATA |

Table 2. UID encryption keys stored in flash memory.

| BLOCK | 4096 FOR 1 BLOCK | | | | | | | | | | | | |
|-------|------------------|--|-----------|--|------------|--|------------|--|------------|-----------|-------|------------|--|
| 0 | UID1 8B | | KEY1 16B | | UID2 8B | | KEY2 16B | | | UID171 8B | | KEY171 8B | |
| 1 | KEY171 8B | | UID172 8B | | KEY172 16B | | UID173 8B | | KEY173 16B | | | UID342 8B | |
| 2 | KEY342 16B | | | | UID343 8B | | KEY343 16B | | | UID512 8B | | KEY512 16B | |
| ... | | | | | | | | | | | | | |
| 64 | | | | | | | | | | | | | |

reader. The operation unit decrypts the received data and verifies RndA. If RndA is successfully verified, the operation unit sends a signal to the lock control module. The lock control module opens the door for a set period. If RndA is not successfully verified or error occurs during the authentication, the operation unit sends a signal to the card reader, which announces an authentication error. The card reader notifies the user authentication error through a sound or a light signal. All of these actions are also stored in the log file in the local and central database (or on memory card of each module).

The communication via RS485 bus is provided through a structured chain of messages ("packet") that are sent among devices. The size of this "packet" is variable. Each byte in the "packet" is significant. The zero position marks the beginning of the "packet" and it is equal to the value 0xFE. The first position is the address of the device, from which the "packet" is sent. The second position is the address of the device to which the "packet" is sent. The third position is the command to be executed (for example, lock opening, lighting up LEDs, loading cards, etc.). The fourth position is the length of the sent data and from the fifth position up are the data itself. A sample structure of the "packet" is shown in the Table 1 below.

Description of each field:

PACKET_START – beginning packet, value is 0xFE

THIS_BOARD_ID – address of the device from which the packet is sent

BOARD_ID – address of the device, to which the packet is sent

CMD – command to be executed by a device that the packet received

LEN – data length

DATA – data

The server accesses to the central database. Data from this database is transferred into the operation unit (local database) using TCP/IP connections. Data transfer from the operation unit to the central database (event record) is done using "HTTP post requests".

Most of the data from the central database uses only server part to better management of the cards and assign permissions. The operation unit uses only connections required for the mapping card readers with lock control modules and authorization access card to zones.

Local database is stored in the operation unit in SPI FLASH memory and RAM. This is not a real database, but only a representation of the data in memory. 8-byte UID card number and 16-byte encryption key are stored in a sequential manner in the memory. The data is entered into the FLASH memory sequentially by ID from the database. It means, that if the card is stored in a central database at the third position and its ID is 3, it will be entered in block 0 to position 3. In this way, all data transferred from the central database to the local database. It avoids overwriting of data in the FLASH memory and speeding up the search process of the UID card and assigning the encryption key to the UID. The next step is the allocation of the information as to whether that card is allowed or not. This RAM memory is used in which the field is created. In RAM memory is created field, which is used just for this. Index of FLASH memory is the position of the field in RAM memory and in this position it is entered as the value of the card is enabled or not. The data stored in RAM memory maps, which card can control which channel to lock (zone). A method of representation of this data also uses field size of 512 elements. Each element occupies 32 bytes of the memory. The ID of the element (position) represents the ID of the card and the individual bits represent the channel of the lock, which it can control. The mapping of the card bookworms and lock control modules are similar to the previous method, unless it is a field size of 32 elements. The size of each element is 4 bytes. The figure (Table 2) shows how is the UID and the encryption key stored in FLASH memory.

5 Conclusion

In this paper we describe the functionality of the generated authentication system based on contactless

technology with encryption. This system is based on quality analysis and the detailed design in fully functional and will be further extended. While implementing we had to deal with many problems that we managed to eliminate. As a result the system is reliable, easy to use, small and inexpensive. In the future, we would like to expand the whole control system with a new functionality controlled via mobile applications and which will also support other types of contactless cards.

Acknowledgement: This work was partially supported by the project „Rozvoj výskumnej infraštruktúry STU, projekt č. 003STU-2-3/2016“ by the Ministry of Education, Science, Research and Sport of the Slovak Republic.

References

- [1] Understanding the Requirements of ISO/IEC 14443, [cit. 28.12.2016].
- [2] <http://www.atmel.com/images/doc2056.pdf>
- [3] W. Rankl, W. Effing. Smart Card Handbook, 3rd edition, New York, John Wiley & Sons, 2003.
- [4] W. Rankl, W. Smart Card Applications: Design Models for using and programming smart cards, New York, John Wiley & Sons, 2007.
- [5] About MIFARE, [cit. 28.12.2016].
- [6] <http://www.mifare.net/en/aboutmifare/>
- [7] Active Load Modulation, [cit.28.12.2016].
- [8] <http://rfid-handbook.de/about-rfid/active-load/modulation.html>
- [9] F. D. Garcia, P. Van Rossum, and R. W. Schreur, “Wirelessly Pickpocketing a Mifare Classic Card,” 2009.
- [10] M. Henzl and P. Hanacek, “Modeling of Contactless Smart Card Protocols and Automated Vulnerability Finding,” 2013.

Estimation of User's Web HCI Literacy

Monika SANYOVÁ*

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

Abstract. User characteristics can influence usability evaluation of a user interface. User research is usually done with approaches such as a questionnaire or interview. Challenge is to replace these time consuming subjective approaches with automatic methods based on processing user's activities. We propose a method for automatic estimation of user's web human computer interaction (HCI) literacy based on his/her interaction with components of user interface. We design an experimental interface and a set of tasks to build a classification model. Since baseline is missing we also design a validation questionnaire for user's web HCI literacy estimation on sample of 117 respondents.

1 Introduction

Nowadays testing usability of a graphical user interface is very common. For usability studies it is useful to know about the user the most relevant information, for instance the level of his HCI literacy estimated due to interaction with the user interface. User interface is supposed to be intuitive and easy to use. We can estimate the level of intuitiveness and learnability by testing the quality of graphical user interface. Good pattern in test participants selection is to choose participants with different computer literacy. If this information is obtained by the participant's interaction with the computer, we can talk about the HCI literacy.

Common user research methods are still questionnaires. Participant's subjective computer literacy estimation is not reliable enough. It is thus necessary to achieve automatic user's HCI literacy estimation. In the work [4] a method for user's computer skills level estimation with use of mouse movements analysis is proposed. We assume, user's gaze analysis can retain more complex and detailed information about the user's interaction, especially when the user is not using computer mouse in carrying out tasks.

Main goal of our work is to design an automatic

method of user's HCI literacy estimation using eye tracking during his interaction with web graphical user interface. In this paper we describe how our method can be used and how we are going to validate it.

2 HCI literacy

Graphical user interface is a communication gate between the user and the system [6]. User communication with the system is reaching his goal by interaction with interface components. System processes these input informations from user and subsequently gives the user response by changing interface components. In context of our work Human Computer Interaction is using web interface components to communicate with the system.

We defined the concept of HCI (Human Computer Interaction) literacy as computer literacy estimation based on interaction with a graphical user interface or a unit of information and communication technologies. As was mentioned in [8] using of information and communication technologies unit requires domain and interaction knowledge. Computer literacy is a basic knowledge of computer interaction required for successful computer usage. In the context of our work this definition expresses the sense of the HCI literacy.

The HCI literate user is familiar with the interface components and aware of their functionality and visual representation, he knows how to use them to interact with the system and he has correct expectations about the response from the system after his interaction with certain components. High level HCI literate user interacts with intuitive user interface more fluently than user with low level of HCI literacy. HCI literacy doesn't depend on the user's domain and terminology knowledge nor on his work with the content.

Theoretically level of the user's HCI literacy can be measured as cognitive distance between his thoughts and intentions about reaching a goal using

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

graphical user interface and its real representation in the system. If we had provided with a perfect, intuitive graphical user interface, level of user's HCI literacy would depend on how highly would his thoughts and intentions about reaching the goal match the reality.

Our work is focused on the web HCI literacy estimated based on user's interaction with web graphical user interface.

3 Related work

In our research area there are quite few relevant works. In some works terms such as computer literacy, computer skills, computer knowledge, computer expertise and formal literacy are discussed [3, 4, 9]. However definitions of these terms are not consistent across these aforementioned papers and term HCI literacy was not mentioned in any analysed work. Two method types of user's computer skills estimation are known - questionnaires and automatic estimation.

There are lots of questionnaires which provides us information about user's computer anxiety¹, attitude towards computer, level of his confidence while using computer or subjective rating of his computer skills^{2,3,4}. We assume that subjective user's opinion is not enough reliable. Other questionnaires like Computer Expertise (CE) questionnaire [1] or Computer Proficiency Questionnaire (CPQ) [2] contain questions for example about computer terminology, files management, email or some problems during computer usage. These and many other questionnaires^{5,6,7} give us information about user's theoretical computer skills. Questionnaire CLS (Computer Literacy Scale) [7] determinates user's knowledge of common symbols and elements used in interfaces of the information and communication technologies. It also determinates diversity of user's computer usage. CLS is the most relevant standardised questionnaire for our work. Unfortunately, it was especially designed for older adults, which means, it is easy for most computer users.

The second type are methods, which can estimate level of user's computer skills automatically with use of a machine learning. In work [4] task dependent and task independent classification methods of user's computer literacy level were designed. They classify users into two groups - novice and skilled users, based

on high frequency events, such as mouse movement during task execution in desktop application. The work focused on estimation of computer skills into two levels (novice and skilled user) based on multiply repeating execution of the same tasks in desktop graphical editor.

4 Graphical user interface components

In our work graphical user interface was analysed and graphical components were identified [5,6]. We identified 45 primary components from which we derived 92 secondary components. For better understanding of different interactions with every component, we divided them into three logical groups:

- Navigational components
- Control components (input control and filtering components)
- Informational components (web content and system response components)

As we mentioned earlier, HCI literacy doesn't depend on user's domain and terminology knowledge nor on his work with content, therefore we have prior interest in navigational and control components. Aforementioned division is not sufficient for a purpose of our work, it is thus we divided components based on the type of interaction with them into following groups:

- Clickable components
- Components, which we can fill in (writeable components)
- "Drag and drop" components
- Components, which we can't interact with (containers, informational components)

Some components matches more than one group, for example spinner, which is clickable and also writeable component.

5 Automatic method for user's HCI literacy estimation

In this paper we describe method for user's HCI literacy estimation we designed. We propose a method which estimates the user's knowledge of many components, especially common components typical for web user interface. Result of the method provides us

¹ <http://www.yorku.ca/rokada/psyctest/companx.pdf>

² <http://www.propofcs.com/quiz-school/story.php?title=computer-skills-assessment-questionnaire>

³ <https://www.jmir.org/article/download/SuppFile/2036/3271>

⁴ <http://documentslide.com/download/link/basiccomputerskillssurveymarch2010>

⁵ <https://www1.dcccd.edu/catalog/ss/transfer/compskillpracttest2014.pdf>

⁶ <https://www.surveymonkey.com/r/?sm=cUIQoj%2Bhe82Ia%20BalUfBCNw%3D%3D>

⁷ http://www.halfbakedsoftware.com/quandary/version_2/examples/computacy.htm

information about user's HCI literacy.

Test participant executes proposed tasks in experimental graphical user interface - e-shop prototype we defined. During tasks execution participant's interaction with the web user interface is monitored. There are several metrics from participant's logged gaze or mouse movements which we consider useful for the HCI literacy estimation. Metrics giving us an information about his HCI literacy are extracted from logged data and used for participant model creation, which is represented as sequence of extracted metrics.

For automatic user's HCI literacy estimation we use existing machine learning method, which uses classification user model as an entry. The method classifies the user into one of five groups which represents level of his web HCI literacy based on his interaction with interface components during the tasks execution.

According to the method description above, our method seems like task and user interface dependent. We want the method to be reusable and applicable also for other user interfaces during different tasks execution. It is thus we designed a methodology for web HCI literacy estimation in which we describe how to use our the method in other graphical user interfaces.

There are two possible ways how to use our method, first way is to use it in usability testing as a pretest in which participant executes our tasks in our experimental user interface. Other way is to construct the classification model by applying methodology we created and use our method in existing user interface which usability is tested. That provides us possibility to do user research and usability testing in the same time, which is time saving.

5.1 Tasks execution

14 tasks divided into five levels according to their difficulty were defined. Task difficulty is determined based on interface components difficulty used for the task completion. Tasks execution requires participant's work with 53 different types of components, we monitored his interaction with 31 components. Tasks for HCI literacy level 1 and 2 are constructed as a sequence of steps (each step is lower level action) and tasks for higher HCI literacy levels are brief goal definitions without help comments.

Test participant carries out how many tasks are required depending on research, for example if in the usability study is reasonable to test only people with their HCI literacy level at least three, but it doesn't matter exactly what level of HCI literacy participant has, participants could be chosen only based on tasks from first three levels.

Tasks are executed by participant in experimental graphical user interface we defined. Each participant interacts with experimental interface for the first time,

so he has no interface knowledge. During tasks execution we consider participant knowledge of components functionality and the type of interaction with it, which are not domain specific. The goal is to make our method for web HCI literacy estimation domain independent because HCI literacy doesn't depend on the user's domain and terminology knowledge. Every domain specific terms required for tasks accomplishment are used and explained in the tasks definition, for example task to display all laptops is defined as *Display all products from category "laptops"*.

5.2 User interaction monitoring

Each task is the sequence of actions consist of finding and recognizing the goal component and of interaction with it. During participant's tasks execution his interaction with user interface components is monitored. Characteristics like his gaze, mouse clicks, mouse movements and keyboard input are logged. Domain independent metrics suitable for HCI literacy estimation we identified are extracted from the logged data. Eye tracking and duration metrics like component findability or recognizability are extracted from each action. Other mouse and keyboard metrics, for example number of clicks and writing velocity are extracted only for clickable or writeable component interaction.

5.3 Classification

After tasks execution and metrics extraction we can represent the test participant as a sequence of metrics - the user model. Classification model will be trained on user models labeled according to user's level of HCI literacy estimated based on our questionnaire result. We will experiment with some existing classifications methods for classification model creation. User model is used as an entry for classification the user info one of five levels of HCI literacy.

6 Questionnaire and method validation

In our research field, there is no relevant method for HCI literacy estimation which we can compare our results with. For results validation of our method we need baseline method which provide verified HCI literacy estimation. Therefore we designed a questionnaire with 42 questions determining user's knowledge of various web user interface components. Questions are focused especially on user's knowledge of these components, which are used in tasks we defined in our method. The aim was designed a questionnaire corresponding with these tasks as much as possible.

The questionnaire was filled by 117 respondents of which 41 were respondents from IT field. Questionnaire was evaluated manually and in this paper we

present following promising results. Figure 1 shows abundance histogram of a success rate of all 117 respondents. Questions were weighted according to their difficulty - question success rate. Respondent success rate was calculated as a sum of correct answered weighted questions. In the figure we can see that lots of respondents gained very high success rate.

Figure 2 shows only results of respondents except IT field. There is big decrease of high success rates, exactly as we expected. There results indicate the normal Gaussian distribution⁸, which is often used to represent real-valued random variables whose distributions are not known. We assume results of our classification method will correlate with results from our questionnaire, what will validate the method.

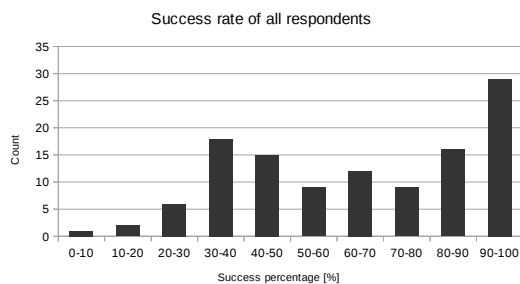


Figure 1. Questionnaire results of all 117 respondents.

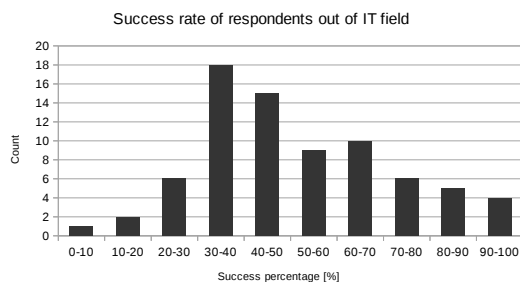


Figure 2. Questionnaire results of 76 respondents out of IT field.

7 Conclusion

We propose the automatic method for user's HCI literacy estimation which is applicable for our experimental user interface and also for other graphical user interfaces thanks our methodology we design. We also propose and validate the questionnaire for user's HCI literacy estimation which was designed as a baseline for our method results comparison but it can be

also used as a standalone approach for user's HCI literacy estimation. We assume results of our method will correlate with results of this questionnaire.

In further work we plan to validate our tasks and user interface we defined. We will train our classification model with data from quantitative experiments and finally we will evaluate our method by comparing its results with results with questionnaire results.

Acknowledgement: This contribution was created with kind support of ČSOB Foundation and is partial result of the project University Science Park of STU Bratislava, ITMS 26240220084, co-funded by the ERDF.

References

- [1] Arning, K., Ziefle, M.: Development and validation of a computer expertise questionnaire for older adults. *Behaviour & information technology*, 2008, vol. 27, no. 4, pp. 325–329.
- [2] Bradlow, E.T., Hoch, S.J., Hutchinson, J.W.: An assessment of basic computer proficiency among active internet users: Test construction, calibration, antecedents and consequences. *Journal of Educational and Behavioral Statistics*, 2002, vol. 27, no. 3, pp. 237–253.
- [3] Di Maio, L.R., Wagner Freitas Cavalcante, N., Leal Ferreira, S.B., dos Anjos Rosa, J.L., Alves, A.S.: Chap. High-Literate and Low-Literate User Interaction: A Comparative Study Using Eyetracking in an Emergent Economy. In: *Universal Access in Human-Computer Interaction. Design for All and Accessibility Practice: 8th Int. Conf., UAHCI 2014, Greece, 2014, Proc., Part IV*. Springer International Publishing, 2014, pp. 89–100.
- [4] Ghazarian, A., Noorhosseini, S.M.: Automatic detection of users' skill levels using high-frequency user interface events. *User Modeling and User-Adapted Interaction*, 2010, vol. 20, no. 2, pp. 109–146.
- [5] Johnson, J.: *GUI Bloopers 2.0: Common User Interface Design Don'Ts and Dos*. 2nd edn. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2007.
- [6] McKay, E.N.: *UI is communication: How to design intuitive, user centered interfaces by focusing on effective communication*. Newnes, 2013.
- [7] Sengpiel, M., Dittberner, D.: The computer literacy scale (CLS) for older adults-development and validation. In: *Mensch & Computer*, 2008, pp. 7–16.
- [8] Sengpiel, M., Jochems, N.: Validation of the Computer Literacy Scale (CLS). In: *Human Aspects of IT for the Aged Population. Design for Aging*. Springer, 2015, pp. 365–375.
- [9] Van Deursen, A.J., van Dijk, J.A.: Modeling traditional literacy, Internet skills and Internet usage: An empirical study. *Interacting with computers*, 2014, p. iwu027.

⁸ https://en.wikipedia.org/wiki/Normal_distribution

Knowledge Harvesting from Slovak Web

Jakub ĎAĎO*

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

Abstract. Majority of data on the Web is stored in unstructured form. In order for machines to utilize the knowledge from the Web we need to transform the data to structured form. Knowledge bases, like DBpedia or Freebase, are being built mainly from English sources, a lot of data from the Slovak Web is missing. In this paper we introduce a knowledge harvesting method for automatic extraction of structured information from Slovak Web sources. It utilizes machine learning, associative rules mining and coreference resolution. We achieved a combined precision of 60 % by evaluating on 8500 Wikipedia articles, with possibilities for further improvements.

1 Introduction

Many data sources on the Web contain data in unstructured form. The idea of the Semantic Web proposes that data should be represented in a structured way so that they can be used by autonomous applications for knowledge discovery and reasoning. Such structured data is represented via RDF standard as tuples composed of *subject*, *predicate* and *object* [1].

The largest existing Slovak knowledge base – DBpedia – was built by extracting information from Wikipedia’s infoboxes (structured tables). DBpedia contains approximately 2 300 000 tuples and this repository is still far from completion. For comparison, Freebase is another knowledge base containing more than 40 million facts.

The domain of Named Entity Recognition (NER) is the dominant part of the knowledge harvesting process [2]. It is being used by the most prominent knowledge bases like YAGO [2], DBpedia [5], FreeBase [1], Knowledge Vault [8]. These systems also enrich the resulting data with relations as opposed to pure NER systems, which only harvest entities (subject, object) from the texts.

The majority of the existing approaches for knowledge harvesting use mainly supervised machine learning, with 3 main steps:

1. Pre-processing
2. Learning
3. Transforming

The pre-processing step is composed of part-of-speech tagging, lemmatization, stemming, sentence segmentation or even sentence dependency parsing [2,3,4,6]. Next, learning relies on already pre-processed data. For example, authors of [8] use this data as input to their automatic extractors that run on the text and yield the results as entities and relations between them. The outcome of the learning step is new information represented as an RDF triple.

2 Related work

We can classify knowledge bases to different types based on the way they are constructed. WordNet [3], which is lexical synonymic dictionary, was created *manually* by domain experts. The Slovak version¹ contains approximately 1/4 of the information of the English version. WordNet contains structure (*class-subclass*), that creates fundamentals for YAGO and DBpedia, which use WordNet as dictionary for knowledge deducing.

DBpedia and YAGO belong to *automatic structured approach* type. They use Wikipedia infoboxes (structured templates) as the main source of data. They use pattern matching techniques simultaneously with regular expressions. YAGO achieved great results as it managed to harvest 100 relations (e.g. *locatedIn*, *bornInYear*) with accuracy of 95% [2]. The last and the most efficient type of knowledge harvesting is *automatic unstructured approach* (e.g. Probase, Knowledge Vault).

Probase [6] is a knowledge base which uses pattern matching method for automatic knowledge harvesting, namely Hearst Patterns. The main idea is to iterate over text and improve the next iteration based on the patterns used in the previous one. Thanks to using Hearst patterns it is easy to create a custom taxonomy (class - subclass), which is the main difference between this and DBpedia. This method

* Master degree study programme in field: Information Systems

Supervisor: Michal Holub, Faculty of Informatics and Information Technologies STU in Bratislava

¹ <http://korpus.juls.savba.sk/WordNet.html>

achieved 92 % accuracy in knowledge harvesting and at the end of day, they managed to extract 85 000 instances, which were included in 40 concepts (actor, book). One of the drawbacks in Probase is the necessity for manually defined patterns.

Knowledge Vault is another method for automatic knowledge extraction from unstructured text. Unlike Probase, this method does not rely on manually created patterns, but it uses machine learning process, specifically classification, to extract knowledge from the text. During experiments, the authors managed to extract 4500 relations and besides that, they extracted over 270 million entities with the precision of 90 %. They use 3 types of extractors (Path ranking algorithm, structured HTML extractor and sentence extractor) and combine them together. These extractors are trained on Freebase. The drawback is that they can learn and extract only relations from a specific knowledge base.

3 Automatic knowledge harvesting from Slovak Web

We propose a method for automatic knowledge extraction from unstructured Slovak text. The results are represented in RDF triples. We focus mainly on the content of Slovak Wikipedia, but our method can work with different languages as well. The method contains 3 major components:

- Mapper – maps tokenized sentences to an existing knowledge base.
- Extractors – extract triples from the text. Each triple is assigned a probability indicating the certainty of extracted knowledge.
- Knowledge evaluator – secures verification of extracted triples with existing knowledge in knowledge base.

Each of these steps relies on Slovak version of DBpedia. In the context of other processes, which are necessary for knowledge harvesting, our method follows these steps:

1. Text pre-processing
2. Text to DBpedia mapping
3. Relation learning and extraction
4. Relation rating, RDF generation

3.1 Text pre-processing and DBpedia mapping

For each text, we need to do a pre-processing phase. The steps realized during this phase are similar to other methods. The only difference is in using the tools² intended for Slovak texts. After the pre-processing part is finished, we continue by mapping the pre-processed sentences from text to DBpedia in order to connect the meaning of the sentence with a

relation stored in DBpedia. We take a sentence and its tokens, as depicted in Figure 2. If we match tokens with subject and object on DBpedia, we assign the predicate from DBpedia to this sentence afterwards. The process of mapping is executed just once, after all of the sentences are pre-processed.

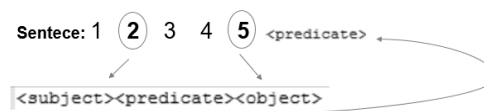


Figure 1. Determination of sentence relation with DBpedia predicate.

3.2 Relation learning and extracting

Our method for automatic extraction of structured information from texts contains 2 types of extractors:

- Extractor for 2 entities
- Extractor for coreferences in text

In the real world, the sentences are not always the same, but they denote the same meaning. Based on this we made the following assumption (A1):

A1: “If one sentence contains similar entities to other sentence, those sentences may denote the same meaning”.

We represent each sentence by a vector of features. We recognize 12 features, which depend on rules and mapped tokens on DBpedia. Features are created by the following rules:

1. Words between 2 mapped entities
2. Number of words between 2 mapped entities
3. N-grams before first mapped entity
4. N-grams before second mapped entity
5. Part-of-speech tags for each word
6. Labeled first mapped entity in sentence

We use the same length of {0, 1, 2} - grams as [4] for the rules 4 and 5. Given an example of sentence (S1), its corresponding feature vector is in Table 1. Bold words in S1 are those, which are mapped to DBpedia, where the relation from DBpedia is in parentheses.

S1: “Important philosopher **Aristoteles** was born in **Greece** in 382 BC”. (*bornIn*)

All vectors are passed as input to the classifier. Before the classification step, all string values are converted to numbers, because logistic regression can work just with nominal values. We use classifier to secure good relation denoting, as many words in the sentences may be similar.

² <http://text.fiit.stuba.sk/>

Table 1. Vector of features for S1 (V-verb; P-preposition; N-noun; A-adjective).

| | |
|----------------------------|-----------------------|
| Set of words S_{middle} | was born in |
| Pos tags S_{middle} | V V P |
| 1gram S_{first} | philosopher |
| Pos tags 1gram S_{first} | N |
| 2gram S_{first} | Important philosopher |
| Pos tags 2gram S_{first} | A N |
| 1gram slov S_{last} | in |
| Pos tags 1gram S_{last} | P |
| 2gram S_{last} | in 382 |
| Pos tags 2gram S_{last} | P N |
| Word count S_{middle} | 3 |
| First entity | 1 |

In process of extraction itself, we use NER system³ for denoting entities in sentences (like locations, persons...), followed by creating vector of features as we described in this section.

Unstructured text is specific in the way how it handles coreferences. Coreferences occur in sentence, when we refer to the same person or thing by the other Noun. For example, we have two sentences, which refer to same person: *Frank is entrepreneur. He also studies*. In our case if we don't consider coreferences we lose a lot of knowledge from text, since we can't map Noun *He* on DBpedia. We approached this problem with the second extractor, based on Wikipedia topic structure.

Wikipedia articles are specific in a way that there is just information per topic. We use name of the topic as referrer to every coreference in text. Second extractor is applied for each sentence, where it finds a match with DBpedia object (DBpedia subject is added from Wikipedia topic name). To better identify relation in sentence mapped just by one entity we use part of feature vector from previous extractor (E1), namely we use n-grams and pos-tagging of words next to mapped entity. We also take the relation mapped to this sentence from the E1. Based on these facts we create new feature vector for the second extractor E2.

E1: “**Wien** has many rivers flowing through, one of them is **the Danube**.” (*hasRiver*)

T1: “**Bratislava** is capital city of Slovakia.”

T2: “**It** has access to 10 km flowing river Danube.”

We have text T1 and text T2 where in text T2 *id* reference to Bratislava. We take a token – Danube from T2 and look for the occurrence in sentences prepared for E1. For every occurrence, we take features (n-grams, pos-tagging) with relation. New feature vector for E2 is created similarly to E1

afterwards. In case of, the text T1 and the text T2, the final triple will look like: (*Bratislava, hasRiver, Danube*).

3.3 Knowledge evaluator

We use knowledge evaluator right before inserting new triples to the knowledge base. We compare triples based on their predicate, in other words we try to determine if predicate of extracted triple may exist with given subject and object. We use 3 rules to determine the existence of triple, where $T(s,p,o)$ is new extracted triple with subject, predicate and object and $R(s, p)$ refers to relation connected to subject:

1. $\exists T(s) \in \text{DBpedia} \wedge R(s, isA) \in \text{DBpedia}$
2. $\exists T(s) \in \text{DBpedia} \wedge R(s, isA) \notin \text{DBpedia}$
3. $\exists T(s) \in \text{DBpedia} = \emptyset$

Firstly, we check if subject exists in DBpedia. If so, then we check whether any *isA* relation exists with this subject. Afterwards, we aggregate all the relations, which are connected to subject into one set. For example, for cyclist – Peter Sagan, DBpedia may contain relation *isA* denoted as *isPerson*. Later on, if we extract new entity from another text with some relation, we check, if these relations are included with all relations connected to person.

For the rule 1 we use associative rules. If some entity is of type person and has relation X, it is probable that another subject of type person will have the same relation.

For the rule 2 we use Jaccard similarity, where we miss the relation *isA*. We try to determine the most similar subject, which has the relation *isA*. We take all the relations of extracted subject and compare them to other subjects in the whole DBpedia. We sort the subjects according to the similarity and pick out the most similar one with existing relation *isA* to our extracted subject.

For the rule 3 we discard new entity, because the subject of that entity is not stored in DBpedia.

4 Evaluation

We conducted the evaluation on the set of 8500 articles from Wikipedia. The articles were from various genres about people, locations, etc. Up to this moment the evaluation was conducted with extractor based on two entities. 8500 Wikipedia articles contained approximately 270 000 sentences. We selected 2000 of them, which contained <subject> <object> values. The average length of a sentence was 15 words. The following experiments were conducted on training (80 %) and testing (20 %) sets of all 2000 mapped relations.

³<http://polyglot.readthedocs.io/en/latest/NamedEntityRecognition.html>

Firstly, we compared SVM and multiclass logistic regression, which are the two most used methods in the domain of knowledge harvesting. We found out that the time for training based on logistic regression increases almost exponentially with the increasing number of relations in the training set. As opposed to SVM, which was trained under 10 seconds in each experiment. The precision differed slightly (2 %) in favor of the logistic regression. That was tiny difference, therefore we decided to use SVM in other experiments as it performs much faster.

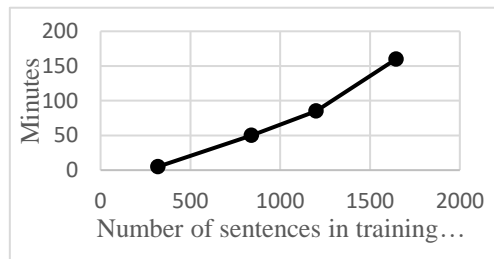


Figure 2. Logistic regression performance.

In the following experiment, we looked at the precision, meaning how efficiently we can determine the correct relation to the sentence. We had 180 specific relations in our training set. Some of them were more significant than others. E.g. relation *location* had 600 unique instances, compared to 2 instances of *hasHDP*. This caused a problem with classifier overfitting. The classifier tends to classify objects more likely with relations from larger group. After trying various values for the classifier parameters we came to conclusion that it is optimal to train classifier, where the difference in the groups' multiplicity is around 30 %. E.g. if relation-A has 150 relations, then optimal minimal count of other relations should be minimal of 20. We also took in mind that we want to avoid classification for scarce relations. Almost 155 relations in our training set had less than 5 occurrences. So, we ended up with 13 final relations. We achieved precision of around 60 % in this setup as opposed to normal setup (without targeting the overfitting problem) where the precision was only 44 %. Table 3 shows examples of precision for couple of relations.

Table 2. Precision distributed by relations.

| Relation | Precision |
|------------|-----------|
| neighbors | 50% |
| city | 50% |
| government | 75% |
| area | 88% |

5 Conclusions

In this paper, we proposed a method for automatic knowledge harvesting from Slovak texts. We must

highlight that to the best of our knowledge there is no similar solution for knowledge harvesting specifically from Slovak texts. Methods, which work for English language texts are all language dependent, therefore they are not directly transferrable to our problem domain.

We have made several experiments with the proposed method. The results are very promising. In the future work we would like to improve them by evaluating the method including other two modules, which are the knowledge evaluator and coreference resolution method.

Compared to other existing methods our is unique as it uses the contents of Wikipedia in order to solve the coreference resolution problem. We also evaluate the extracted entities with associative rules and Jaccard Similarity.

Acknowledgement: This work was partially supported by the Science and Technology Assistance Agency under the contract No. VG 1/0646/15.

References

- [1] Son, B.N., Hoang, P.H.: New Approach for Extracting Information on Freebase Based on User's Preferences. In *Proc. of 16th Int. Conf. on Information Integration and Web-based Applications & Services*. ACM Press (2014), pp. 54-58.
- [2] Fabian M., Suchanek, G. K., Weikum, G.: Yago: A Core of Semantic Knowledge. In *Proc. of the 16th Int. Conf. on World Wide Web*. ACM Press (2007), pp. 697-706.
- [3] Heredero, G.G., Bandyopadhyay, S., Laha, A.: Extraction of Contextual Information from Medical Case Research Report Using WordNet. In *Proc. of the Fourth Annual ACM Bangalore Conference*. ACM Press (2011), pp. 1-8.
- [4] Mintz, M., Bills, S., Snow, R., Jurafsky, D.: Distant Supervision for Relation Extraction Without Labeled Data. In *Proc. of the Joint Conf. of the 47th Annual Meeting: Vol. 2*. Association for Computational Linguistics (2009), pp. 1003-1011.
- [5] Auer, S., Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak, R., Ives, Z.: DBpedia: A nucleus for a Web of Open Data. LNCS 4825, Springer (2007), pp. 722-735.
- [6] Wu, W., Li, H., Wang, H., Zhu, K.Q.: Probbase: A Probabilistic Taxonomy for Text Understanding. In *Proc. of 2012 ACM SIGMOD Int. Conf. on Management of Data*. ACM Press (2012), pp. 481-492.
- [7] Dong, X., Gabrilovich, E., Heitz, G., Horn, W., Lao, N., Murphy, K., Strohmman, T., Sun, S., Zhang, W.: Knowledge Vault: A Web-scale Approach to Probabilistic Knowledge Fusion. In *Proc. of the 20th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining*. ACM Press (2014), pp. 601-610.

Segmentation of Multiple Sclerosis Lesions in Radiological Images Using Methods of Computer Vision

Juraj UHRIN*

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

Abstract. Methods of computer vision and image processing have growing importance in the medicine for the processing of radiological images. Manual segmentation by neurological experts is very time consuming. Hence, an automatic multiple sclerosis lesions segmentation is developed as an alternative solution to the manual segmentation. In this paper, we analyse the methods of image processing for medical evaluation of the patient's condition. We are focusing on the automatic segmentation of multiple sclerosis lesions from radiological images obtained by magnetic resonance (MR). We propose segmentation methods that analyse two-dimensional FLAIR sequence images. The segmented regions provide information about the location of lesions in the brain. Proposed methods have been implemented and evaluated using real medical data. Preliminary results are promising.

1 Introduction

In the processing of radiological images in medicine there are methods of computer vision and image processing whose importance is rapidly increasing. Manual segmentation of multiple sclerosis lesions by neurological experts is time consuming and subjective. A lot of publications appeared in recent years, which support the idea of automatic multiple sclerosis lesions segmentation as an alternative solution to manual segmentation.

Particularly interesting in analysis of two-dimensional images from magnetic resonance imaging (MRI) are methods that use natural properties of the brain, such as symmetry. The use of these features in the analysis of MR images can provide information relevant for assessing the patient's condition.

In this paper we analyse the methods of image processing and we focus on the automatic

segmentation of multiple sclerosis lesions from MR images. We propose segmentation methods that provide information about the location of lesions in the brain. Proposed methods have been implemented in C++ using libraries suitable for the processing of medical data. The achieved results are evaluated using real medical data.

2 Segmentation

One of the first steps of many algorithms in image analysis is the process of segmentation. It is a process of division of the digital image into parts or areas according to certain features and characteristics of individual pixels. Image segmentation is important in a lot of computer vision areas, like filtration of noisy images, processing of images from medical modalities, objects identification on satellite images, automatic monitoring of traffic situation or object detection in general applications of computer vision.

Many published articles analysed different segmentation methods and approaches, yet image segmentation is still considered to be complex and difficult task [9]. A suitable segmentation technique depends on specific problem and its specific features. The level of segmentation and methodology of problem solving is chosen according to these features. Image segmentation may be used for detection and identification of specific objects in different images [4].

The most often used segmentation methods are [7]:

- Edge based methods
- Region based methods
- Threshold based methods
- Methods based on neural networks

None of these techniques is however a general solution for complex segmentation problem. It is

* Master degree study programme in field: Information Systems
Supervisor: Assoc. Professor Vanda Benešová, Institute of Computer Engineering and Applied Informatics, Faculty of Informatics and Information Technologies STU in Bratislava

necessary to choose a suitable technique and adjust it to the specific problem, if we want the best results. To combine more segmentation methods to create one suitable method seems to be the best solution so far.

3 State-of-the-art MS lesion segmentation

Based on the assumption that different brain tissues have different grey-level intensities, the most common feature used for MS lesion segmentation is the voxel intensity [3]. Furthermore, the appearance of the tissue and the lesions depends on the sequence used to capture MR images. For segmenting MS lesions, the FLAIR sequence is better at differentiating between lesions and healthy tissue despite the fact that it introduces deceptive artefacts into the image which complicates the lesion segmentation in the sub-cortical regions. We may distinguish between two basic approaches of MS lesion segmentation: single-channel and multi-channel. Single-channel approaches use only one MR image and multi-channel are those that combine several images from several MR sequences.

Kamber et al. [5] developed a brain tissue model for segmentation of MR images. The offered model was three-dimensional, voxel based and provided a priori probabilities for white matter (WM), grey matter (GM) and cerebrospinal fluid (CSF). They suggested restricting search area for MS lesions to white matter only.

Udupa et al., suggested a semi-automatic segmentation method [8] utilizing fuzzy objects and fuzzy connected sets. In this method, sample points for white matter, grey matter, and CSF were selected by user. In the next step, the void pixels were offered to the user, for selection as MS lesions. They showed that their methodology was highly reliable and consistent, with a variation coefficient of 0.9% for volume.

Admasu et al. [1] suggested an improvement to the fuzzy connected sets approach by using neural networks. User was no longer asked to make decisions about void pixels, because neural network did this part for him. The evaluation was carried out by comparing the quantity of lesions through the proposed method and the manual segmentation by an expert. Compared were also false positive and false negative percentages. The results showed 90% correlation with the segmentation by the specialist.

3.1 MS Lesion Segmentation Challenge 2008

At MICCAI 2008, a workshop was organized with the objective of comparing different segmentation algorithms in the form of a competition. One of the segmentation challenges was the segmentation of multiple sclerosis lesions from MR data. The data has been acquired at the Children's Hospital in Boston (CHB) and the University of North Carolina (UNC).

The dataset was divided into two sets: training and testing data.

Anbeek et al. [2] proposed to segment white matter signal abnormalities by using an intensity based k-nearest neighbors method with spatial prior and a fuzzy inference system. This method provides a voxel-wise probabilistic classification of MS lesions and was rated as one of the top-ranked methods of this challenge. Evaluation results can be seen in Table 1.

Table 1. Evaluation results of segmentations from Anbeeks work.

| Datasets | UNC Rater | | CHB Rater | |
|-------------|------------------------|-------------------------|------------------------|-------------------------|
| | True pos. [%] score | False pos. [%] score | True pos. [%] score | False pos. [%] score |
| CHB Case5 | 40.7 | 98.1 | 78.3 | 79.5 |
| UNC Case2 | 58.8 | 74.1 | 36.4 | 17.2 |
| All Average | 48.9 | 83.2 | 59.1 | 78.5 |

4 Proposed method

In this paper, we propose a method for multiple sclerosis lesion segmentation from medical images. We decided to use numerous computer vision methods and algorithms to properly identify and locate lesions in MR images. Our method analyse two-dimensional FLAIR sequence images, therefore is preprocessing of raw MR data required. Proposed method follows these six steps:

1. Load two-dimensional MR image to be analysed.
2. Mask is created to remove cranial regions from analysed image.
3. Morphological reconstruction is used to obtain local intensity maxima.
4. FAST corner detector [6] is applied on processed image and keypoints are detected.
5. Each keypoint is checked for symmetry pair and filtered accordingly.
6. Dominant contours are extracted from image using Canny edge detector.
7. Keypoints are mapped to contour positions and overlapped areas are segmented from image.

4.1 Preprocessing

There are two main reasons for preprocessing – transformation raw MR data to two-dimensional

slices and removal of a cranial area from extracted images.

Two-dimensional slices are obtained by using ITK library and are cropped to resolution of 512x512 pixels.

Cranial region on analysed image can have exactly same grey level intensity as lesions that occur in brain tissue. We propose simple method for masking cranial region and extracting region of interest from analysed images. This method creates binary mask from an input image and using erosion adjusts the mask to represent relevant region of interest (see Figure 1). The binary mask obtained from input image is iteratively eroded to achieve suitable circular region of interest that does not contain any cranial regions.

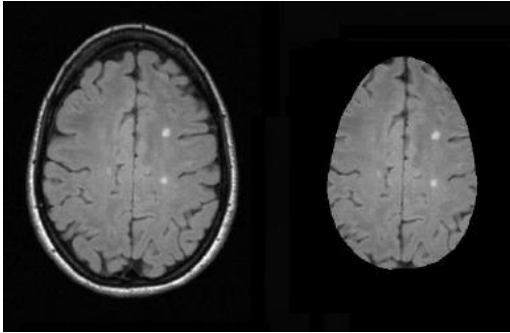


Figure 1. Input image compared with extracted region of interest.

4.2 Lesion segmentation

Our approach to lesion segmentation is to apply morphological reconstruction and highlight areas with local grey level intensity maximum (see Figure 2).

To identify lesions on the processed images we decided to use FAST corner detector to obtain positions of keypoints on analysed image. Keypoints are then located as follows:

1. Pixel p with intensity I_p is selected and it is to be identified as an interest point or not.
2. We define appropriate threshold t that depends on mean intensity value in input image.
3. We consider a circle of 16 pixels around pixel p .
4. Pixel p is identified as keypoint, if there exists a set of 12 contiguous pixels in the circle which are all brighter then $I_p + t$ or all darker then $I_p - t$.

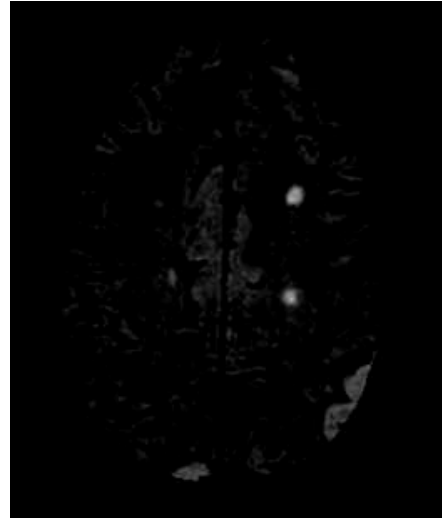


Figure 2. Local intensity maxima obtained by morphological reconstruction.

For better comparison we also propose a blob detection approach for extracting keypoints from analysed images. Many lesions in early stages of multiple sclerosis illness appear on MR scans as circular blobs.

To segment multiple sclerosis lesions we extract image contours and map their positions with keypoint positions detected earlier. Contours that match keypoints are then extracted and declared segmented lesions (see Figure 3).

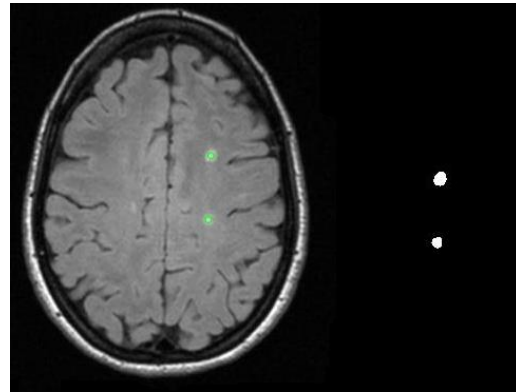


Figure 3. Keypoints detected on analysed MR scan and extracted lesions.

5 Current results

In this paper, we present the results of our segmentation methods. We implemented and

evaluated method that includes morphological reconstruction in combination with FAST corner detection and blob detection method. Symmetry check for each keypoint was not implemented at this stage of the work. Table 2 represents results of these segmentations.

These results are not the final results yet because this evaluation does not include the complete pipeline. We have not implemented symmetry check yet and this will be done in the earliest future. Final results will be provided at the conference.

Table 2. Results of proposed segmentation methods.

| Datasets | Morphological reconstruction | | Blob detection | |
|-----------|------------------------------|-------------------------|------------------------|-------------------------|
| | True pos. [%] score | False pos. [%] score | True pos. [%] score | False pos. [%] score |
| CHB Case5 | 48.87 | 71.01 | 55.14 | 55.47 |
| UNC Case2 | 45.88 | 40.47 | 46.24 | 25.30 |
| Average | 47.38 | 55.74 | 50.69 | 40.39 |

6 Conclusion and future work

To date we proposed complex method for multiple sclerosis lesion segmentation. To better understand specific problems appearing in this area, we analysed currently popular approaches of lesion segmentation and possibilities of processing medical data obtained from MR.

We achieved promising results in lesion segmentation and our estimated true positive rate is close to reference methods we analysed. We believe that implementing keypoint symmetry verification will provide even better results. We are going to continue with evaluation on more MR volumes and complete overall segmentation process.

Acknowledgement: The authors would like to thank for contribution to the Research and Development Operational Programme for the project “University Science Park of STU Bratislava”, ITMS 26240220084, co-funded by the European Regional Development Fund.

References

- [1] Admasu, F., et al. Segmentation of multiple sclerosis lesions from MR brain images using the principles of fuzzy-connectedness and artificial neuron networks. In: *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*. IEEE, 2003. p. II-1081.
- [2] Anbeek, P.; Vinccken, K. L.; Viergever, M. A. Automated MS-lesion segmentation by k-nearest neighbor classification. *MIDAS J*, 2008.
- [3] Cuadra, M., B., et al. Comparison and validation of tissue modelization and statistical classification methods in T1-weighted MR brain images. *IEEE transactions on medical imaging*, 2005, 24.12: 1548-1565.
- [4] Jain, A. K.; Zhong, Y.; Lakshmanan, S. Object matching using deformable templates. *IEEE Transactions on pattern analysis and machine intelligence*, 1996, 18.3: 267-278. Henrio, L.: Asynchronous Object Calculus: Confluence and Determinacy. PhD. thesis, Universite' de Nice-Sophia Antipolis, (2003).
- [5] Kamber, M., et al. Model-based 3-D segmentation of multiple sclerosis lesions in magnetic resonance brain images. *IEEE transactions on medical imaging*, 1995, 14.3: 442-453.
- [6] Rosten, E.; Drummond, T. Machine learning for high-speed corner detection. In: *European conference on computer vision*. Springer Berlin Heidelberg, 2006. p. 430-443.
- [7] Sharma, N., et al. Automated medical image segmentation techniques. *Journal of medical physics*, 2010, 35.1: 3.
- [8] Udupa, J.K., et al. Multiple sclerosis lesion quantification using fuzzy-connectedness principles. *IEEE Transactions on Medical Imaging*, 1997, 16.5: 598-609.
- [9] Varshney, S. S.; Rajpal, N; Purwar, R. Comparative study of image segmentation techniques and object matching using segmentation. In: *Methods and Models in Computer Science, 2009. ICM2CS 2009. Proceeding of International Conference on*. IEEE, 2009. p. 1-6.

Virtual Graffiti Using Smartphone

Martina HALAJOVÁ*

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

Abstract. Nowadays, a set of motion sensors is commonplace in smartphones, for example an accelerometer or a gyroscope, which has a great potential to enrich the development of mobile applications by adding new interactive and interesting features. One of such features may be drawing graphical objects using motion tracking of a user's smartphone. Currently, any implementation of a system for drawing or writing on large flat surfaces requires dedicated equipment, such as a financially expensive interactive whiteboard. Our innovative application, called Virtual Graffiti, unburdens the user from this limitation by using only a smartphone and a projector. A server receives data from the accelerometer of a smartphone, calculates the relative position of the smartphone and performs rasterisation of graphical drawings.

1 Introduction

With use of current technology, we are able to create street art such as graffiti in a virtual environment. The purpose of the application is to provide a way of drawing large-scale graphic objects, without the need of dedicated equipment. The need of the special equipment usually results in additional costs, limiting the availability of the system for a wider spectrum of users.

Given the drawbacks of the special equipment, we propose an application called Virtual Graffiti that uses a set of commonly available technologies, such as a smartphone and its built-in sensors, a computer for calculations and a display device. The main challenge of the system is to maintain the precision of the calculated position.

The paper is structured as follows. Section Architecture describes the architecture of Virtual Graffiti, section Accelerometer discusses the characteristics of the accelerometer sensor on smartphone. Section Calculation of Position describes the usage of a double

integration as the main method for the calculation of smartphone's position. Section Evaluation discusses an experiment designed to verify the calculation used to determine current position of the smartphone. Section Conclusions summarizes the results of this paper.

2 Architecture

The architecture of Virtual Graffiti and connections between its individual parts is represented in Figure 1. The system is divided into two separate parts – client side and server side. The client side is implemented as an Android application, used to capture and send data from the smartphone's sensors, as well as transmit user defined actions. The server side represents a plug-in to the GIMP image editor written in Python language. The server receives data from the Android application and interprets them as graphical objects visible for the user on the display device.

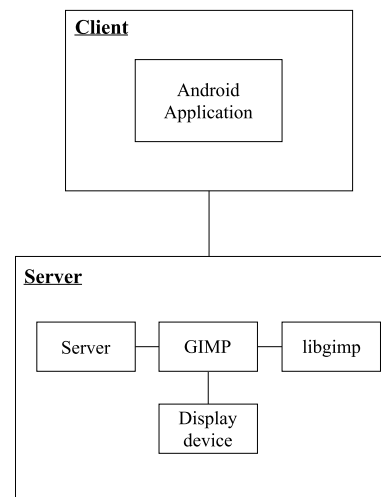


Figure 1. Architecture of Virtual Graffiti.

* Bachelor study programme in field: Informatics

Supervisor: Kamil Burda, Institute of Informatics, Information Systems and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

3 Accelerometer

For the rasterisation of graphical drawings based on the motion of the smartphone, it is necessary to obtain the current position of the smartphone. The most suitable built-in component of the device for this purpose is an accelerometer.

The accelerometer measures the acceleration applied to the smartphone, including the gravitational force [1]. However, the signal received from the accelerometer also contains noise, therefore it must be smoothed, usually by using the moving average [3]. In our case we compute the moving average from 16 samples, while the sampling rate of the Android application is set to 50 Hz.

4 Calculation of Position

After extracting and preprocessing raw data from the smartphone's accelerometer, we use the double integral of the mobile device acceleration over time to determine the relative position of the smartphone [2].

The integral is generally defined as the area below the curve. In this case the curve represents data retrieved from the accelerometer. Because the retrieved data are discrete values unevenly distributed in time as shown in Figure 2, we are using the Trapezoidal rule for the approximating the area.

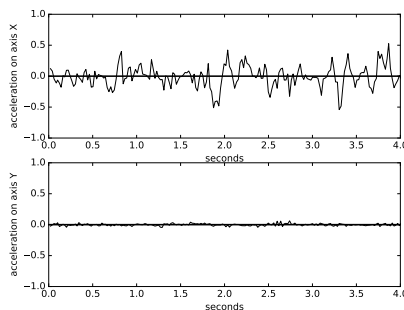


Figure 2. Sampled accelerometer values from a mobile device during slow movement along a horizontal line from left to right.

However, there are several issues associated with the discussed method. Besides the fact that the raw data from accelerometer include the aforementioned noise and the force of gravity (which can be filtered on side of the Android application), there is another issue

regarding the implementation of the double integral calculation.

Because we work with the discrete values and the Trapezoidal rule is only an approximation of the area under the curve, it does not precisely match the real conditions. This error, although small, becomes a significant issue, especially considering the accumulation of the errors. The accumulation error may be avoided by a higher sampling frequency [3].

5 Evaluation

The testing of Virtual Graffiti will be executed in the following scenarios. The tester's task will be to follow a straight line using the mobile device, firstly along the x-axis from left to right, then along the y-axis from bottom to top. Next test will consist of drawing diagonal line from upper-right corner to lower-left corner. Final task will be testing the ability of the application to correctly calculate device's position while drawing a circle. The computed position will be compared to the template shape the tester was supposed to draw.

6 Conclusions

In this paper we proposed an application and a method for the calculation of position and motion of the smartphone by the double integration of data from the smartphone's accelerometer and the errors associated with it. We have designed Virtual Graffiti which takes advantage of this approach and provides users the way to draw the graphic objects by the motion of their smartphones. Future work includes testing the solution on several users to determine the precision of calculated position.

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

References

- [1] Google, Inc.: Sensors Overview. https://developer.android.com/guide/topics/sensors/sensors_overview.html, [Online: accessed February 15th, 2017].
- [2] Neto, P., Pires, J.N., Moreira, A.P.: 3-D position estimation from inertial sensing: Minimizing the error from the process of double integration of accelerations. In: *IECON 2013 - 39th Annual Conference of the IEEE Industrial Electronics Society*, 2013, pp. 4026–4031.
- [3] Seifert, K., Camacho, O.: Implementing positioning algorithms using accelerometers. *Freescale Semiconductor*, 2007, pp. 1–13.

Evaluation of Adaptive Prediction Methods

Tomáš GAŠŠO*

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

Abstract. Information hidden in data is most valuable on their arrival and its value declines over time. Methods, which are able to process data online or incrementally, are used for stream analysis and prediction. To evaluate the methods, which are dealing with stream-specific problems, such as concept drift, limited processing time or storage, we need evaluation methods to verify their abilities. In this paper, we propose an adjustment of measures, which express the quality of a concept drift detector. By adjustment of a confusion matrix's construction for outputs of methods detecting the concept drift, the measures, that were originally designed for categorical data, are also applicable to numerical type of data. This solution is based on idea that we need to know whether the method is able to discover concept drift in a window defined by time or count of instances from data stream. Moreover, we propose a methodology for evaluation of adaptive prediction methods, which we implemented in a tool for validation and comparison of adaptive prediction methods.

1 Introduction

Each device, smartphone, service, network, system, software creates lots of data. These huge amounts of data have typically features as volume, velocity and variety. Such data can be processed offline or online. When they are processed offline, they are already stored. But when we are thinking about huge amounts of data which arrive as a stream, we can employ online methods of data mining.

One of the tasks of data mining is prediction. Methods predicting from data stream have to deal with such problems as changes in data over time (concept drift), restrained memory, decaying accuracy. That's why they should be adaptive and robust. Generic modules of adaptive learning algorithms are memory, learning, loss estimation and change detection. The memory module defines how and for how long algorithm store temporary data. The loss estimation module provides feedback on

prediction error. The change detection detects the concept drift in data over time and gives the feedback to the learning module, which performs prediction.

In online data stream processing, there is a problem with evaluation because of concept drift, time, stream structure, latency, etc. The standard approach to evaluation of adaptive prediction methods is using the standard methods for learning systems' evaluation, such as holdout, Kappa statistics, McNemar test, prequential analysis.

The change detection module of online adaptive prediction methods can be considered as a type of classifier. One of the standard evaluation methods for classifiers is construction of confusion matrix which outputs are used to compute measures as accuracy, recall, precision, sensitivity or specificity.

In this paper, we present a framework for adaptive stream prediction methods and the modified measures we utilize for concept drift detection evaluation. The paper is structured as follows: in section 2, we mention existing tools for stream prediction and concept drift detection and their evaluation, in section 3 we describe our solution, section 4 presents our implementation of proposed methodology and testing of this solution, and the last section is the conclusion and future work.

2 Related work

The open source software Weka (Waikato Environment for Knowledge Analysis), from New Zealand, is a collection of algorithms for various data mining tasks, e.g. classification, regression, clustering, rules mining, and visualization.

Massive online analysis (MOA) [2] is related to the Weka environment. MOA is an open source framework for data stream processing. It also includes sets of machine learning algorithms for classification, regression, concept drift detection; frameworks for evaluation and supports building of a data stream. Currently, MOA provides sets of concept drift detection evaluators and does not allow to user insert his own algorithms via user interface.

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

3 Concept drift detection evaluation

We designed a framework for evaluation of a user's algorithms for concept drift detection. In the next subsections, we describe the evaluation measures and methods we utilize.

3.1 Adjusted confusion matrix

The standard way of confusion matrix's construction is a mapping of monitoring method's or specific measure's outputs onto real data and their evaluation as positive or negative.

If we consider some more complicated problem where a method evaluates any hypothesis as a positive or a negative, then a confusion matrix is built to reflect its correctness [5]. In such case, the confusion matrix consists of the outputs of the evaluation method, e.g. true/false values for positively/negatively labelled data instance vs. the real dataset instances.

In the standard confusion matrix's construction, it is important to detect the concept drift in the exact place of the drift appearance. Our adjustment of the confusion matrix's construction is based on the assumption that in data mining, and especially in the problems of data stream mining, it is important to detect the concept drift in some time or in a number of instances horizon.

Our adjustment of confusion matrix's construction is based on a window represented by time or count of instances from data stream. The size of the window is based on the type of data stream environment and the method used to detect concept drift. Imagine a part of data stream in time and the instances of data represented by dots (see Figure 1). Let say the instance with number 1 represents the point of real concept drift in the data stream. The confusion matrix constructor creates virtual window with size of 8 instances of data stream. This window creates space for the concept drift detector to identify the drift. If the detector identified the concept drift, for example, in the instance with number 2, this detection is labelled as true positive and the confusion matrix's true positive detection cell value is updated. In such case, the virtual window is closed and any subsequent detection, for example, in the instance with number 3, will be labelled as false positive.

3.2 Prequential analysis

The prequential method is used to evaluate learning algorithms [4]. It computes the error of the model with three possible estimators. Each of the estimators, holdout, sliding window and fading factors, define how is the sampling provided.

We use the holdout test, because we have a large dataset and the holdout test is used when there is enough data for testing.

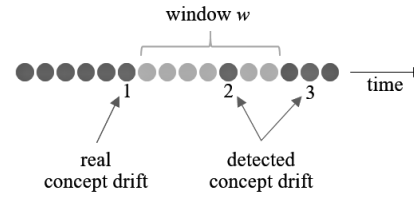


Figure 1. An example of confusion matrix's construction utilizing a window.

Sliding window and fading factors are used when in process of testing there is limited feedback from the environment or a lack of memory or time. We simulate limited access to data so it can be tested by these methods.

3.3 Matthews Correlation Coefficient

The Matthews Correlation Coefficient is used as one of the measures that are based on confusion matrix and its asset is in true/false classification problems evaluation [8]. The value of coefficient fits into interval $(-1,1)$ where -1 represents completely different result between predicted and real instance of data, 0 represents prediction, which is better than random predicted value and 1 represents the same predicted and real instances.

We use the Matthews Correlation Coefficient in a standard way as the measure of confusion matrix and also for evaluation of differences between the standardly constructed confusion matrix and the adjustment of confusion matrix's construction.

3.4 Kappa statistics

Kappa statistics is used to measure the reliability of unbalanced data stream [1]. We use a new measure, k_m statistics, derived from kappa statistics to compare one classifier to another classifier. This new measure is defined by equation (1):

$$k_m = \frac{p_o - p_m}{1 - p_m} \quad (1)$$

where p_o represents prequential accuracy of the actual classifier, and p_m represents prequential accuracy of another classifier.

3.5 McNemar test

The McNemar test is used to find the differences between two classifiers [3]. It is a statistical test and is computed with two variables: the first number defines the number of wrong classifications of the first classifier and the second number defines the number of correct classifications of the second classifier [1]. We are using the McNemar test to compare the methods for concept drift detection.

4 Framework for evaluation of adaptive prediction methods

For implementation of the presented evaluation framework with all of the evaluation measures and methods, we chose the R language and the Shiny framework to build a web application. Web application is easy to use. It accepts algorithms written in R language as an input and it allows users to set constraints for the evaluation.

The application provides the proposed methodology consisting of methods and measures that are suitable for

- evaluation of adaptive prediction methods that predicts numerical values of a stream/time series, and for
- evaluation of methods for concept drift detection.

It uses anonymized dataset of electricity usage in Slovakia regions. However, the application allows users to insert their own dataset, from which the application can create a data stream.

The main functionality of the application is the proposed methodology and its set of methods and measures to run and evaluate algorithms on input. The application processes the algorithm and then runs it on a data stream created from the selected dataset. While the algorithm is processing the data stream, the application constructs a confusion matrix, computes measures and evaluates accuracy of the algorithm. The outputs of evaluation are displayed in real time as tables (seeFigure 3) and plots.

| | Drift | NO drift |
|------------|-------|----------|
| Detected | 1 | 15 |
| UNdetected | 1 | 3151 |

Figure 2. Results of confusion matrix from test.

4.1 Framework testing

We test our solution by implementing common concept drift detectors, such as: CUSUM, Page Hinkley [6], geometric moving average, MAPE [9] and Pettit test [7]. We test the solution's ability to process the algorithm on input, to perform a prediction (or concept drift detection) by it and to evaluate the algorithm, i.e. to compute and visualize the results (seeFigure 3) in real time.

Test was performed on electricity data that consisted of 33 days with 2 concept drifts. Each day consists of 96 measures from which we create a data stream. The MAPE detector detected 1 drift correctly and the other did not. Another 15 drifts were detected

but incorrectly and the rest of the instances was correctly labelled without concept drift. These statistics were captured by our confusion matrix (see Figure 2) and on the other hand the standard confusion matrix had statistics with 0 correctly detected drifts and 2 incorrectly detected as no drift. Figure 3 shows the results of testing algorithm with MAPE as concept drift detector. These measures were computed from values in confusion matrix (see Figure 2). The best error rate is 0 and represents all incorrect predictions divided by the total number of dataset. PLR and NLR represents positive, respectively negative, likelihood ratio. DP represents disease prevalence what in our case means how often the drift occurred. PPV and NPV represents probability of positive or negative predicted drift label. MCC (Matthew's Correlation Coefficient) represents the quality of detector.

| | Value |
|-------------|--------|
| Error Rate | 0.0051 |
| Accuracy | 0.9949 |
| Precision | 0.0625 |
| Sensitivity | 0.5000 |
| Specificity | 0.0556 |
| PLR | 0.5294 |
| NLR | 9.0000 |
| DP | 0.0006 |
| PPV | 0.0625 |
| NPV | 0.0556 |
| MCC | 0.1755 |

Figure 3. Output measures of web application testing based on confusion matrix's values.

5 Conclusions and future work

The main contribution of our work consists of

- the proposed adjustment of confusion matrix's construction, and
- evaluation methodology for adaptive prediction methods and methods detecting concept drift.

In this paper, we have mentioned problems of data stream mining and evaluation algorithms dealing with data stream. We presented our adjustment of confusion matrix's construction based on a creation of a virtual window when the concept drift is present in the dataset. The window is defined by time or number of instances of data stream. We have collected existing methods and measures into methodology for evaluation of data stream algorithms and implemented a web application, which provides this methodology.

Our future work is going to consist of implementation of the remaining parts of the proposed application design, and testing of the final application. We want to perform the tests between students who are dealing with open challenges of data stream or batch mining. The ultimate contribution of the application should be a comparison of various prediction/concept drift detection methods in the power engineering domain, specifically, methods used for power demand forecasting from a data stream of smart meters' measurements.

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

References

- [1] Bifet, A. et al.: Efficient Online Evaluation of Big Data Stream Classifiers. In: *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '15*, (2015), pp. 59–68.
- [2] Bifet, A. et al.: MOA: Massive Online Analysis. *The Journal of Machine Learning Research*, (2010), vol. 11, pp. 1601–1604.
- [3] Gama, J. et al.: A survey on concept drift adaptation. *ACM Computing Surveys*, (2014), vol. 46, no. 4, pp. 1–37.
- [4] Gama, J., Sebastião, R., Rodrigues, P.P.: On evaluating stream learning algorithms. *Machine Learning*, (2013), vol. 90, no. 3, pp. 317–346.
- [5] Hamid, Y., Sugumaran, M., Journaux, L.: Machine Learning Techniques for Intrusion Detection. In: *Proceedings of the International Conference on Informatics and Analytics - ICIA-I6*, (2016), pp. 1–6.
- [6] Page, E.S.: Continuous inspection schemes. *Biometrika*, (1954), vol. 41, no. 1/2, pp. 100–115.
- [7] Pettitt, A.N.: A Non-Parametric Approach to the Change-Point Problem. *Applied Statistics*, (1979), vol. 28, no. 2, p. 126.
- [8] Powers, D.M.W.: Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation. *Journal of Machine Learning Technologies*, (2011), vol. 2, no. 1, pp. 37–63.
- [9] Vrablecova, P., Rozinajova, V., Ezzeddine, A.B.: Incremental Time Series Prediction Using Error-Driven Informed Adaptation. In: *2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW)*, (2016), pp. 414–421.

Testing Platform for Memory Management in Real-time Systems

Dávid BEŇO*

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

Abstract. In this paper, a new testing platform designed for evaluation of dynamic memory allocation algorithms focused on real-time systems is presented. The platform producing statistics about the allocation algorithms. A graphic user interface is included in this platform, which is launched on an external computer connected to the testing platform. The platform is performing the testing according to various test scenarios. For hardware, Raspberry Pi 3 microcomputer is used. The test results are sent via USB-UART to the external computer, where the results are displayed with GUI-based application. The main goal of the new testing platform is to be able to evaluate various new memory allocation algorithms and their implementations and to compare their attributes and performance to other existing algorithms. In addition to that, it will be possible to see online changes in memory during the execution of the testing scenarios. Because of repetitive allocation and freeing of the memory blocks of various sizes and locations, one can observe how fragmentation appears in the memory. After all of this, one can also compare different memory allocation methods and algorithms according to their response times and determinism.

1 Introduction

The difference between a classic “office” operating system and a real-time operating system is in its deterministic behaviour. Determinism means that we always get the same output each time as long as there are the same inputs and conditions. In the field of real-time operating systems, it means that we can predict the behaviour of these systems. Not only output of the system is guaranteed but the timing of providing the output is guaranteed too.

In the operating system, many applications, which will access the memory, will be launched. Allocation and assignment of memory space for an application, which is requesting this space, is one of

the main responsibilities of operating systems. In addition to that, precise memory management is also a critical area, which can disrupt determinism of the real-time systems. Because of the fact that the time complexity of classic memory allocation algorithms is not constant in respect to the size of the memory, these algorithms cannot be used in real-time operating systems because the response time of the whole real-time system would be too varying. Thus, dynamic memory allocation is usually forbidden in real-time operating systems. Therefore, real-time applications do not use dynamic memory allocation at all.

However, there is an ongoing research focused on dynamic memory allocation algorithms for real-time systems too. Although there is no ideal solution yet, some existing algorithms are providing a complexity near to constant time complexity and in consideration of the memory size and memory utilization, they are relatively good too. Due to this, a real-time testing platform for evaluation of real-time memory allocation algorithms would be very beneficial.

The goal of this work is to show how memory space transforms in time due to allocation of memory blocks of various size and how fragmentation appears due to freeing different blocks of memory.

2 Related work

Memory manager and process manager are two basic functions of operating system kernels. The memory manager allocates memory for each process and allows exchange of addresses between processes that are in progress.

Each variable has a memory space assigned, which corresponds to its type during the whole its “life”. Variable name is a symbolic address of the memory block, where the value of the variable is stored. Whenever a new variable is created, the function for assigning the memory to variable is called – allocation [1].

* Bachelor degree study programme in field: Informatics
Supervisor: Lukáš Kohútka, Institute of Electronics and Photonics, Faculty of Electrical Engineering and Information Technology STU in Bratislava

There are two basic types of allocation:

- static
- dynamic

2.1 Static memory allocation

This type of memory allocation is used when we know the amount of memory the process needs before it is executed. For example, if we add two numbers, we know that a space for two *integer* variables is needed. Compiler specifies needs of the memory for all variables and the bootloader allocates the memory in time of the execution. After this execution, while process is running, it doesn't manipulate with memory – there are neither allocations nor freeing of blocks – static allocated variables exist to the end of the process. Then the operating system takes control and destroys these variables.

However, the static allocation is not enough every time. For example, when calling a recursive function, we cannot predict how many times the function will be called and how much memory will be needed. Another case is that whenever a file is loaded to the memory, we don't know how big this file is. The idea of allocating so big block of memory that the whole file would be in it, is not correct. Such memory space may not be available at all. Moreover, it is wasting available resources too. In these and similar cases, it is recommended to use dynamic memory allocation instead. [1,7]

2.2 Dynamic memory allocation

The principle of dynamic memory allocation is that we can allocate memory block, while the process is running. We can access this memory via pointers.

Each programming language has its own functions for dynamic memory allocation. All we have to do is to call these functions. They are called *runtime* functions. Input parameters are not needed to be known in compilation time.

Basic functions used when working with dynamic memory (C language):

- malloc
- calloc
- free
- realloc

Malloc

The most used function in common programs. This function has only one parameter – *unsigned integer*, which is the amount of bytes to be allocated. The malloc function returns a pointer to *void*. Because of that, it is recommended to cast this pointer to suitable data type [1].

However, dynamic allocation takes more memory space than we send in function parameter. It is because the operating system needs some space for the memory management too. How much exactly depends on the particular operating system that is used [1].

Calloc

Similar function as malloc. However, this function has a feature that it initializes the allocated array to default value.

Free

This function frees allocated block of memory. In fact, it is a reverse function to malloc. In general, it is recommended to free the memory blocks that are no longer used. The parameter of this function is a *void* pointer, which points to an address of the allocated block of memory.

Realloc

This function reallocates memory – that means it changes the size of the existing memory block. By using this function, we can make array bigger or smaller. When increasing the size of an allocated array, the whole array is copied to the new block of memory with the new size. When decreasing the size, the size is simply changed [1].

2.3 Fragmentation

It is a situation when the memory is not effectively used or huge memory space is wasted. Either RAM and HDD can be fragmented. Fragmentation is making the system performance worse.

There are two basic types of fragmentation:

- external
- internal

External fragmentation

Because of many requests, the free space is fragmented in a way, that there is possibly sufficient free space to satisfy an allocation request, but the free space is not arranged in a contiguous fashion. In the worst case, the memory manager evaluates that there is not enough memory space to satisfy the request. A memory allocation request can fail due to the external fragmentation when there are “holes” between each two allocated blocks [2,4].

Internal fragmentation

This fragmentation type appears in a case, when the memory blocks with fixed size are allocated. They do not have to be same all the time – bigger blocks are assigned to processes with higher demands and smaller blocks are assigned to processes with lower demands.

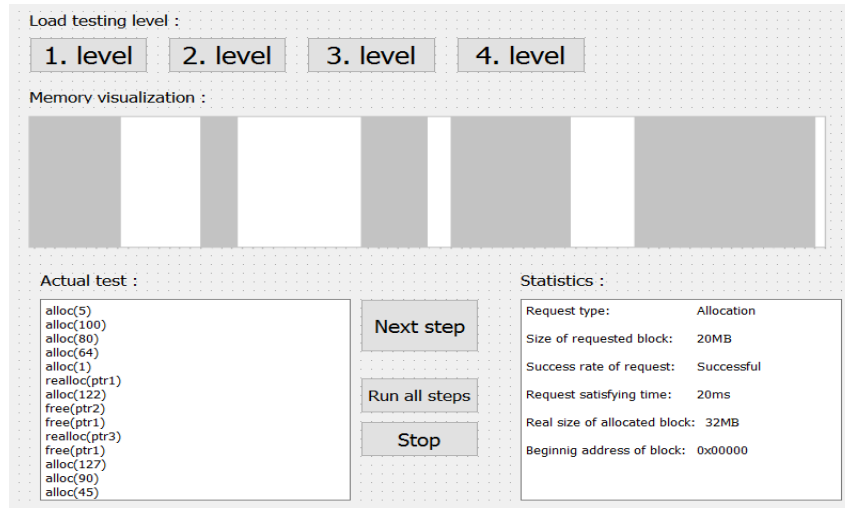


Figure 1. GUI part of platform.

Internal fragmentation is difference between the size of the assigned block and the size that is actually used. Although such a memory space is allocated, it is in reality wasted.

Internal fragmentation can appear in special implementation of the malloc function. The memory manager might prevent the external fragmentation. When allocating a block, it checks how much space remains for the next block. If the memory manager evaluates this space as too small, it assigns such space to the block being allocated [2].

2.4 Memory allocation algorithms

There are several algorithms, which are used for finding a free block of memory. These algorithms are implemented in functions like malloc. They search through a list of free blocks stored in some data structure. It is often a linked list, FIFO, LIFO or an address order. Free blocks are allocated the way the algorithm chooses them. Combination of data structure and algorithms results in different effectivity of allocation. For example, Best fit algorithm can be implemented as tree of lists with the same blocks or First fit can use Cartesian tree [3,5,6].

3 Graphic user interface

For better visualization of the current memory state, an application with graphic interface was designed. This is suited for an external computer that is connected to the testing device. This application is based on a cross-platform framework called Qt. The primary goal of the Qt framework is to be used for applications written in C++ language. However, it is also possible to connect the Qt GUI with an application logic written in another programming language, in our case it is plain C. The C language is

also the core language for FreeRTOS operating system. The GUI is displayed in Figure 1.

The graphic part of the application has three main purposes:

- Visualize an actual memory state
- Control of the testing platform
- Display test results and statistics

Visualization of an actual memory state

The application shows the state of the memory in real time. The memory is represented by a rectangle in full width of the application window. Each block of memory has one of white or blue color, white blocks are free and blue blocks are allocated. After clicking on a memory block, information about this block is displayed. This information contains the first and last address of the memory block and size of the block too.

Control of the testing platform

For better visualization of the tested algorithm, the user can pause the testing anytime and see how the memory is being changed step by step. Each request is stored in a FIFO queue and the application pulls the stack and executes the pulled request. The user can manually execute each request, which provides the user with the best control of the system and the best view of impact for the memory state.

Display test results and statistics

When the testing is finished, the GUI-based application provides results and statistics of the testing. These statistics show a number of requests for allocation (in detail), a number of successfully allocated memory blocks (this means a number of

successful requests), a number of unsuccessful requests, relative fragmentation and utilization of the memory.

4 Platform architecture

Our testing platform consists of two basic parts. The first testing part is running on Raspberry Pi 3. The second GUI-based part is launched in an external computer. This division is needed because of the fact, that the graphic interface needs dynamic accesses to the memory when working. However, if the memory is accessed when the tests are running, the result would be distorted (see Figure 2).

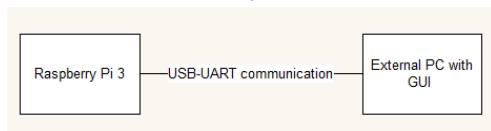


Figure 2. Map of platform.

Because our platform is divided to two parts, it is necessary to secure communication between these two parts. Sending and receiving information between is based on UART (Universal Asynchronous Receiver / Transmitter) for asynchronous communication. After execution of each request for memory, data are sent by UART to external GUI part. It does not matter if request has been successful or not, data are sent anyway. Moreover, it is important to get accurate data about memory state and it is important to show information about failed requests.

Here is information about requests. This information is used to create statistics and show the state of the memory:

- type of request (whether it is an allocation request or free request)
- size of requested block
- success rate of the request
- request satisfying time
- real size of the allocated block
- the beginning address of the block

5 Testing

An advanced functionality of our platform is comparing several allocation algorithms to each other. Each of these algorithms will be tested by several tests. These tests are divided to several levels, where each level has a specific scenario.

1. level – testing basic functionality
2. level – testing fragmentation with requests for small blocks
3. level – testing fragmentation with requests for large blocks
4. level – stress testing

The first level tests are testing if all functions of the memory allocation algorithm are working properly and are ready for next tests.

The second and third level tests are testing memory fragmentation. These tests are trying to find out how long it takes to fail any request due to the fragmentation. The difference between these two levels is that the second level contains requests for small size blocks. On the other hand, third level contains requests for large blocks.

The fourth level is stress testing and shows how the allocator handles many various requests. There are many requests for different size blocks of memory. The aim of this test is to ballast system with the most fragmentation as it is possible, so it is hard to allocate free block of memory and satisfy request. Result of this test shows how much time does it take to fail any request for due to the fragmentation.

6 Conclusion

In this paper, a new testing platform for evaluation dynamic memory allocation algorithms with the graphic user interface included was presented. With the help of this platform, it is possible to create statistics from the requests and show how the memory allocation algorithm works in time, while satisfying many requests. The main goal is to be able to compare various allocation algorithms.

We have described how the communication in this platform works and which information we show in statistics, as well as the state of the memory. Then, we have described the GUI of the testing platform, which shows the information gathered from the Raspberry Pi 3, and levels of testing.

Acknowledgement: This work was supported in part by the Slovak Research and Development Agency under grant APVV-15-0254 and by the Slovak Republic under grant VEGA 1/0905/17.

References

- [1] Herout, P.: Učebnice jazyka C (1. díl), 2009.
- [2] Štefanovič, J.: Základy operačných systémov, Správa pamäte, 2005.
- [3] Johnstone, M. S., Wilson, P. R.: The Memory Fragmentation Problem: Solved?, 1998.
- [4] Qualcomm Incorporated, Memory and Heap Technology Guide for Developers, 2012.
- [5] Robson, J. M.: Worst case of fragmentation of first fit and best fit storage allocation strategies, 1977
- [6] Andersson, K., Andersson, R.: A comparison between FreeRTOS and RTLinux in embedded real-time systems, 2005.
- [7] Pettersson, M., Svensson, M.: Memory Management in VxWorks compared to RTLinux: Linköping University, 2006.

Virtual Money Replacing Meal Vouchers

Marián ČARNOKÝ*

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

Abstract. In this paper we present a novel approach in payment system for the use of colored money. Nowadays are people hasty and they want to manage their tasks and obligations quickly and easily. Most of us have a smartphone, which is perfect gadget that makes our life easier. This device is already used in different payment systems. We design a paying system, which consists of Android and web application, that provides customers the possibility to pay for a meal or food in restaurants or supermarkets with their meal vouchers. People won't be able to use the paper vouchers, but only smartphones with NFC to pay at cashier. The aim of the work is to transform the paper form of vouchers to electronic form, to provide a faster and more secure way to pay and giving a solution for user credit management.

Unlike competing solutions, our application provides an opportunity to send credits to another person, pay with exact amount of credits, blocking the purchase of the selected type of goods.

1 Introduction

Nowadays people want to make their life easier and they wonder how this could be achieved even with the smallest changes. For example, people don't want to carry many of store loyalty membership cards, so they are adding them into mobile applications [1]. This can be applied even at meal vouchers and thus help humans to carry and worry less. Paying by paper form of meal vouchers at cashier is very slow process, because cashier needs to count the number of vouchers, stamp the vouchers and calculate the difference to pay. If the value of bought goods is less than the value of vouchers, the rest money forfeit to the store. The bought goods are controlled only by the cashier, so if the cashier does not reject the payment, you can buy tobacco or alcohol while vouchers are limited to buying food.

In our proposed system the vouchers are stored in Android application and voucher credits are managed through web application. The payment is initiated by

NFC communication between customer and cashier Android application.

2 Our proposed system

Android mobile application is designed for use by customer and cashier. The customer's part of application is responsible for transaction confirm and users account overview. The Cashier is using application to create a payment request and send a bought goods list to the server for verification. These goods must be verified, because the system allows to set rules for goods category purchase. It means, people cannot buy some kind of goods, if it is on a prohibited goods list of an account. Application uses Internet connection to send bought goods list, verify login information and download user details of an account.

NFC is chosen for transaction confirmation, because of fast and safe way of communication between two devices. Another possible solution was to use QR codes, but this solution was very slow way to initiate the transaction. Transaction is now initiated just by attaching devices to each other.

Web application is placed on web server. It's duty is provide an interface for meal voucher management. This application is web page created by HTML and Javascript. The application can be used only by system administrator and employer. The employer is managing voucher credits for his employees, creating new employees, adding new goods restrictions. The application is using PouchDB, which enables applications to store data locally while offline, then synchronize it with CouchDB and when the application is back online, keeping the user's data synchronized.

Database server stores data about users, transactions, credits and companies. The database is created by using CouchDB NoSQL database system. Data is represented by documents, which are sets of key value types. Data is gathered by MapReduce functions and data consistency is ensured by Multi-Version Concurrency Control (MVCC).

* Master degree study programme in field: Software Engineering
Supervisor Dr. Tomáš Kováčik, Institute of Computer Engineering and Applied Informatics, Faculty of Informatics and Information Technologies STU in Bratislava

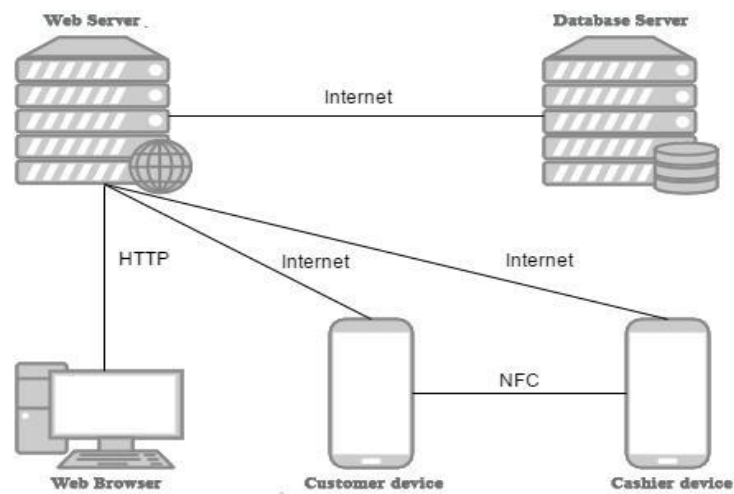


Figure 1. System architecture.

The system architecture is shown on Figure 1. The Table 1 contains a comparison of selected features and functions of our system called iMOPS with existing electronic system Ticket Restaurant Card [2] and GexPay [3]. These two solutions are chosen, because their features are similar to our system, they are deployed and used by people. The table also includes a column with comparison of classic paper voucher system.

3 Conclusions and future work

Our system called iMOPS, allow you to pay for food in supermarkets and restaurants with usage of Android device with NFC functionality. The system allows to manage tickets, employees and add new goods restriction by employer. Customer application has a functionality to get an overview of current account balance and to send a credit to another person. The system verifies the purchase of allowed goods types. Payment for goods can be done by exact amount of credit. Transaction is initiated by NFC

communication, which is the fastest possible way of connection establishment. This technology is not included in all smartphones, so in the future, we'll have to figure out a way like the rest of smartphones quickly establish a connection with another device.

References

- [1] K. Slegers, „Payment with Mobile (NFC) Vouchers – Using Co-design Techniques to Identify User Requirements,“ rev. *Mobile HCI*, Amsterdam, 2008.
- [2] Ticket Restaurant, „O Karte Ticket Restarant,“ [Online]. Available: <http://www.trkarta.sk/o-karte-ticket-restaurant>. [Cit. 1 February 2017].
- [3] GexPay, „Stravné listky v mobile,“ [Online]. Available: <https://www.gexpay.com/sk-SK/PublicPages/Home.html>. [Cit. 1 February 2017]

Table 1. Functionality comparison of solutions.

| | iMOPS | Ticket Restaurant card | GexPay | Paper voucher solution |
|-------------------------------|------------------------------|----------------------------------|------------------------------|------------------------|
| Pay | <i>By mobile application</i> | <i>By card or SMS</i> | <i>By Mobile application</i> | <i>Paper voucher</i> |
| Internet connection needed | <i>Cashier side</i> | <i>Cashier side</i> | <i>Cashier side</i> | <i>Not needed</i> |
| Transition initiation | <i>By NFC</i> | <i>By NFC or magnetic stripe</i> | <i>QR code</i> | <i>Manually</i> |
| Pay the exact amount | <i>Yes</i> | <i>Yes</i> | <i>Yes</i> | <i>No</i> |
| Send credit to another person | <i>Yes</i> | <i>No</i> | <i>Yes</i> | <i>No</i> |
| Overview of the balance | <i>Yes</i> | <i>Yes</i> | <i>Yes</i> | <i>Count manually</i> |
| Prohibition of goods | <i>Yes</i> | <i>No</i> | <i>No</i> | <i>No</i> |

Visualization of Student Activity in an Interactive Web-based Application

Lukáš MEŇHERT*

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

Abstract. Visualizations can significantly help the system in becoming more friendly by presenting information, that may otherwise stay hidden from users. For example, comparing the users' with the teachers' expectations can offer greater insight to the user. In this paper, we propose a designed set of visualizations, aiming to enhance adaptive learning system ALEF the new generation with explicitly opened learner model. Believing that visualizations can support both self-reflection and navigation in the system, we set out to design visualizations that add value to the learning process for both the students and the teachers.

1 Introduction

As stated in [7] the goal of visualization is to aid our understanding of data by leveraging the human visual system's highly tuned ability to see patterns, spot trends, and identify outliers. Well-designed visual representations can replace cognitive calculations with simple perceptual inferences and improve our comprehension, memory, and decision-making. The challenge, however, is to create effective and engaging visualizations appropriate to the data.

Our contribution is based on a simple premise of opening up a learner model to each individual student and teacher of the course. The model we are going to open is recorded by the system known as ALEF [10]. To open the system means giving access or showing the users the underlying data gathered by the system. These data create user's model. Model in short represents the current state of the student's domain knowledge.

Aforementioned system ALEF or Adaptive Learning Framework is designed to provide ease of creation for adaptive and interactive web-based educational environments. Right now, it serves as a learning resource for students in classes such as PSI (Principles of Software Engineering) and FLP (Functional and Logic Programming) on Faculty of

Informatics and Information Technologies in Bratislava.

One of the cornerstones with which ALEF was designed and constructed is the belief that students who interact with a teacher and other students in the class, should be able to interact with each other in the virtual environment as well. ALEF supports this by allowing students and teachers to comment on a particular question in the system. However, students cannot communicate their progress in the system with each other. Thus, students do not know how they compare to the other students in the course. One way to address this problem is to show them a comparison of their activity to the overall activity of students in the system. This can be also done by the form of visualizations.

2 Related work

Growth and popularity of visualizations prove, that it is simply one of the best ways to interpret data in a way that can reduce cognitive strain for the user. This claim is supported by a collection of other works such as MasteryGrid [2], SQL-tutor [8], OLMlets and Flexi-OLM [3], Narcissus [10] and others, which found out that opening a learner model, and therefore the usage of visualizations, supports the learning process. There is a variety of reasons for opening the learner model [4], and here we summarize reasons, that are relevant for our work:

- Addressing the user's right to view data about themselves;
- Aiding navigation directly from the open learner model;
- Raising awareness (of knowledge, progress, difficulties, etc.);
- Facilitating planning;
- Facilitating competition amongst learners;

* Bachelor degree study programme in field: Informatics

Supervisor: Professor Mária Bielíková, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

- Helping the learner to take greater control over their learning decisions.

In a research study done by A. Mitrovic [9] it was confirmed, that even simple open learner models consisting of skill-meters, have a positive impact on learning and metacognitive skills. For example, even though less-able students ended up with worse results at self-assessment, their performance increased in comparison to their peers. It was also confirmed, that students reflected more on their knowledge by selecting problems that suited their level of understanding.

3 Designing process

The opening of a learner model could easily be reduced to giving access to raw tables found in a database. This, however, would not be very helpful to anyone except a few individuals, that would be really interested, and had the needed capacity to understand the shown dataset. In this example, we can see the improvement visualizations can provide over raw data.

Every process, not just that of a visualization design, should start by identifying our goals. Without this, we wouldn't know what our efforts are leading us towards. We should pay attention to the size of the dataset, its complexity, quality, needs of end users; students and teachers while doing so.

There are a few guidelines to help us choose between strategies when creating a design for a new problem [6]. Each of them has its trade-offs and these designs are built from the three primary building blocks: juxtaposition, superposition, and explicit encodings.

3.1 Choosing the right visual representations

Since our main goal is to design visualizations for two types of users, identified goals are not the same for both of them. Speaking of visualizations in a context of educational applications, preferences vary from user to user, but a few features stand out from the rest and are preferred more than the others [5]. We also embraced these features as a goal for student visualizations:

- Navigation aid;
- Raised awareness of knowledge;
- A right to view data about oneself;
- Raised awareness of misconceptions and problematic areas;
- Comparison to current expectations and domain.

While identifying goals for teachers' set of visualizations, we originated from consultations with our supervisor:

- Effective selection of questions suitable for weekly students' tests of the course;
- Effective identification of students, that display frequent and quality preparation;
- Detection of problematic concepts or questions among students;
- Overview of the students' activity.

Our questionnaire, in which 25 students took part, was aimed at students' preferences regarding visualizations in an educational environment, and it revealed that 80% of students preferred interactive visualizations over static ones. We presume that the reason for this is, that dynamic visualizations can tell stories in a way static ones cannot.

3.2 Designed set of visualizations

To have as much freedom in visual design as possible, we chose to implement our visualizations in a D3.js [1]. It is a java script library similar to jQuery in a way it can manipulate DOM structure with a few perks such as transitions and typical selections.

Figure 1 shows the first proposed visualization. This visualization compares the student with a course average of either time spent in the system or ratio of successfully answered questions. Its purpose is to motivate a student and offer him a tool for self-reflection. On the x-axis, we can see pair of bars, each corresponding to one week in semester. Height of the thin bar represents average value of students while the wider bar represents value of one student. It is coloured on scale from red to green according to student value, average value ratio. If the students value is higher than average, the colour will be in shades of green, otherwise it will be coloured in shades of red.

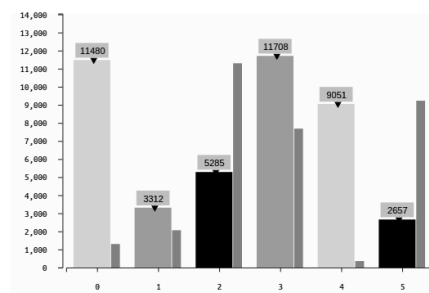


Figure 1. Visualization for students: time spent in the system.

To break down concepts to those which students grasped and those they had a problem with, we propose the second visualization shown in Figure 2. A zoomable map of concepts supporting navigation

to related questions. ALEF contains 67 different concepts in PSI course spanning from ‘activity diagram’ to ‘project management’. By clicking on a rectangle representing the concept, a student will be shown details with a list of questions related to this concept. This way, the student can immediately navigate to questions bound to the problematic concept.

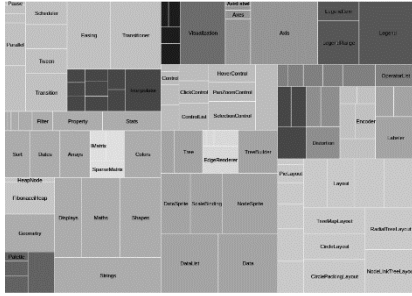


Figure 2. Visualization for students: understanding of the different concepts in a tree map.

Size of the cell in the tree map is proportional to the number of related questions, while the colour of the cell will be either red or green according to whether the student understood the concept as expressed by equation (1). $Succ$ is number of successfully answered questions from the last twenty, $nsucc$ number of incorrectly answered question from the last 20 and rqs is ratio of questions that were correctly answered at least once and all of the questions. If the $u > 1$ the concept will be labelled as understood.

$$\frac{succ}{nsucc} * rqs = u \quad (1)$$

Visualization in Figure 3 should help the teacher to identify questions suitable for the students in the upcoming test. The advantage being, that it shows relationships between all the questions. With appropriate filters, we assume it will be easy to use and will prove to be useful for identification of inaccurately formulated questions. Questions represented by dots, that have comments in discussion are highlighted. Dimension of dot volume is not used, every dot has the same volume.

Visualization presented in Figure 4 is supposed to provide basic stats for the teacher in the form of skill meters. With functional sorting mechanism, this visualization is mainly designed to identify a certain group of students, for example, the most studious ones.

Together, this set should be sufficient for a view of the most basic statistics or comparisons and a little more. Teacher will be able to download the data in .csv and .tsv format, for more advanced processing if needed.

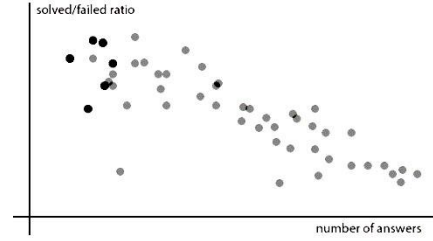


Figure 3. Visualization for teachers: identification of problematic questions.

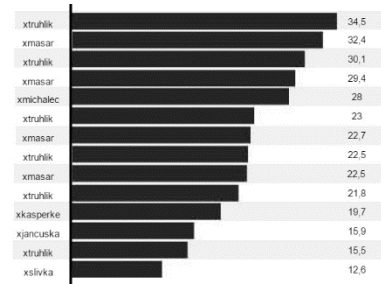


Figure 4. Visualization for teachers: statistics of users.

4 Evaluation

With visualizations designed to support the learning process of students we plan to monitor students throughout the year with a web application, that will function alongside ALEF. The application will monitor user activity and supply us with proper data. With these, we will be able to spot relations between students that use visualizations and students that do not.

In another experiment, we plan to gather 20 students in controlled environment with eye-tracking devices. We will give students assignment with simple tasks to evaluate usability of proposed visualizations measuring number of mouse clicks needed to fulfill the task, time it took to fulfill the task and number of fulfilled/unfulfilled tasks. In the end students will fill questionnaire to record their experience.

5 Conclusions

In this work, we pointed out different aspects of visualizations and the associated data that should be considered while designing visualizations. Facing endless options to visualize data, we believe we focused on the ones, that are most important in the domain of education. With this, we introduced a set of visualizations and described the thinking process that went into designing it. Visualizations and options

of interacting with them may change over time according to our new findings and results of the experiment. Finally, we outlined the next phase in which we will be implementing the web application, that will help us in the evaluation of the usefulness of the designed visualizations.

Acknowledgement: This project is the partial result of the collaboration within the SCOPES JRP/IP, No. 160480/2015 and the project „Rozvoj výskumnej infraštruktúry STU, projekt č. 003STU-2-3/2016“ by the Ministry of Education, Science, Research and Sport of the Slovak Republic.

References

- [1] Bostock, M., Ogievetsky, V., Heer, J.: D³ data-driven documents. *IEEE transactions on visualization and computer graphics*, 2011, 17.12: 2301-2309.
- [2] Brusilovsky, P., et al.: Open social student modeling for personalized learning. *IEEE Transactions on Emerging Topics in Computing*, 2016, 4.3: 450-461.
- [3] Bull, S., et al.: Supporting interaction preferences and recognition of misconceptions with independent open learner models. In: *International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems*. Springer Berlin Heidelberg, 2008. p. 62-72.
- [4] Bull, S., Kay, J.: Student Models that Invite the Learner In: *The SMILI Open Learner Modelling Framework*. *Int. J. of Artificial Intelligence in Education* 17(2), 89-120 (2007)
- [5] Bull, S.: Preferred features of open learner models for university students. In: *International Conference on Intelligent Tutoring Systems*. Springer Berlin Heidelberg, 2012. p. 411-421.
- [6] Gleicher, M., et al.: Visual comparison for information visualization. *Information Visualization*, 2011, 10.4: 289-309.
- [7] Heer, J., Bostock, M., Ogievetsky, V.: A tour through the visualization zoo. *Commun. Acn*, 2010, 53.6: 59-67.
- [8] Mitrovic, A.: A knowledge-based teaching system for SQL. In: *Proceedings of ED-MEDIA*. 1998. p. 1027-1032.
- [9] Mitrovic, A., Martin, B.: Evaluating the effect of open student models on self-assessment. *International Journal of Artificial Intelligence in Education*, 2007, 17.2: 121-144.
- [10] Šimko, M., Barla, M., Bielíková, M.: ALEF: A framework for adaptive web-based learning 2.0. In: *Key Competencies in the Knowledge Society*. Springer Berlin Heidelberg, 2010. p. 367-378.
- [11] Upton, K., Kay, J.: Narcissus: group and individual models to support small group work. In: *International Conference on User Modeling, Adaptation, and Personalization*. Springer Berlin Heidelberg, 2009. p. 54-65.

Classification of Short Texts from an Interactive Web Application

Michal HRUTKA*

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

Abstract. This paper describes experiments with various machine learning algorithms applied on short texts coming from an interactive web application in order to classify them as appropriate or not. We provide a description of the algorithms we have experimented with along with the results they provided on our dataset.

1 Introduction

The problem we are dealing with in this paper can be defined as follows:

The interactive web application is used on events with speakers, for example a panel discussion, where attendees can ask questions. They can ask their questions by sending them using this application for everyone else to see. Each question has its own score, which is a number of votes it received from the attendees. In some cases, inappropriate questions are submitted, which can negatively influence the event as well as the usability of the application.

To counter this problem, the application offers a feature called moderation, which when enabled requires an administrator to manually approve or discard each incoming question before it is displayed for everyone else to see.

In this paper, we describe how machine learning can be leveraged to perform such moderation automatically, by performing a simple binary classification for each question, which labels them as appropriate or inappropriate.

We have achieved very good results, which show us that this problem can be solved using simple approaches.

2 Related work

2.1 Papers

In the paper [5] the authors tried to apply methods usually used for long text classification to short text

messages. They compared the performance of *Naïve Bayes* classifier, *k-nearest neighbors* and *Support vector machine*.

They have found that short texts require different classifiers than longer texts to achieve best performance. *Naïve Bayes* and *SVM* outperformed the *k-nn*.

The authors in [6] explore the effects of weighing word-based features in binary text classification problems. They compare the improvement of model performance after applying *tfidf* to a dataset labelled for different classification tasks. The authors have found that when this technique is applied to a domain specific dataset it can negatively influence the performance of a classifier because it values rare features and diminishes the impact of frequently occurring features. Therefore, some of the most important features can be discarded by this technique.

2.2 Algorithms

The *perceptron* belongs among the first machine learning algorithms used for classification. It is described as a simple logic gate with multiple inputs and binary outputs. The perceptron sums inputs (features), each feature is multiplied by its weight, and if a set threshold is reached, the output is 1, otherwise -1. This makes the perceptron capable of binary classification [1].

Adaline, short for *Adaptive Linear Neuron*, is the improved version of the Perceptron. The main difference between Adaline and Perceptron is that a linear activation function is used to update the weights of inputs instead of the unit step function [2]. Adaline uses the output of the *Activation function* for updating the weights of inputs instead of the predicted class.

Logistic Regression is used for classification problems, even if the name of this algorithm suggests otherwise. It is also one of the most popular classification algorithms. The main difference

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

between Logistic Regression and Adaline is that Logistic Regression uses the Sigmoid function instead of the Activation function [1].

Support Vector Machine, also known as *SVM*, is a popular machine learning algorithm, an improved version of the Perceptron. The difference is that the Perceptron minimizes the misclassification errors, while the SVM maximizes the margin [2] – the distance between the decision boundary and the data samples (support vectors) nearest to it.

Decision Tree models are popular because they are easy to understand. This machine learning model separates a given dataset by creating conditions. Depending on the features of the dataset, the Decision Tree model finds several conditions in order to predict classes of all data samples.

Random Forest is a combination or an ensemble of decision trees. The goal of this algorithm is to create a strong learner by combining multiple weak learners together. The advantages of this approach are that it is less likely to overfit and has a better generalization error [1]. It performs very well on classification problems, is easy to use and scale.

Naïve Bayes is a very simple linear machine learning algorithm, which has become popular for performing very well despite its simplicity. It is called naïve because of the fact that it does not take any dependency among features into consideration. In some cases, this algorithm performs as well, or even better, as more complex machine learning algorithms. Naïve Bayes classifier comes in two versions. The Bernoulli model takes into account only the presence of features in data samples. The Multinomial model takes into account not only the presence but also the number of occurrences of features in data samples [4]

3 Comparison of different approaches to short text classification

3.1 Dataset

The dataset we have been working with contains over 27.000 questions from 96 events. The selection of these events was based on correct use of the *questions feature* (an example of incorrect use of this feature is when a speaker asks a question and the attendees write their replies via this application), the topic of the event, number of attendees and the use of the *moderation feature*.

For each question the dataset contains the following information:

- text of the question,
- label (appropriate / inappropriate),
- whether a question was sent with a name or anonymously,
- the score of the question and

- timestamp of when the question was sent, approved, and answered.

There are multiple sources of noise in the dataset, which can cause a decrease of the trained model precision. There is no precise information whether the *moderation feature* was being used or not. We can only deduce it from the fact that the timestamps of the question's creation and approval differ. Questions are often discarded not only because they are inappropriate, but also because a similar or identical question has already been asked (in fact, there are many questions discarded because of such duplicities). It is also necessary to take into account, that moderating questions is subjective, and two different persons can have different opinions, one might find an inappropriate question appropriate and vice versa.

3.2 Data pre-processing

To transform the text of each question to a form suitable for a machine learning model, a so-called *bag of words*, we experimented with three different vectorizers.

A vectorizer takes text of questions as input, applies a tokenizing function and returns a sparse matrix representation of it. The tokenizing function splits each question into ngrams, applies a stemmer on the ngrams and removes stop words – words that hold no meaning by themselves or are too common.

The sparse matrix represents the dataset as follows: each column represents a single feature (word, n-gram, etc.) and each row represents a single data sample (question). Depending on the configuration of the vectorizer, the values can represent either the presence of a feature (binary values) or the number of occurrences of a word in a data sample. In such case, we define this as a *Count Vectorizer*.

Similarly to *Count Vectorizer*, a *Hashing Vectorizer* also transforms text into a sparse matrix with values representing the number of occurrences of features in documents. The difference is, that the hashing vectorizer does not create a vocabulary, but it uses a hashing trick to find the token string name (word) to feature integer index mapping.

In text classification, there are often words that appear in many data samples which belong to different classes. These words decrease the performance of machine learning models, because they do not contain useful information. The *Tf-idf vectorizer* is designed to lower the importance these words have in the machine learning process. Tf-idf is short for *term frequency-inverse document frequency*.

3.3 Methodology

To evaluate the performance of the machine learning models, we have used *k-fold cross-validation*, which divides the dataset into *k* subsets, marks one subset as the *test* subset and trains the model on the remaining data. This process is repeated *k* times and an average of the performance metric is returned as the final score of the model.

The input for each machine learning model is the sparse matrix returned by a vectorizer and a vector of labels – appropriate or inappropriate. The model is then evaluated on whether the predicted class matches the corresponding label.

We chose the *Matthews correlation coefficient* [3] as the performance metric. It takes into account true and false positives and negatives and is generally regarded as a balanced measure which can be used even if the classes are of very different sizes.

We have also experimented with adding *bigrams* to the *feature vector*. A *feature vector* is the collection of features that are observed for each data sample. When a feature consists only of one word, it is called *unigram*. When a feature consists of two consecutive words, it is called *bigram*.

Next, we have applied *feature selection* to remove irrelevant features that do not contribute much to the classification problem solving. The simplest approach to feature selection is to evaluate the relevance of each feature separately, sort the features according to this metric and select *n* best. The best performance was reached after reducing the number of features to *one fifth* of the original feature vector size.

Finally, we were also experimenting with adding one additional class to the dataset by labelling the questions that received a very high score with a different label in order to separate the questions into three categories: inappropriate, appropriate and great questions. During the evaluation process, the predicted results were converted back to fit binary classification evaluation, as it was not necessary to differentiate between appropriate and great questions. However, the results showed that for some of the algorithms the performance did not change and for other it even decreased.

3.4 Results and Discussion

All of the above-mentioned techniques were applied using all of the algorithms mentioned in section 2.2. For this project, we chose python as the programming language and used the implementations of machine learning algorithms from a library called scikit learn¹.

We have found that the worst performing algorithms were the *perceptron* and *Adaline*.

The *random forest* achieved good score, but the amount of time it took to train a random forest classifier was significantly higher than any other algorithm we were experimenting with.

Surprisingly, the *Naïve Bayes* performed very well, achieving better performance than the *random forest* and even though the training process was very fast.

The best performing algorithms were *Logistic regression* and *Support vector machine*. Both of these algorithms achieved the highest score and the time of the training period was still low, higher than the *Naïve Bayes* but significantly lower than the *random forest*.

We have observed that the *k-fold cross-validation* minimizes the risk of overfitting a model to a single training subset of the data and provides a much clearer result than randomly choosing the training subset.

We experimented with adding more features to the sparse matrix to improve the performance, specifically whether the question was sent anonymously or not and the length in words, but both have proven to be negatively affecting the performance of the classifiers.

To include the question score, we could not simply add it to the sparse matrix, because with new questions, the score is not known. Therefore we experimented with adding one additional class for questions with above average scores. After the prediction, the labels were converted to fit the binary classification – the two labels representing questions that were appropriate, with both below and above average scores were united. However, this approach also proved to degrade the performance of the classifiers.

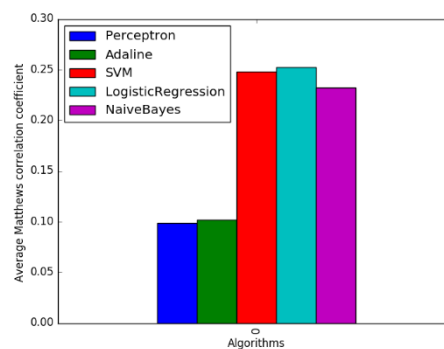


Figure 1. Comparison of average Matthews correlation coefficient for used algorithms.

We found that the *Matthews correlation coefficient* is a suitable metric for evaluating the model

¹ <http://scikit-learn.org/stable/documentation.html>

performance since the dataset contains mostly appropriate questions and focuses also on false negative and false positive classifications.

After adding *bigrams* to the *feature vector*, we have observed a decrease in the model performance. We assume that the information gain from a bigram is not greater than a random combination of unigrams, but adding bigrams to the *feature vector* increases variance [5].

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

References

- [1] Raschka, S. *Python Machine Learning*. Birmingham : Packt Publishing Ltd., 2015
- [2] Murphy, K. P. *Machine learning : a probabilistic perspective*. Massachusetts : s.n., 2012. ISBN 978-0-262-01802-9. Henessy, J., Patterson, D.: Computer Organization and Design: The Hardware/ Software Interface. Morgan Kaufmann Publishers, San Mateo California, (1994).
- [3] Pedregosa *et al.*, *Scikit-learn: Machine Learning in Python* ., JMLR 12, pp. 2825-2830, 2011.
- [4] Downey, A. B. *Think Bayes: Bayesian Statistics Made Simple*. Needham : Green Tea Press, 2012. ISBN 978-1449370787.
- [5] Bekkerman, R., & Allan, J. (2003). Using bigrams in text categorization (Technical Report IR-408). Retrieved February 16, 2017 from the Center for Intelligent Information Retrieval, University of Massachusetts website: <http://ciir.cs.umass.edu/pubfiles/ir-408.pdf>
- [6] Martineau, J. et al.: Improving binary classification on text problems using differential word features. In Proc. of the 18th ACM conference on Information and knowledge management (CIKM '09). ACM, 2009, pp. 2019- 2024

Determination of the Eye-catching Parts in Graphical Interfaces

Patrik BEKA*

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

Abstract. Eye-catching and graphically attractive design is an extremely important part of every website, although it might not appear to you at first sight. Our main purpose is to develop a neural network, capable of learning from given data and eventually predicting eye-catching and important parts of the websites. We selected method based on the convolution neural network, which is able to make predictions from the given images of the web pages. Final prediction is shown in the form of the heatmap, which determines the most engaging parts of the given web pages.

1 Introduction

Design of the good graphical interface for web pages is not an easy thing to do at all. Nowadays, when people are looking for informations almost only via internet, the good design has even bigger importance. Whether the visitor of page is interested in the content or not, it is not as important as his attraction to the design of web page and the fundamental availability of information. Therefore, design of the web page is certainly one of the most important preconditions for the future comeback.

Our task is to develop the system, that is capable of determining the eye-catching parts of web pages just from the web page image (screenshot). In this thesis, we try to describe our method of determining important parts of the web pages using neural network, specifically convolution network, along with other similar solutions.

2 Related work

Convolution neural networks [2, 7] are widely used almost in every domain, where is the need for image recognition. Whether it is automatic face detection

on Facebook, autonomous cars capable of self driving using autopilot by Tesla, Google, or software for sorting and classification of cucumbers in the Japanese farm¹. Prototype of self driving car by Google called Dave-2 [1] has a model of convolution neural network, which processes frames from cameras placed on the car. These cameras are recording environment in 10 frames per second and after some modifications, these frames are passed to the neural network. This network consists of 9 layers, one normalization layer, 5 convolution and 4 fully-connected layers.

Related work about judging the importance of different parts in a web page can be classified into two main approaches.

First approach is using HTML code of the web page to divide the whole web page to the main parts with operation called segmentation. Segmentation uses either segmentation algorithms to divide page into the blocks by different features, or data structures to represent components and elements of HTML document. Most commonly used approaches to segmentation are:

- DOM-based segmentation:

HTML document is represented as a DOM [4] (Document Object Model) tree. Tags are representing block of pages, for example P-paragraph, TABLE-table. Very accurate representation of the structure of HTML document, but not accurate enough for dividing of visual different blocks.

- Location-based segmentation [10]:

Page is divided into 5 parts: center, left, right, bottom, top. Problematic may be scrollable pages, when page has different layout of the objects after scrolling.

* Bachelor study programme in field: Informatics

Supervisor: Dr. Mária Šajgalík, Institute of Informatics, Information Systems and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

¹ <https://cloud.google.com/blog/big-data/2016/08/how-a-japanese-cucumber-farmer-is-using-deep-learning-and-tensorflow>

- VIPS (Vision-based Page Segmentation [3]) algorithm:

Combination of the previous two examples of segmentation. Dividing page by color, size of blocks, etc. At first, the appropriate nodes, which imply the horizontal and vertical lines of web page, are found in the DOM tree. Based on that, semantic tree is created. It is a tree, where every segment is a separate node. Continuity of the segmented page is controlled by permitted degree of coherence (pDoC [9]), which ensures keeping content together, while semantically different blocks apart.

After segmentation, the importance of blocks is determining. It could be done using heuristics [8], human assessors (who manually label the blocks), etc.

As an example of working solution can be mentioned the work of Microsoft researchers [10]. They used VIPS algorithm for page segmentation and 5 human assessors to label each blocks of web pages. Blocks were numbered by importance from 1 to 4, from insignificant information (such as adds) to the most important part of web page (news, products, etc.). Authors of this solution presume that people have consistent opinions about the importance of the same block in a page. After every block was labeled, the model of importance was created as a mapping function of every block and its importance:

$$\{\text{block features}\} \rightarrow \{\text{block importance}\} \quad (1)$$

For estimation of the block's importance was used neural network, where blocks are represented as tuple $\{x, y\}$. Number of block is x and its importance is y (real number). The network type was RBF (Radial Basis Function) and the neural network used standard gradient descent.

Second approach to determine important parts of web pages is based on the image of web page and a neural network that predicts the most engaging parts of web page in the form of heatmap, saliency map, etc.

Solution using this approach was created by Shen Chengyao and Zhao Qi [11]. They created a dataset divided into the categories by the content of web page (pictorial, text, mixed). Each category contains 50 images, which were shown to the 11 subjects. Their sequences of views were recorded using MATLAB with Psychtoolbox and Eyelink 1000. MKL (multiple kernel learning) was used for views prediction, which was trained as a binary regression problem.

3 Method of predicting eye-catching parts of the websites

After closer research of the problem domain, we decided to use second approach of determining important parts of the web pages. Our method is based on neural network, particularly convolution neural network, which is able to predict the eye-catching parts just from the image of a web page. After experimenting with various architectures of neural networks, we decided to use architecture shown in Figure 1, because it provides the best results. Shown architecture contains one convolution layer and one max pooling layer followed by normalization, fully-connected and output layer.

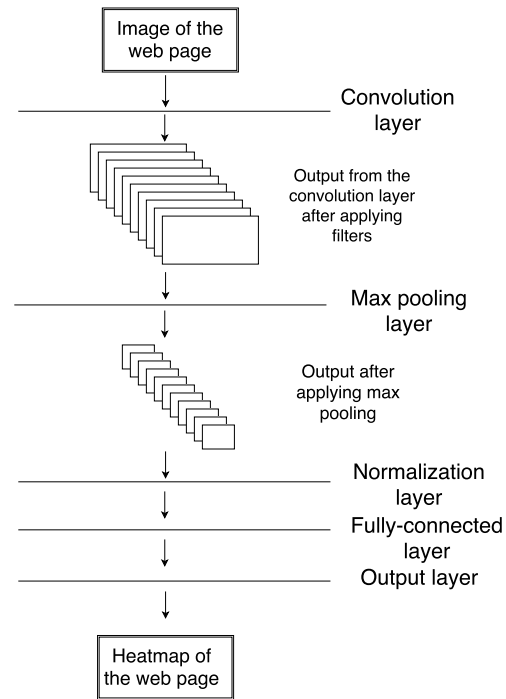


Figure 1. Diagram of the neural network's architecture.

Convolution layer contains convolution filter, which size is 5x5. Activation function on this layer is standard ReLU:

$$f(x) = \max(0, x) \quad (2)$$

After processing, data is passed to the max pooling layer, where the output from previous layer is processed using window, which size is 2x2, and MAX operation. Output of all filters is merged into one wide flat layer, normalization layer. This layer is followed by fully-connected and output layer, where the

whole prediction is made. The output layer contains final predicted heatmap for image, which size is $m \times m$, where m is the size of both, image and heatmap.

4 Evaluation

Our dataset was collected as a part of experiments, that are described in bachelor thesis of Mária Dragúňová [5]. Now it contains 15 images of different web pages with 44 sequences of views (fixations) on these images in first 5 seconds.

Images sizes are 1920×1080 , what is too much. Therefore, the size is reduced to 256×256 , which is also more convenient for neural network.



Figure 2. From left: original image, heatmap, heatmap on image.

Fixations consist of X and Y coordinates for every view (range from 0 to 1) and duration of views. Number of fixations is variable for every image, the range is from 3 to 20. Heatmaps are calculated from these fixations using normal (Gaussian) distribution. First of all, every fixation is converted into the point on

image of web page, as it is shown in equation 3. That ensures same size of heatmaps and images.

$$[x, y] = [\text{fixation_x} * 256, \text{fixation_y} * 256] \quad (3)$$

The probability density function of normal distribution is shown in equation 4, and it is calculated for every fixation from all other fixations.

$$f(x | \mu, \sigma^2) = \frac{1}{\sqrt{2\sigma^2\pi}} * e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (4)$$

x - current point for which the normal distribution is used

$(x - \mu)^2$ - distance between current point and other fixation

σ^2 - duration of fixation For visualization of heatmaps on images, simple alpha blending is used, as you can see in Figure 2.

Before passing data to the convolution neural network, we divide the data (images, heatmaps) into 16 smaller parts (buckets), so every image and heatmap is basically 4×4 grid (shown in Figure 3), consisting of squares of size 64×64 . This division of data is done mostly because of the need for a larger training dataset.

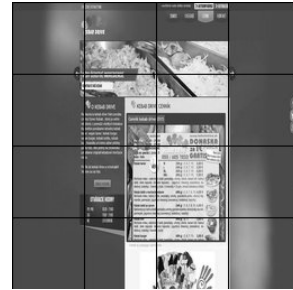


Figure 3. Visualization of the grid on image.

The parts of divided images are the input data, labels are equally divided heatmaps. Currently, accuracy of predictions is about 20%, what is not much.

5 Conclusion

This thesis describes a prototype of neural network, that is capable of predicting the eye-catching parts of the given web pages in form of the heatmap. It does not need any source code of the web sites, only screenshot is required. That may be considered as a solid advantage compared to some other solutions of similar problem.

Prototype is written in Python using Tensorflow² framework for neural networks. Results will be compared to the saliency model of Itti and Koch [6], but only after some modifications and more training. In future, we should consider optimization for network, maybe try to add more hidden layers. We may also perform another experiment with different web pages and let more people look at them, analyze it so we will have more data for training our convolution neural network and hopefully the dividing of data to 4x4 grid wouldn't be needed.

As a real-life application of our solution, I could imagine a web page, where people just upload screenshot of the web pages design and get a visualization of predicted engaging parts in form of the heatmap. That could provide solid information for administrators of web sites, so they would be able to decide where to place important elements and where adds should be placed in order not to cause distraction.

Acknowledgement: This contribution was created with kind support of ČSOB Foundation and is partial result of the project University Science Park of STU Bratislava, ITMS 26240220084, co-funded by the ERDF.

References

- [1] Bojarski, M., Del Testa, D., Dworakowski, D., Firner, B., Flepp, B., Goyal, P., Jackel, L.D., Monfort, M., Muller, U., Zhang, J., et al.: End to End Learning for Self-Driving Cars. *arXiv preprint arXiv:1604.07316*, 2016.
- [2] Britz, D.: UNDERSTANDING CONVOLUTIONAL NEURAL NETWORKS FOR NLP. <http://www.wildml.com/2015/11/understanding-convolutional-neural-networks-for-nlp/>.
- [3] Cai, D., Yu, S., Wen, J.R., Ma, W.Y.: VIPS: a Vision-based Page Segmentation Algorithm. Technical report, 2003.
- [4] Chakrabarti, S.: Integrating the document object model with hyperlinks for enhanced topic distillation and information extraction. In: *Proceedings of the 10th international conference on World Wide Web*, ACM, 2001, pp. 211–220.
- [5] Dragúňová, M.: Vyhodnocovanie používateľského zážitku analýzou pohľadu a emócií. Bachelor thesis, Bratislava: FIIT STU, 2016.
- [6] Itti, L., Koch, C.: A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, 2000, vol. 40, no. 10–12, pp. 1489 – 1506.
- [7] Li, F.F., Karpthy, A., Johnson, J.: CS231n: Convolutional neural networks for visual recognition, 2015.
- [8] Liu, Y., Wang, Q., Wang, Q. In: *A Heuristic Approach for Topical Information Extraction from News Pages*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006, pp. 357–362.
- [9] Mehta, R.R., Mitra, P., Karnick, H.: Extracting semantic structure of web documents using content and visual information. In: *Special interest tracks and posters of the 14th international conference on World Wide Web*, ACM, 2005, pp. 928–929.
- [10] Ruihua Song, Haifeng Liu, J.R.W., Ma, W.Y.: Learning block importance models for web pages. In: *Proceedings of the 13th international conference on World Wide Web (WWW '04)*, New York, USA, 2004, pp. 203–211.
- [11] Shen, C., Zhao, Q. In: *Webpage Saliency*. Springer International Publishing, Cham, 2014, pp. 33–46.

² <https://www.tensorflow.org/>

Indoor Free Time Places Availability

Tomáš LACH*

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

Abstract. In today's great complexes of buildings it is important to track people's movement. We propose extension to faculty application VirtualFIIT, which visualize availability of free time places such as buffet, hammocks or library. For localization we use Bluetooth Low Energy beacons. Important part of proposed method is calculation of actual user position, which is obtained by least squares trilateration method. We have already achieved localization accuracy 3 meters. Location of students will be visualized in application indoor maps part. Our prototype aggregates data about students movement, what can be basis for many other use cases such as app personalization.

1 Introduction

It is hard for new students of faculty to orientate in a new environment. Not only for these students, there is VirtualFIIT application¹. This app helps complete everyday tasks of student. It displays student's timetable, maps of faculty with individual rooms, student canteen menus and public transport timetables which departs near to faculty.

The majority of students of our faculty does not visit faculty building only for lectures and seminars. They also spend there part of free time, for example between lectures. There are many free time places in our faculty, for example buffet, library, hammocks or other sitting places with tables. It would be really helpful for students to know how much occupied these places are and how many students are there. These facts could be visualized in maps part of Virtual FIIT application [1,2].

Main part of functionality implemented in this use case can be reused for many other use cases. For ex-

ample, faculty app could be personalized according to aggregated data, summary visualizations of all students can be generated, or participation at lectures can be checked.

2 VirtualFIIT

Assumption of working prototype is Virtual FIIT application running in background, working Internet connection and Bluetooth turned on. Bluetooth is needed to estimate users' location, because proposed application use Bluetooth Low Energy beacons already deployed in faculty building. Although mobile users do not have Bluetooth always turned on nowadays, it is expected as standard to have Bluetooth turned on all the time in the future. Location of user could be estimated using Wi-Fi technology, however standalone Wi-Fi based solutions are not accurate enough. Another disadvantage is fact, that many users are using mobile Internet connection instead of faculty Wi-Fi access points.

Virtual FIIT is faculty application to help students. Actual version of application [1] was developed within Team Project subject under Dr. Alena Martonová leadership. This app use architectonic pattern client-server. Client is thin, it only displays data received from server. Client part of app use Ionic frame², which use built-in JavaScript frame AngularJS³.

Server part of application runs on NodeJS⁴ environment and uses facade design pattern. This part of application uses multiple external API, which provide information about student's courses, timetables, canteens menus and public traffic departments. To decrease overall time of client-server requests, these

* Bachelor study programme in field: Informatics

Supervisor: Dr. Alena Martonová, Institute of Informatics, Information Systems and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

¹ <https://play.google.com/store/apps/details?id=sk.stuba.fiit.virtfiit>

² <https://ionicframework.com/>

³ <https://angularjs.org/>

⁴ <https://nodejs.org/en/>

data are cached using key-value storage Redis.

3 Problems to solve

First problem to solve is localization of user's device. It is important to determine user location accurate enough. Reasons of using BLE beacons are mentioned in [1] and [2]. Firstly we used naive 2D Trilateration to solve this, and this method achieved average accuracy of 5-6 meters. We have already used another localization method, trilateration using non-linear least squares method [3]. First prototype of app with this method used with default parameters achieved accuracy of 3 meters. We expect increase in accuracy in about 0.5 - 1 meters in next prototypes. We used Gimbal Proximity Beacons and mobile phone Huawei Y7 with Android system for testing. Beacons were deployed approximately 5 meters apart and testing was static, without movement. Each measurement was repeated 50 times.

Second problem is to resolve limitations of existing faculty app. All technologies must stay preserved, so we used compatible libraries and plug-ins. We use AngularJS wrapper Ng-cordova-beacon⁵ and Cordova/PhoneGap iBeacon plugin⁶ for nearby BLE beacons scanning. For computing trilateration we are using ml-curve-fitting library⁷. This library uses Levenberg-Marquardt method of solving non-linear least squares problem [3].

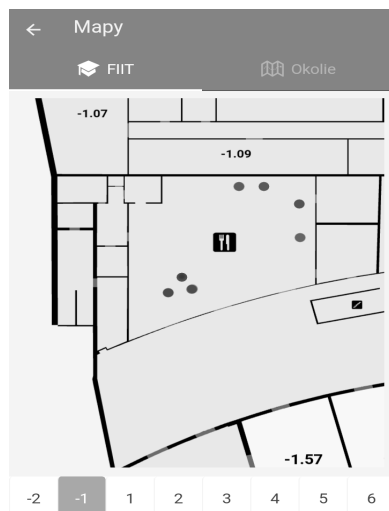


Figure 1. Visualization of free time place availability. Dots represent other students.

Third main issue is visualization of students in free time places. Mobile device users will be visualized into already existing VirtualFIIT map of building part. We are planning to display users as color dots. Figure 1 shows visualization of students in buffet free time place.

We have already implemented localization part, with limitations of existing faculty app resolved. We are testing this prototype. Visualization part is not yet implemented.

4 Conclusion

We propose innovative part of faculty application for students, which visualize how busy are free time places in our faculty. We have already tested localization part of our prototype, and it achieved accuracy of 3 meters. This idea is not binded only for VirtualFIIT application, it is possible to implement it also for another public places, another schools and faculties, shopping centers and libraries for instance. This prototype aggregates indoor movement data, which can be also used for another indoor movement evaluations and visualizations. These free time places users movement data can be interesting not only for another users, but also for building managers to manage capacities of these places.

References

- [1] Karas, J., Kučera, M., Mazán, F., Olešová, V., Příbul, D., Šoltés, F.: Beacon Indoor Localization. In: *IIT.SRC 2015: Student Research Conference*. Proceedings in Informatics and Information Technologies, Vydavatelstvo STU, 2015, pp. 441–442.
- [2] Kovarova, A.: A Building as a Context for Multidomain Information Service – Case Study Virtual FIIT. In: *9th workshop on intelligent and knowledge oriented technologies*. WIKT 2014, Nakladatelstvo STU, 2014, pp. 72–77.
- [3] Levenberg, K.: A Method for the Solution of Certain Non-Linear Problems in Least Squares. *Quarterly of Applied Mathematics*, 1944, p. 164–168.

⁵ <https://github.com/nraboy/ng-cordova-beacon>

⁶ <https://github.com/petermetz/cordova-plugin-ibeacon>

⁷ <https://www.npmjs.com/package/ml-curve-fitting>

Weather Prediction using LSTM

Jakub KAZIMÍR*

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

Abstract. The weather forecast is important problem in society because it affects the live of every person on Earth. The wrong prediction can cause destruction of property or health. Our purpose is to develop neural network, that can forecast weather by actual situation of weather conditions. For this purpose, we choose LSTM, which is a special architecture of recurrent neural network. A final forecast contains the most important part of weather and it is shown as weather situation for the next time period. In prediction for the next time period we get accurate results without long training.

1 Introduction

From the beginning of life, we are trying to predict a future. One of the oldest event, which we predict, is weather. In this case, we step forward from primitive methods of watching sky to complex computation models, which can predict weather with better accuracy then ever before.

Weather affects the life of every person on Earth and this is the main reason, why it is predicted. For someone, wrongly predicted weather means only unpleasant surprise. For others, wrong prediction means destruction of property, loss of harvests or even death. Therefore, it is important to predict the weather correctly.

Nowadays we can predict weather with various methods: numerical, synoptic and statistical. The final prediction is made by combination of these methods in common. Meteorologist predicts short-term and local weather forecast by the application of meteorological laws. This method is called synoptic and it is not used very often. Statistical method is based on statistic. This method is used for places where numerical models can not be computed for its complexity. Input is the weather prediction from numerical models and output is the weather prediction for

all other places, which are not covered by numerical method. The last method is called numerical method. It is based on solving differential equation containing a lots of parameters. These parameters contain data about temperature, pressure, wind speed, wind direction, dew point temperature, etc. These equations present processes in weather because atmosphere is only a mass of air, which behave according to physical laws.

The accuracy of numerical model is based on completed data from sensors over the world. Numerical models are not very accurate in forecasting, because data are always incomplete. This issue is solved by making more predictions with slightly changed data. Numerical method for its complexity can not be computed for all locations. This method calculates weather only for locations with big distance between them. The final prediction is average of numerical predictions with computed weather prediction for other locations by statistical method. [7]

This work is trying to predict weather with a smaller computing power. To achieve this goal, we use neural network, which is for computing power easier.

2 Related work

Related work about predicting weather with using neural network can be classified in three categories. [2, 3]

In the first category, they want to try to predict unusual weather phenomena such as monsoon rains or high temperature warnings.

In the second category, they predict one parameter in weather such as temperature, optionally also other parameters in weather such as humidity, cloud, etc.

In the third category, they predict parameters, like in the second, but they examine various types of neural networks and their settings for the best possible prediction of weather.

* Bachelor study programme in field: Informatics

Supervisor: Dr. Peter Lacko, Institute of Informatics, Information Systems and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

We found that they used data from one specific point in the world. Data for this point contain temperature or more parameters like humidity, wind speed, wind direction and pressure. The atmosphere is mass of air and it has different temperature at different locations. This behavior, along with other features of air, causes that mass of air is moving over the world and it makes fronts. This phenomenon causes, that temperature can rise or fall rapidly. Watching only one location cause, that neural network can not see these fronts, which cause mistakes in prediction.

After review of related works, we collected informations about neural networks, activators, optimizers, parameters of weather, etc. These experiences were used in the design of our neural network. So we could avoid errors, which they made.

2.1 LSTM

Long Short Term Memory network, usually called LSTM is a special type of recurrent neural network, which is able to learn long-term dependencies. It was designed to remember information for unlimited duration. Construction of module is more complicated than standard recurrent neural network. It has more layers like on figure 1. This module can pass data to another module with some new informations from the past. Information in modules can pass from the beginning to the end and it is not limited how long can piece of information stay in data. [6]

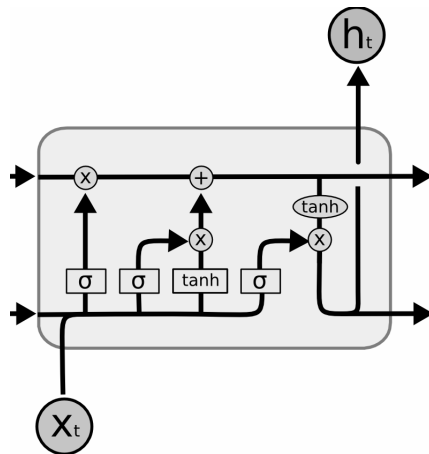


Figure 1. LSTM module [6].

In figure 1: X_t - input of module, H_t - output of module, σ, \tanh - activators, $\times, +$ - numerical operation for multiplying and summing.

LSTM is usually trains with backpropagation through time algorithm, which process error from output layers back to input layers. This algorithm can not learn very well with a lot of layers in network because

small error is multiplied by small error, which causes slow learning in deep network.

3 Method of predicting weather

The first thing what we needed is to get the weather data. One of the best dataset was from organization called European Centre for Medium-Range Weather Forecasts (ECMWF), specifically dataset is called ERA-40. Dataset contains collected meteorological data from the world. We choose only parameters, which are the most important in weather, specifically temperature, dew point temperature, pressure, wind speed and wind direction for European continent with data from 1. 9. 1957 to 31.8.2002. These data were collected every 6 hours and they were preprocessed for missing or deviated data. For our experiments we choose this dataset.

After review of problem domain, we tried few experiments to find out the best accuracy with respect to simplicity of neural network. For the first experiment we choose only one parameter to predict the weather. For a kind of neural network was used a special type of recurrent neural network, Long Short Term Memory, LSTM. This kind was chosen after review of related work and types of neural networks. Prediction was made only for one location, without data from other locations. Our first model contains data from past hours. LSTM was developed for this kind of data. Time is not present directly in data, but in their sequence. Our model of network has only 5 positions back in time from predicted value. Results were not acceptable.

For the next experiment we need more data and knowledge about weather. After review of weather and how it works, we choose parameters, which inflicted prediction of weather the most. Method is based on neural network, specifically the type of recurrent neural network called LSTM, like the first experiment. Data are collected from the special map like in figure 2. We collect parameters like temperature, dew point temperature, pressure, wind speed and wind direction. Difference between the first experiment and the current version is, that data are collected from map with specific grid represented by square where size is 13 data points vertically and 13 data points horizontally. Gap between points is 0.5 latitude and 0.5 longitude, but we collect every second point, so gap between collected points are 1 latitude in vertical and 1 longitude in horizontal way. For better idea, figure 2 show data collection we described, where:

- - places, from where we have data
- - places, from where we collect data
- ★ - place, where we can predict weather

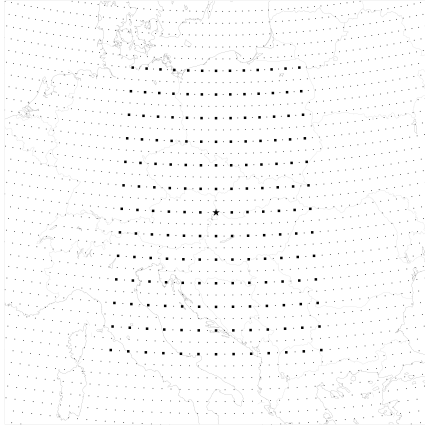


Figure 2. Plot of map where • - places where we have data, ■ - place where we collect data, ★ - place where we predict weather.

These collected data are converted to standard units in Slovakia like Celsius or Pascal for better visualization and they are transformed to scale from 0 to 1 because of the neural network. Data are at the end of preprocessing random arranged for disabling neural network to learn seasonal changes. Input layer of neural network contains 8 time-stamps, where each time-stamp contains the data of all parameters we collected. It means, the input is 3 dimensional array, where the first dimension is used for examples, the second for making sequences and the last dimension contains all transformed data from the map. We chose 8 time-stamps because it gave us the best results in terms of performance. Output layer contains one time-stamp with predicted data like temperature, dew point temperature, pressure, wind speed and wind direction. Activation function of last layer was sigmoid with optimizer function RMSprop. Error rate was evaluated with mean absolute percentage error. Neural network contains 2 LSTM layers and one Multilayer perceptron. The first layer is LSTM layer with 200 units, the second is LSTM layer too with 300 units. The last layer is Multilayer perceptron with 5 units for output where activation function is sigmoid.

This setting of data collection and neural network

was chosen after experiments with various data collection and neural network settings. This setting has the best accuracy with reasonably performance.

For some experiments we used feed forward neural network called multilayer perceptron with same data, which has not acceptable results and terrible power.

4 Evaluation

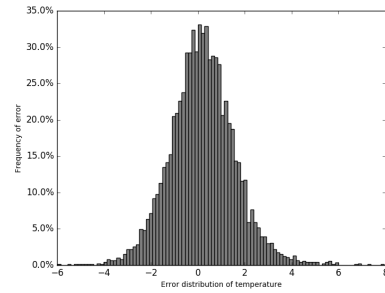


Figure 3. Error distribution of temperature in Bratislava.

Evaluation for this neural network is statistical method called "mean absolute percentage error", which is for weather forecasting very useful. Training of neural network is setting this error to minimum. For better imagination we use histogram of errors for each parameter. In figure 3, we can see, that temperature errors are situated near 0 and they are mainly in range from -2 to 2. Training of this neural network was on 80% of data, other data were used for testing and evaluating predictions. In table 1, we can see accuracy of other predicted parameters. Neural network was tested after 1000 epochs.

Our main goal was not only to make weather prediction for one place, but this model of neural network must be suitable to train on every place. For testing we choose another two places in Slovakia: Poprad and Kamenica nad Cirochou. In table 1 we can see accuracy for all places and all parameters.

Table 1. Table of mean squared error and mean absolute error of each parameter for all three places.

| | Bratislava | | Poprad | | Kamenica nad Cirochou | |
|-----------------------|--------------------|-----------------|--------------------|-----------------|-----------------------|-----------------|
| | Mean squared error | Mean abs. error | Mean squared error | Mean abs. error | Mean squared error | Mean abs. error |
| Temperature | 1.98 | 1.08 | 1.87 | 1.06 | 1.99 | 1.10 |
| Dew point temperature | 1.62 | 0.97 | 1.65 | 0.98 | 1.60 | 0.99 |
| Pressure | 3.48 | 1.42 | 3.60 | 1.47 | 3.43 | 1.42 |
| U component of wind | 2.06 | 1.09 | 1.90 | 1.08 | 1.86 | 1.07 |
| V component of wind | 2.08 | 1.11 | 1.80 | 1.05 | 2.03 | 1.10 |

Accuracy of all three places are similar and very accurate.

5 Conclusion

This work describes a prototype of neural network, which is able to predict weather at location, where was trained. This neural network works on 6-hour weather forecasting for two cities in Slovakia called Bratislava and Poprad and a small village named Kamenica nad Cirochou. Model shows, that recurrent neural network can be trained to predict weather with a good accuracy. Accuracy of model is evaluated by mean absolute percentage error, mean squared error and histogram, for better visualization. This prototype is written in Python using Keras framework on Tensorflow backend.

In the future, we should change training data with smaller periodic time. New data with better refresh time can improve accuracy of neural network.

This work can be used in real-life by making weather forecasting on places, which are not covered by numerical prediction for reasons we described.

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

References

- [1] Bernacki, M., Włodarczyk, P., mgr inż. Adam Gołda: Principles of training multi-layer neural network using backpropagation. http://home.agh.edu.pl/~vlsi/AI/backp_t_en/backprop.html, 2005.
- [2] Chalabala, M.: Weather Prediction Using Neural Networks. Master's thesis, The Technical University of Košice, Košice, 2010.
- [3] Culclasure, A.: Using Neural Networks to Provide Local Weather Forecasts. Master's thesis, Georgia Southern University, Georgia, 2013.
- [4] Hochreiter, S., Schmidhuber, J.: Long short-term memory. *Neural computation*, 1997, vol. 9, no. 8, pp. 1735–1780.
- [5] Nielsen, M.A.: *Neural Networks and Deep Learning*. Determination Press, 2015.
- [6] Olah, C.: Understanding LSTM Networks. <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>, 2015.
- [7] Pastirčák, V.: Krátkodobá predpoveď počasia na šmú. <http://www.shmu.sk>, 1999.
- [8] Shiffman, D.: *The nature of code*. 1 edn. The Nature of Code, 2012.
- [9] Česká meteorologická společnost: Meteorologický slovník výkladový a terminologický. <http://slovník.cmes.cz>, 2016.

Question Generation from Educational Text

Tomáš GÁBRŠ*

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

Abstract. One of the key factors of learning is evaluation of student's knowledge. Knowledge tests are the most common form of testing. With increasing number of studying materials in electronic form increases also interest in automatically generated tests from these materials. We focus on development application for teachers, which can help during test preparation to automatically generate questions from text. The main objective is to make the teacher's work more efficient and to increase the quality of the tests.

1 Introduction

The field of information technology is one of the most rapidly expanding disciplines. Nowadays, it is also common part of the educational process and it is important to use it as effective as possible. Regular part of educational process in schools and universities is paper exam. Exams are often created manually, so teacher has to identify target parts of text, construct the question and finally generate the whole exam in suitable format for printing.

Our goal is to simplify this process and generate questions for exam automatically based on the input text. This should be done with our proposed application. Questions can be edited by user and the final exam can be saved or exported. We obtain important information from the articles by various tools like in [1, 2, 3].

2 Proposed system

The system is divided into three components. First is external tool Stanford CoreNLP, which is used for parsing plain text and retrieving information about the structure of the sentences [4]. These information are necessary, because questions are generated on the basis of these information. With use this tool to enrich information about tokens from sentences, as well as information about whole sentences. Besides that, we are able to retrieve information about relations between the tokens. We also work with OpenIE (Open information extraction) elements, which refer

to the extraction of triples: subject, relationship and object.

There are many possibilities how to use CoreNLP. For us is the best option to use it as a standalone application, CoreNLP Server, which provides an API for our application. The most important advantage of this usage is the response time, because it keeps the annotators loaded after the initial request and there is no need for loading it separately for every request. There is also a library for Java client, which is very comfortable to use.

Second component of the application, server-side application, is responsible for business logic and question generation based on information received from CoreNLP. It provides an API for the client application and main parameter of the request is input text. There are not any restrictions for the length of the text – it can be simple sentence as well as whole paragraph. There are also optional parameters and they can be used for specifying requirements by user. It is possible to specify the need of generating distractors for the questions or the maximum number of generated questions. If input text is the article with title, it can be also sent within parameters.

After receiving request, CoreNLP API is called via provided CoreNLP client. It is configured according to the needs of the request and then API call is made. Response time of CoreNLP is based on the length of input text. When response is received, we bind provided information to a model classes. There are multiple classes, which represent various levels of information. These classes contain required attributes, which are filled in pursuance of CoreNLP response and their structure can be represented as tree with one root element, which is model class for article. This structure can be easily transformed to JSON string and passed to the other components of application.

When all information about the input text are available and set as parameters of the model classes, we focus on implementing algorithms for question generation as effective and reliable as possible. First

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

step is to identify the target parts of input text. We implemented multiple types of questions and every question has its own rules. These rules can have various complexity. All rules are based on the information provided by CoreNLP and saved within the model classes. Algorithm iterates through the tokens and every token is tested for every question rule. If rule is marked as fulfilled, all required information about question are written into the output. During iterations, there are another information written into the output, for example every token representing person included in input text is written into the list of names. These additional information can be used for generating distractors, because words with similar meaning are stored in the same list.

Our system generates fill-the-blank questions. Target part of question is blanked and student has to fill correct answer. Answer can consist from one or few words. Distractors are selected on the basis of information about the tokens. One of the key information for distractors selection is named entity tag. For example, when correct answer is person, distractors are also selected from the list of persons. After generation of all possible questions from the input text, the output is provided to the client.

Client application is available for the teachers and it provides various functionality. Teacher has to register to get access to all functionality of system. After registration and login, teacher can use system for exam generation. When he sends the request with the input text, server-side application is called to generate new exam. When the exam is completely generated, it is returned to the client application and displayed in the interface.

Exam is composed of the fill-the-blank questions. The test screen shows the original text with gaps for putting the correct answers. When the teacher selects one of these blanked parts, system shows the possible answers (correct answer and distractors if they were generated). Teacher can edit them, delete them or add new options. There is also possibility to create questions manually. Teacher can select part of text and generate new question related to marked part. After adding this question, it is equivalent to questions generated by application, so next options can be also added.

At the end, text contains multiple questions with correct answers and optional distractors and teacher can save the exam. It is inserted into database and it can be load in the future.

When exam is completed and ready for exporting, teacher can configure various output parameters. He does not have to manually create multiple variations of exam for every student, but he can simply set number of students and for every student is generated

unique exam. There is also option to specify if distractors will be exported too.

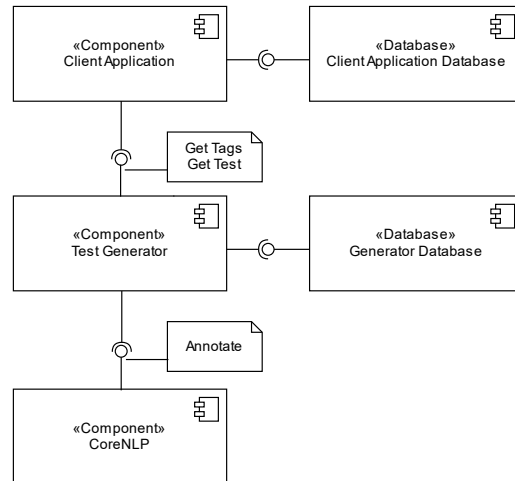


Figure 1. Component diagram of proposed question generation system.

3 Conclusions

Our application focuses to factual question generation from text. It was designed as a tool for teachers. It allows to create, save and reuse questions for tests. The tests can be also automatically evaluated.

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

References

- [1] Ali, Husam Deeb Abdullah Deeb. Automatic question generation: a syntactical approach to the sentence-to-question generation case. 2012. PhD Thesis. Lethbridge, Alta.: University of Lethbridge, Dept. of Mathematics and Computer Science, c2012.
- [2] Aharwal, Manish; MANNEM, Prashanth. Automatic gap-fill question generation from text books. In: *Proceedings of the 6th Workshop on Innovative Use of NLP for Building Educational Applications*. Association for Computational Linguistics, 2011. p. 56-64.
- [3] Beinborn, Lisa; ZESCH, Torsten; GUREVYCH, Iryna. Candidate evaluation strategies for improved difficulty prediction of language tests. In: *Proceedings of the Tenth Workshop on Innovative Use of NLP for Building Educational Applications*. 2015. p. 1-11.
- [4] Manning, Christopher D., et al. The stanford corenlp natural language processing toolkit. In: *ACL (System Demonstrations)*. 2014. p. 55-60.

TP Cup – The Best Student Team Competition Showcase at IIT.SRC 2017

Mária BIELIKOVÁ*

*Slovak University of Technology in Bratislava
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 9th 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 13 student teams presented at IIT.SRC 2017 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 satisfying 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 industry partners) 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). Moreover, last two years students can compete to win their own project theme. This year 5 teams participated and two teams get their proposed theme.

We let other 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 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. Eight 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 our 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

* Institute of Informatics, Information Systems and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

Accompanying Events - TP CUP

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 13 teams with accepted applications; all 13 teams advanced to the second stage. These 13 teams presented their projects in form of showcase at the IIT.SRC 2017.
- 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/>

TP Cup Sponsors

- | | | |
|-------------|-----------|-------------|
| – Accenture | – Softec | – robime.it |
| – Ditec | – Tempest | – Unicorn |
| – QBSW | | |

eMotion: Emotion Management and Mood Recognition Framework

Zuzana BOBOTOVÁ[†], Dávid ČERNÁK*, Veronika GONDOVÁ[†],
Tomáš MATLOVIČ[†], Tomáš PAVLOVIČ[†], Ján ŠMIHLA[†]

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

Emotions influence people's everyday decisions, choices, and relationships. During the day we have to deal with many emotions, which can be triggered by various situations or objects. Mood is rather a longer state of the person, which can last for a days or even weeks. Although the emotions are generally short-lasting events, they also affect the overall mood, which have a bigger impact on the quality of our life and health.

A good example of emotion influence is a study from Lerner et. al in [2]. Authors conducted an experiment with 200 participants. The participants were in the role of merchants and their task was to trade with economists. Negative emotions such as disgust or sadness were triggered to participants while they were buying and selling various products. The results of the experiment showed a dramatic impact of the emotions on the economic behavior. If people want to make smart decisions, knowing yourself and management of emotions is almost necessary for them.

In our work we present an application for emotion management and mood recognition. We monitor user emotions in real time using the context data from the mobile phone and the short questionnaires. Collected data are used to compute different characteristics about user, such as level of socialization or physical activity. They are presented using the charts. Furthermore, we propose a method for mood recognition using the machine learning techniques. We can detect user's negative mood and help him to deal with it by suggesting him some of his favourite activities.

There are several approaches of detecting emo-

tions automatically. We can divide them by the input they use: EEG [3], ECG, electrodermal activity, speech and voice intonation, facial expressions, body language and text [5]. Based on how they work and what try to achieve, we can divide them into three categories:

1. Applications collecting data about users that try to find the cause of their negative emotions (e.g Mood Tracker).
2. Applications collecting data about the users that try to get rid of negative emotions by recommending them to do some physical activities, listening music etc. (Pacifica, Happify).
3. Applications that do not collect any data but already expect users to have the negative emotions. They focus on dealing with negative emotions.

There are more options how to present results of the collected data to the user and also many ways how to relieve the negative mood. Therefore, we created a questionnaire, which was filled by 492 participants - 289 women and 203 men. The most of them were from the two age categories: people between 20 and 26 years (340 participants) and people between 15 and 19 years (110 participants). We asked people whether they experience negative emotions often, whether and how they would like to solve it. Figure 1 shows a result of the first two questions - people feel stress in their life and they would like to solve it. In the Figure 2 is a result of the key question which strengthen our conviction that charts are the way we should give a feedback to the user. Our product spans across all

* Master study programme in field: Software Engineering

† Master study programme in field: Information Systems

Supervisors: Peter Gašpar, Institute of Informatics, Information Systems and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

Dr. Fedor Lehocki, Institute of Robotics and Cybernetics, Faculty of Electrical Engineering and Information Technology STU in Bratislava

three earlier mentioned categories and in addition we introduce completely innovative approach in the field of the emotion management.

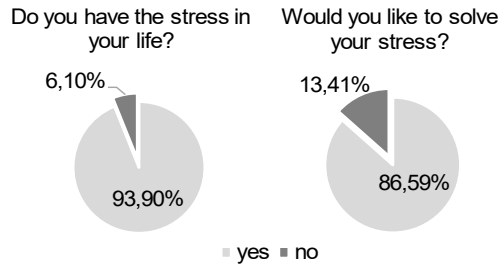


Figure 1. Answers on the questions about stress in people's life.

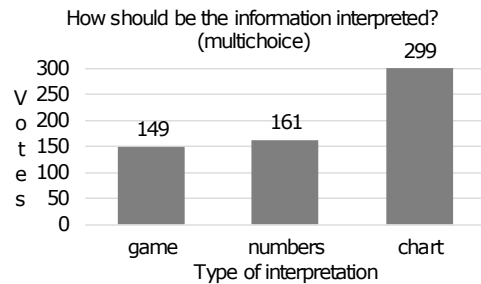


Figure 2. Answers on the question about interpretation of the information.

The main idea behind our application is inspired by the research team from the University of Cambridge. They created an application which was intended to regularly collect data about user's behavior [4]. Data included information about the user activity (through GPS and accelerometer), the socialization of user (through a number of contacts and organizing of activities), and the sentiment of communication. We also added biometric data using certificated medical sensors (such as an oximeter and an activity tracker). However, our approach is not limited to those sensors only. We proposed an architecture that enables a simple connection to any new sensor into the mobile application and an automatic processing of this data, as well. Inspired by the fact that emotions and mood are influenced by the personality traits [1] too, we started a cooperation with the team of psychologist.

Data collected by our application are used to determine four different features:

- Level of socialization (frequency of texts/calls, texts sentiment, calendar events, WiFi, BT)
- Mood (questionnaires, model)
- Physical activity (GPS, accelerometer)

- Heart activity (HRV)

User is able to explore this information using our mobile or web application.

Architecture of our framework is composed of two main components. The first part is a web application implemented in Python (Django) that is primary designed for data processing. Real time data are effectively stored using Redis. Web application also provides chart analysis for the user. The second part of the system is a mobile application for Android platform. It is used to collect mobile usage data and also allows pairing and collecting raw data from biometric sensors (AM3, PO3). In addition user can analyse calculated statistics using charts and also get recommended activity. Communication between mobile and web applications is handled by the API.

We have evaluated our approach by conducting two controlled experiments. First one was with approximately 40 participants. During the experiment they were using our mobile application and the data were labelled using the short questionnaires. Second one was only with 3 people where the biometric sensors were also used to label the context data from the mobile phone.

The main goal of our work is to provide people a tool for the management of their emotions. User's mood is detected using the proven scientific researches. Our main contribution is that the application provides personalized feedback to the user. We are also planning to use our application for medical treat by doctors or psychologists. Moreover, there are also other domains where the emotion recognition plays an important role, such as: e-learning, entertainment, marketing or law.

References

- [1] Buss, A.H., Plomin, R.: *Temperament (PLE: Emotion): Early developing personality traits*. Volume 3. Psychology Press, 2014.
- [2] Lerner, J.S., Keltner, D.: Beyond valence: Toward a model of emotion-specific influences on judgement and choice. *Cognition & Emotion*, 2000, vol. 14, no. 4, pp. 473–493.
- [3] Petrantonakis, P.C., Hadjileontiadis, L.J.: Emotion Recognition From EEG Using Higher Order Crossings. *IEEE Transactions on Information Technology in Biomedicine*, 2010, vol. 14, no. 2, pp. 186–197.
- [4] Rachuri, K.K., Musolesi, M., Mascolo, C., Rentfrow, P.J., Longworth, C., Aucinas, A.: EmotionSense: a mobile phones based adaptive platform for experimental social psychology research. In: *Proceedings of the 12th ACM international conference on Ubiquitous computing*, ACM, 2010, pp. 281–290.
- [5] Shivhare, S.N., Saritha, S.K.: Emotion Detection From Text Documents. *International Journal of Data Mining & Knowledge Management Process*, 2014, vol. 4, no. 6, p. 51.

WebX: A System for Automated Web Extraction

Ján BRECHTL*, Tomáš JUHANIAK⁺, Martin KALUŽNÍK*, Rastislav KRCHŇAVÝ⁺,
Michal KREN⁺, Martin LACEK*, Andrej VACULČIAK*

*Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
tim.16.fiit.1617@gmail.com*

Web extraction describes a process of parsing data from the web into a structured form for further use. It is an essential part of many applications across many domains, for both scientific and commercial use. Researchers are able to analyse the data produced by users of the Social Web platforms to study and predict human behaviour. In the commercial field, this process - also known as Competitive Intelligence - helps companies to analyse the activity of their competitors.

The rate of generating new data by users on Web 2.0 is increasing every year. Extracting all this data manually is no longer viable, either due to the high cost of employing domain experts or simply because of the sheer amount of the data [1]. Also, creating an ad-hoc solution in many cases might not be robust against ever-changing web pages (e.g., changes in the HTML structure) [2].

Data analysis is part of our everyday life. Each project, whether it is academic or commercial, needs a structured data set which might not always be available. In that case, the data has to be extracted from various web pages. This can be achieved by web wrappers, which transform data from unstructured sources, i.e. web pages, into structured data [3], usually using XPath, regular expressions, etc. However, creating the wrapper manually is difficult, time-consuming and usually undesirable. Many students from our faculty have to solve a similar problem while working on their projects, when they have to build their own dataset. We believe that a system, such as the one we are proposing, would help them and other researchers to collect the necessary data in an easy to use and automated manner.

Our goal is to introduce a system for automated data extraction, which significantly reduces the human effort usually required for setting up the extraction process.

Based on these facts, we have defined the goals for our project - a tool that is able to automatically

extract unstructured data located on the web and store it in a structured form for later processing, and which also provides a simple interface for managing the extraction process. This process is best characterized by a wrapper life-cycle [3]:

1. Wrapper induction. Creating a wrapper usually takes a long time and some degree of programming knowledge is required. Our browser extension provides an intuitive interface, that allows people with little technical skill to set up the wrapper.
2. Wrapper execution. The extraction script should be executed repeatedly at the precise time. Our system handles this by using a scheduler, which executes the scripts automatically according to a user-specified schedule.
3. Wrapper maintenance. Extraction script can eventually break down and has to be repaired. Extraction status can be found in the extraction logs and when a wrapper fails, it can be adjusted accordingly. Intuitive user interface allows the user to easily modify, i.e. repair the failing script.

We divided our extraction system into several logical modules (shown in Figure 1), of which each is responsible for one of a major part of the system.

User Management provides registration and login for new users, administration of the existing users and securing the data from unauthorized access.

Project Management module enables the user to create new projects, scripts, data schemas or extraction intervals. These are important for setting up a crawler. At first, the user must create a project to define some kind of domain (e. g. types of web pages), in which the extraction will be performed. Next, the user creates one or multiple scripts and defines the data schema, which is an array of key-value pairs specifying the field names and data types.

* Master degree study programme in field: Software Engineering

⁺ Master degree study programme in field: Information Systems

Supervisor: Dr. Ivan Srba, Institute of Informatics, Information Systems and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

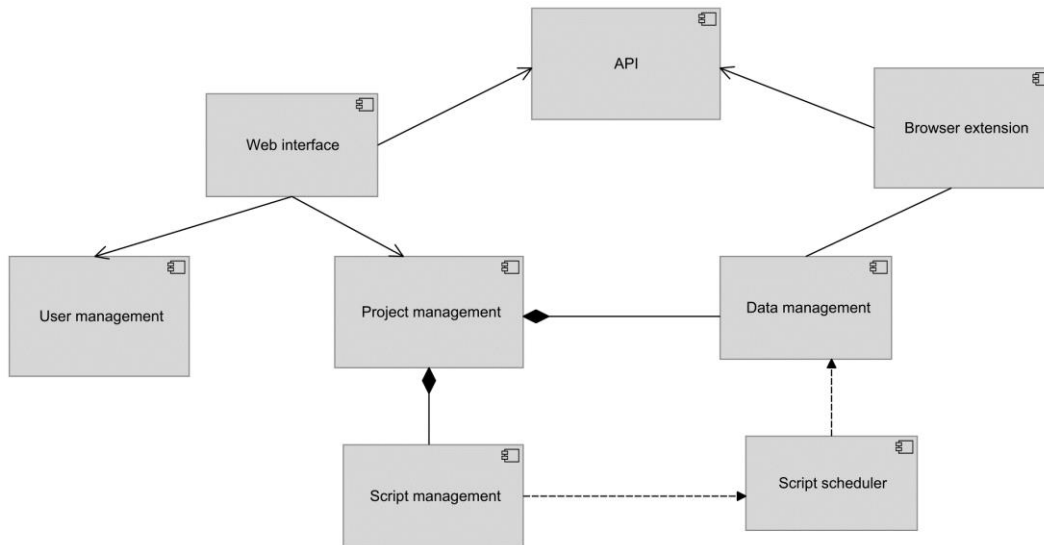


Figure 1. System modules and their relationships.

Crawler is the key part of our system. Its main purpose is to traverse the specified web pages and extract the desired data, that are specified via XPath, according to the instructions in the script. It also creates logs for each extraction, so the user can easily check, if extraction passed or not.

Browser Extension is the user-interface for annotating the data on a web page that is to be extracted. The user maps the data schema to HTML elements, that are represented by their XPath expressions (as mentioned above).

Data management displays the output of performed extractions, e.g. the extracted data structured according to the schema. This module also provides the extraction log.

Scheduler handles the planning and execution of the crawler. By setting up the frequency, scheduler executes the crawler repeatedly at the precise time. This is one of the main advantages of our solution compared to other web extraction systems, which are mainly focused on one-time extraction.

An interface for the access to the application's browser extension or data export is provided by API.

In comparison to the other available alternatives, our system offers more sophisticated functionality. Users can easily create multilevel page extraction model, which can be executed periodically and in precise time intervals defined by the user. More technically skilled users can edit the script behaviour, which is represented in a JSON format, otherwise created or edited via our browser extension.

There are options to do different types of postprocessing on the extracted data, such as extracting not just the textual value but also attributes

of a give HTML element. Other postprocessing options can help clean up the data, which is significant for data analytics. Extractions are displayed in structured, statistical view, supported by logs created during the extraction process. Data can be either downloaded as a .csv file or via our API.

References

- [1] T. Furche, G. Gottlob, G. Grasso, C. Schallhart, A. Sellers, T. Furche, G. Gottlob, G. Grasso, C. Schallhart, · A Sellers, and A. Sellers, "OXPATH: A language for scalable data extraction, automation, and crawling on the deep web," *VLDB J.*, vol. 22, pp. 47–72, 2013.
- [2] T. Furche, J. Guo, S. Maneth, and C. Schallhart, "Robust and Noise Resistant Wrapper Induction," *Proc. 2016 Int. Conf. Manag. Data - SIGMOD '16*, no. 2, pp. 773–784, 2016.
- [3] E. Ferrara, P. De Meo, G. Fiumara, and R. Baumgartner, "Web data extraction, applications and techniques: A survey," *Knowledge-Based Syst.*, vol. 70, pp. 301–323, 2014.

StoryTeller: Software Requirements Gathering and Evaluation

Jakub ONDIK*, Adam NEUPAUER*, Patrik JANUŠKA*,
Martin OLEJÁR*, Ondrej HAMARA*, Miroslav HURAJT†

*Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
storyteller-04@googlegroups.com*

Software requirements gathering process is one of the key parts of software development. For software analysts, it is important to gather as many details about customers' software requirements as possible and transform them into comprehensible and efficient form suited for programmers. The fact that customers rarely knows what exactly they want makes this process complicated. Furthermore, it is necessary to define acceptance tests along with customers, that should check fulfilment of gathered requirements [2]. As a solution to the problems in this area, we need to provide future software users with an easy-to-understand way to gather their software requirements and to define, run and provide the details and the results of defined acceptance tests.

Effective software requirements gathering and automated generation of acceptance tests could be very helpful during the whole life cycle of agile development [1]. Nowadays, there are some products such as Calabash¹ and Screener² that focus only on writing and execution of automated acceptance tests for mobile applications or web applications. However, they do not deal with the software requirements gathering process, nor the generation of automated acceptance tests from gathered requirements.

The aim of our project StoryTeller is to provide a solution for efficient requirement gathering and traceable acceptance test execution and generation in an effective and automated way. In the context of V-model of software development, we deal with the phases in the upper layer (requirements analysis and specification, acceptance testing) and provide inputs for the phases in the lower layers. In compare with process of requirements specification in which requirements are gathered during interview with

customers and user forms are drawn only on a paper or in a tool, our solution integrates these activities into one – user forms with its scenarios are created in a form of sketches and can be used by programmers directly.

To achieve this goal, we provide the method for software requirement gathering based on visual sketching and commenting of user scenario screens. The method consists of the following steps:

1. creation of user screen forms (sketches),
2. automatic generation of specific scenarios and definition of custom constraints for scenario steps,
3. automatic generation of acceptance tests.

A user – usually a software analyst, can create projects, which consist of user scenario sketches, acceptance tests and support chat rooms for involved team members. After creation of the project, users can create various user screen forms with predefined visual tools. These tools consist of basic user interface form elements found in various wireframe sketching applications and include elements such as textbox, label, button or checkbox. After adding selected element to canvas, users can modify their visual properties. These properties are based on standard CSS properties, such as text font, colour and size. Users also have to specify a name of added element, which is in cases of text box and label same as their content. This name is used as a placeholder in scenario step and denotes an interactive element used in acceptance test. In addition to standard visual element properties, there is a system property used for system messages to show state messages, such as a success and a failure.

* Master degree study programme in field: Software Engineering

† Master degree study programme in field: Information Systems

Supervisor: Dr. Karol Rástočný, Institute of Informatics, Information Systems and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

¹ <http://calaba.sh/>

² <https://screener.io/>

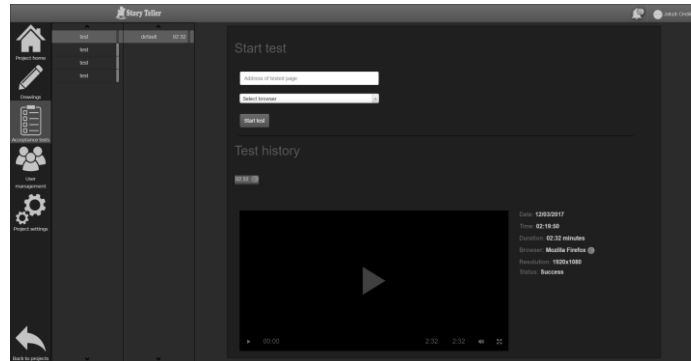


Figure 1. Screen shown after test execution.

Specific scenario related to a sketch is being automatically created at the time of an element addition to the sketch. After the element addition to the sketch, a scenario step specific to the added element is added at the end of the scenario. Content of the step consists of a basic text form and domain specific language form (DSL) used for automatic acceptance test generation. As a DSL we chose Gherkin³, a business readable DSL, used to describe user scenarios in a form of features.

Gherkin allows us to specify custom preconditions, postconditions, actions and results in a form of human readable text. This allows us to define various constraints specific to each visual element, such as an email validation or a regular expression condition for text box element, that enhance acceptance tests abilities. Gherkin also allows us to map system messages, either generated by exceptions thrown during tests or user defined messages in a visual element marked as element for system messages described above. Order of automatically added scenario steps can be freely modified. StoryTeller also supports creation of multiple scenarios per sketch and branching of scenario based on user specified conditions. This allows users to create and execute multiple tests with different data without duplicating the sketch of user interface.

Automatic generation of acceptance tests is based on generated feature files in Gherkin notation and sketches transformed to HTML in case of web applications. This feature file can consist of multiple defined scenarios and their steps. However, each scenario needs to have defined at least one action and at least one postcondition or one result. If these requirements are met, an acceptance test can be executed. As an execution tool we chose Behave⁴, a

python based behavioural driven development tool, which supports Gherkin notation. This allows us to create modular test execution module for various types of tested applications, e.g. desktop, smartphone or web applications.

The first type of applications we decided to focus on are web applications. To support multiple types and version of web browsers, we chose to use Selenium⁵, which allows us to easily manipulate HTML DOM elements. After defining and executing acceptance tests, users can preview the execution in a form of simple videos or to schedule automatic execution of all defined tests at specified time, which is illustrated on Figure 1. Results of tests are published to project dashboard by default. Additionally, users can setup an integration with collaborative tools such as Slack⁶ and receive test result in a form of notifications.

To evaluate the properties of our solution, we plan to make our project accessible to one of the local incubators in order to obtain feedback for the described solution.

References

- [1] Haugset, B., Stalhane, T.: Automated acceptance testing as an agile requirements engineering practice. In: System Science (HICSS), 45th Hawaii International Conference on System Sciences. IEEE, (2012), pp. 5289–5298.
- [2] Melink, G., Maurer, F.: Multiple perspectives on executable acceptance test-driven development. In: International Conference on Extreme Programming and Agile Processes in Software Engineering. Springer Berlin Heidelberg, (2007), pp. 245–249.

³ <https://github.com/cucumber/cucumber/wiki/Gherkin>

⁴ <http://pythonhosted.org/pytest/>

⁵ <http://www.seleniumhq.org/>

⁶ <https://slack.com/>

EduSim – Multi-domain Educational Simulations

Adam BLAŠKO*, Martin CVIČELA⁺, Ivan GULIS⁺, Tomáš LIŠČÁK⁺,
Branislav MAKAN⁺, Marek MATULA⁺

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

It is well known people remember things better through practice, experiments and visual information [1]. There are many tools used in education nowadays which try to improve the quality of the lectures, e.g. Alef on FIIT STU [2]. It is seldom that one application is used for multiple lectures and even less so for multiple courses. Students and teachers often need to learn to use every such application correctly before they can use it to learn or to teach. Sometimes, this learning process to use such educational application can be too long to be worth the return value we get from it. It can also be troublesome to keep multiple applications updated. This is another reason for a generic simulation creator which would be easy to use, accessible and usable for multiple domains.

Our goal is to create an application for creating educational simulations. This application needs to be generic enough to be used in multiple domains, e.g. physics, biology, geology, etc. To make such an application generic enough, a modular approach is required. Therefore, almost everything in our application is an independent module which provides a piece of functionality. Some of the modules rely on each other due to logical implications, e.g. you cannot delete an item before you select it. Besides modularity which allows genericity, the main strength of our application is the capability to export created simulations in the HTML format. This will allow teachers to create interactive simulations and spread them among their students who can run them simply in their internet browsers without requiring any dependencies.

There are many advantages to educational computer simulation. Schools usually need to buy various requirements for various courses. These all cost money and there is almost never enough of them for every student to use. In a computer simulation, every component can be duplicated almost

indefinitely and every student can experiment with them. The only requirement is computer access for every student. Given the popularity and spread of smartphones, tablets and computers today, this is easier to accomplish than buying requirements for every student for every course.

EduSim application is built upon Unity Game Engine which provides us with a strong base with plenty features already built in. It also supports both 2D and 3D graphics, which means it's suitable for both 2D and 3D modules which can be added in the future.

Almost *everything* in our software is modular. Two main module groups are the following:

- General editor modules
- Domain modules

General editor modules of the software represent the main logic behind the editor for creating simulations. They provide the following functionality:

- User interface
- Creating, saving and loading projects files
- User interaction
 - Creating object instances
 - Selecting, deleting, moving, rotating, and duplicating objects
 - Connecting objects
 - Setting values to the objects, e.g. battery voltage
 - Workplace navigation
 - Pausing and running the simulation
- HTML5 export of the simulation
- Localization
- Keyboard shortcuts

* Master degree study programme in field: Software engineering

⁺ Master degree study programme in field: Information Systems

Supervisor: Marek Lóderer, Institute of Informatics, Information Systems and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava
External consultants: Michal Mjartan, Atos, Vladimír Broniš, Atos

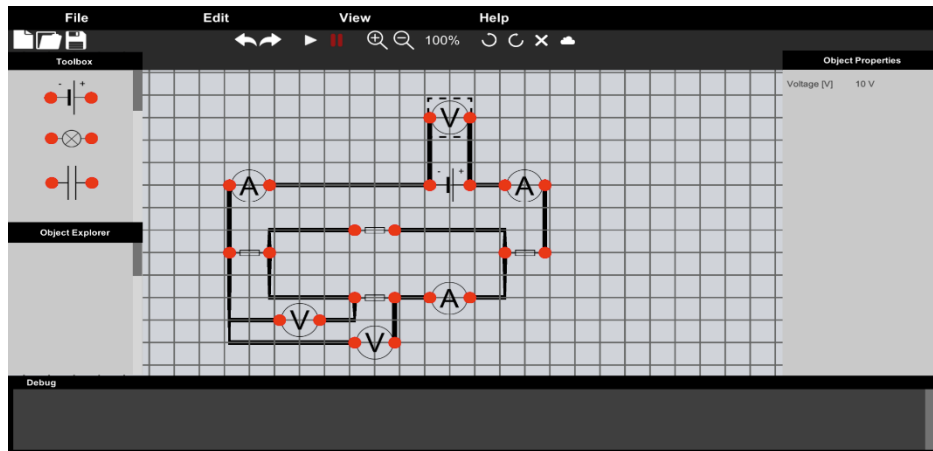


Figure 1. Screenshot of running simulation in our application.

The most interesting and the unique part of the core modules, is the capability to export any simulation created with the software to the HTML5 which can be hosted on a web page or shared among the students to experiment with. Paired with its multi-domain support, this feature becomes very powerful as it enables these simulations to be highly accessible on the web.

In the future, these exports can also be used as tests and exams with automatic evaluation systems built in, which would save teacher's time, ensure fair test evaluation and remove human error. Students could be given a working simulation with pre-set components and their goal would be to calculate some aspect of the simulated system. They could also be given a half-working simulation they need to complete for it to function properly. In a more static environment, such as hierarchical structure of animals or a geographic map, a whole different set of tasks could be implemented. These tasks would bring the educational tests closer to puzzles and games, which would make learning process more fun and attractive. These are, however, out of agreed scope of our course.

The second large module group are the domain modules. Domain modules are packages which can be loaded in the core software and thus provide all graphical and logical components needed to create working simulations for the given domain, i.e. images of the components and scripts to handle the simulations for the module. As these modules are separated from the general module group, they can be loaded as needed without affecting the stability of the program.

In this course, we worked on electronics module for creating electric circuit simulations. This module allows users to create electric circuits using predefined components and arbitrary set their parameters.

For the logical part of our electric circuit module we used freely available source code written in C#¹. This library was wrapped with our own logical part which connects it with Unity which serves as the front-end. We also use our wrapper to save and load the state of the project to and from an external file.

In the future, EduSim can be upgraded by adding new educational modules and adding support for student testing with automatic evaluation. Since the application was made with modularity in mind, many more functionalities we didn't think of can be easily added without impairing existing ones.

References

- [1] Ronen, M.; Eliahu, M.. Simulation—a bridge between theory and reality: the case of electric circuits. *Journal of computer assisted learning*, 2000, 16.1: 14-26.
- [2] Šimko, M.; Barla, M.; Bielíková, M.. ALEF: A framework for adaptive web-based learning 2.0. In: *Key Competencies in the Knowledge Society*. Springer Berlin Heidelberg, 2010. p. 367-378.

¹ The source code for the electric circuit is available at: <https://github.com/Mervill/SharpCircuit>

Advanced Repeatable Service Module for High-altitude Balloon

Ján PÁNIS*, Jakub FINDURA*, Maroš FRKÁŇ*, Tomáš URBAN*,
Martin ORAVSKÝ*, Dominik PISAROVICH*

*Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
timovyprojekt10@gmail.com*

Our research falls into the area of space engineering. In particular, the paper focuses on the problem of applications for stratospheric balloon, its higher reliability and repeatability. Many studies have already been devoted to the stratospheric balloons. The authors show many ways how to insure the security and provide sufficient results in data acquisition. Our approach is based on the redundant communication, secure data acquisition and easy repeatability. The applicability of the proposed approach is demonstrated on building a reliable service module for student's experiments in stratosphere environment, which is the main contribution of our work.

Area of stratospheric balloon flights is popular nowadays and also a bit specific in some way [1]. Alongside suborbital rockets, sounding rockets and CubeSats, stratospheric balloons are considered as first steps in education of space engineers. Slovakia is in process of achieving full membership in the European Space Agency (ESA) in these days. ESA coordinates the financial and intellectual resources of its members, therefore it is important to raise awareness of possible involvement in space research and to raise new generation of space engineers during the integration.

At this moment, National Aeronautics and Space Administration (NASA) presents the state-of-the-art approach in this area in collaboration with National Science Foundation (NSF). They realized many long-term experiments [2], which explain a lot of important questions about universe, atmosphere, the Sun and the space environment. These balloons provide many benefits, they:

- can be released from locations worldwide to support scientific needs,
- can be prepared for flight in as little as six months,

- offer a low-cost way of conducting science research,
- provide a stable platform for longer flight durations.

Among all the advantages in stratospheric balloons, the best one is repeatability and reliability. Both these features area are acquired by high redundancy in data acquisition, communication and security.

The goal of our one year work is to create reliable service module suitable for frequent and repeatable releasing balloons with experiments for graduate and undergraduate final works or any other academic research.

Our approach is based on existing solutions used for commercial flights into stratosphere [3]. This led us to design portable service module for control the flight, tracking and data acquisition (see Figure 1).

We propose tracking using two GPS sensors to provide higher reliability by space redundancy. Measured data is transmitted two-way via UHF radio and GSM communication network. Arduino GPS sensor provides us with actual information about position of our balloon. However, this information is not as precise as we need. Therefore, for tracking after the landing we use radio communication inspired by Amateur Radio Direction Finding (ARDF). For this purpose, we use two RF transceiver 433 MHz modules with directional antenna. It helps us to easily find the balloon after landing on the ground in less hardly accessible area.

Approach to the data acquisition is based on Global System for Mobile Communications (GSM) and storing data on-board in flash memory card. At first, we need to acquire required information. In our solution, we concentrate on acquiring information about temperature, time, speed and altitude. In case of temperature, we analysed many sensors and tried to fit it to our domain, but as we need to endure lower

* Master degree study programme in field: Software Engineering
Supervisor: Michal Valiček, Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

temperatures, we can not use modern digital sensors. Hence we use resistance temperature detector (RTD) Proffuse PT100-1020. It provides us the scale from -70°C to 500°C. As we expect the lowest possible temperature -60°C, this solution is more suitable for us. During data acquisition, the measured values are tagged with time, therefore real-time clock (RTC) is implemented into. RTC functionality is based on control chip, crystal and back-up battery or supercapacitor.

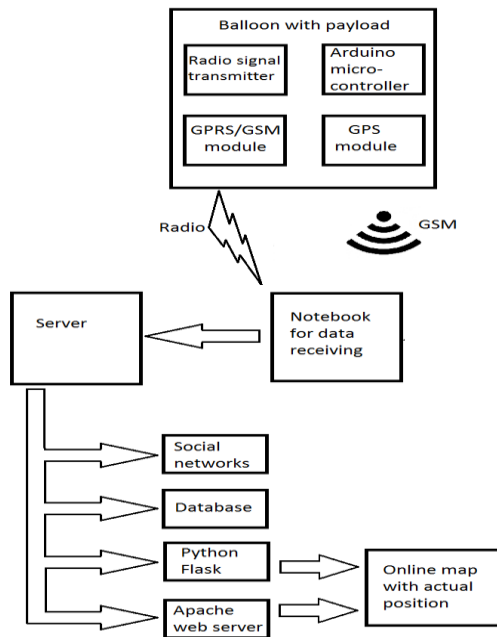


Figure 1. Architecture.

Used Arduino Mega microcontroller provides through 16Mhz crystal in cooperation with processor ATmega 328p results precise enough. Oscillator with crystal is stable and could be used as a trigger for obtaining required data. The data are acquired in exact periods. As in case of measuring speed and altitude, we propose usage of GPS Logger shield from Sparkfun company. It contains slot and control unit for the flash memory card so the redundancy to sending data is provided. However, this shield cannot measure GPS coordinates and write data to flash memory card at the same time. Due to long GPS module reading time we cannot solve this problem with any software lock yet, so we decided to add another memory module.

Our system is also capable to propagate itself during the flight. It stores all received information and share them on social networks periodically. The system also generates a webpage with actual position of our balloon in real time. In architecture of our

server are used principles of multilayer architecture (see Figure 2).

Each layer has a defined functionality and has isolated options in communication just for the nearest layers.

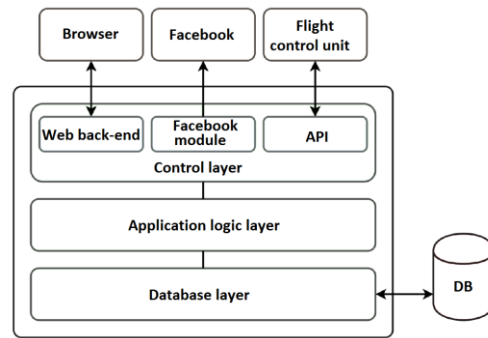


Figure 2. Proposed multilayer server architecture.

In order to obtain better results as higher vertical speed, we are working on releasing the gas from the balloon continually. It allows us to fill the balloon with more gas on the ground and ascend with more speed while not bursting out our balloon so early.

Our approach to stratospheric balloons may be useful mainly from the aspect of frequently releasing balloons for academic experiments mostly due to its flexibility, adaptability and low cost. We have concentrated on two features in our work. The first is redundancy and reliability of prospected solution. The second is to easily repeat and adapt experiments in stratosphere environment.

We are going to verify our solution by experimentally extracting maximum period of every task and using theorems verify ability to finish all the tasks before deadline even in the worst scenario. It is also possible to enhance our proposed methods with controlled fall, which we are considering for our future work.

Acknowledgements: This work was supported by the Grant of Tatra banka Foundation.

References

- [1] Akerman, D.: PIE1 – Raspberry Pi Sends Live Images from Near Space, July 16, 2012, [Online; accessed October 7, 2016]. Available at: <http://www.daveakerman.com/?p=592>.
- [2] Johnson, M.: Scientific Balloons [Online; accessed December 12, 2016]. Available at: <https://www.csbf.nasa.gov/balloons.html>
- [3] Zavodsky, O.: Balón Universum 2 zhodnotenie, May 5, 2011. [Online; accessed October 15, 2016]. Available at: <http://svetelektro.com/clanky/balon-universum-2-zhodnotenie-399.html>.

SECMON - Security Events Collerator Done Right!

Norbert DANIŠIK[†], Matej GURÁN[†], Matúš JURIKA[†], Štefan KADLIC[†],
Roland LANG^{*}, Milan POLIAK[†], Matej PUK[†]

*Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
talented-otters@googlegroups.com*

Security of computer systems is becoming more complex as Internet grows and amount of security threats are increasing. Nowadays, having antivirus software is not sufficient from the perspective of cyber security, so adding additional layers of defense is recommended and often required. Mainly, we are speaking about monitoring and preventing security threats, detecting security incidents using tools supporting “Defense in Depth” methodology. Early identification of attacks provides an advantage while protecting your personal and organisation networks. Even if attack is successful, these software solutions can help system and network administrators to identify cause of it. Critical part of “Defense in Depth” is logging activities and monitoring security and this is done by SIEM (Security information and event management) solutions. These are based on a principle, that all security related data produced by various devices, are easier to analyse, when viewing from the single point of view. For spotting advanced threats and security patterns is crucial to implement correlation mechanism. Activity of these solutions is based on real-time analysis of security alerts generated by network hardware and software. There are various solutions on the market, for example ArcSight, AlienVault, etc., but their problem is often high price, complexity in required knowledge and proprietary technology. However, there are a lot of potential users, who seek advanced monitoring and alerting solutions, but they lack resources.

This was motivation for us - to develop SIEM system - SECMon, which will satisfy needs of home and SMB users. The main advantages, which our product leans on, are open-source technology core, price availability, interactive and modular user interface. To meet these needs, as the core of our system, we have

chosen a free correlation tool SEC (Simple Event Correlator) [2, 4]. It is command-line utility intended to analyse log files generated by various network nodes. RSYSLOG is used to collect and send all the logs to the respective central device, on which is running SEC instance that processes all the information. One of the challenges is to prepare a normalization middleware, which is converting all types of logs into an agreed format - CEF (Common Event Format). Normalization of logs improves scalability of the solution into new types of monitored nodes, which can have different formats of logs that are converted into CEF format. This format was chosen due to its wide usage among different systems (mainly based on Linux kernel) [1, 6].

SECMon as a software is far more than just a graphical user interface for its core - SEC instance. In order to provide comfortable and effective solution, we added different layers into our architecture, that is shown in the figure 1. As mentioned above, all the acquired logs with network and usage activity are required to be normalized into CEF format. Then SEC processes these logs and finds patterns of attacks based on different rules, which are managed by system administrator. After successful identification of correlated security event, it is recorded into database. As long as security monitoring and event correlation is obliged to be real-time, usage of distributed database is critical in this solution. It helps to process streamed data at very short response time and though provides opportunity to fast reaction to malicious events.

With increasing amounts of data to be processed, multiple instances of SEC run parallelly. Synchronizing these instances is able thanks to databases, which are middle element between SEC and model of web

* Master study programme in field: Information Systems

† Master study programme in field: Software Engineering

Supervisor: Dr. Ján Laštinec, Institute of Computer Engineering and Applied Informatics, Faculty of Informatics and Information Technologies STU in Bratislava

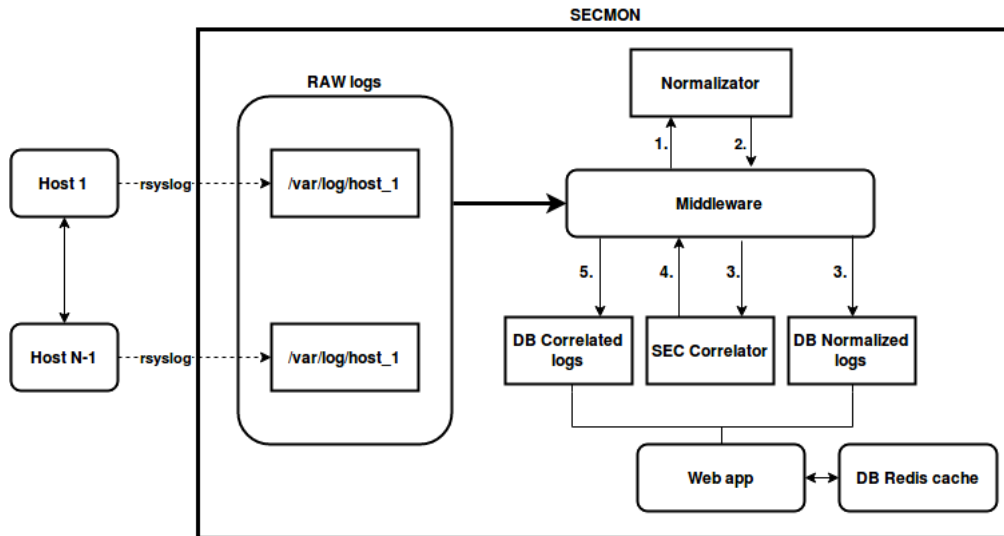


Figure 1. Architecture of SECMon.

backend. The other interesting feature of SECMon is ability to monitor number of connections from each IP address, TCP ports, etc. This provides strong analytic insight into ongoing intersections into monitored systems. Due to these facts, we are using Redis as storage to effectively process incrementing the needed values based on their keys (e.g. IP addresses or ports) [3, 5].

SECMon is designed to be a web application with intuitive user interface based on Material design. Every organization seeks for different types of monitoring activities, so the dashboard, which is in control of user, is fully modular and its elements are draggable, resizable and able to be persistent in their settings. Application has set up access policies in order to segregate users into groups with privileges.

Our aim is to build a security tool which is based on open source technologies and focuses on homes, small and middle sized companies which seek advanced security in their organizational networks. Usually, SIEM systems are complex and require a lot of resources. In terms of innovation, our solution is designed to be lightweight and scalable with interactive environment that does not expect only experienced users. It comes with several predefined monitoring rules commonly used by security professionals, which we consulted the process of development with. Although, for advanced users, there is a possibility of full adjustment of SEC rules in order to accommodate SECMon to their desired settings. Other innovative approach is the normalization of data which provides

potential for effective correlation of security events from multiple types of devices. In terms of extensibility, new normalization settings could be implemented in moments. In conclusion, SECMon is meant to be configurable, adjustable and adaptable to all kinds of networks supported by user interface with various personalized dashboards and statistics. Planned upgrade in the future is anomaly detection, which is based on machine learning model that could detect the potential threat even before it eventually happens and notifies security personnel.

References

- [1] Lang, D.: Using SECCorrelator, Simple Event. *Sysadmin*, 2012, pp. 38–43.
- [2] Myers, J., Grimaila, M.R., Mills, R.F.: Log-Based Distributed Security Event Detection Using Simple Event Correlator. In: *2011 44th Hawaii International Conference on System Sciences*, 2011, pp. 1–7.
- [3] Vaarandi, B.R., Grimaila, M.R.: Security Event Processing with Simple Event Correlator. *ISSA Journal*, 2012, vol. 10, no. 8, pp. 30–37.
- [4] Vaarandi, R.: SEC documentation pages, accessed: 15.02.2017, <https://simple-evcorr.github.io>.
- [5] Vaarandi, R., Blumbergs, B., Çalışkan, E.: Simple Event Correlator - Best Practices for Creating Scalable Configurations. Volume 2015., 2015.
- [6] Vaarandi, R., Vaarandi, R.: SEC – a Lightweight Event Correlation Tool SEC – a Lightweight Event Correlation Tool. *Proceedings of the 2002 IEEE Workshop on IP Operations and Management*, 2002.

BeaCode: Indoor Navigation Providing Position Related Information

Peter AUGUSTÍN*, Veronika BALÁŽOVÁ*, Marek BRUCHATÝ⁺, Juraj FLAMÍK*,
Ondrej KIPILA*, Sandra KOSTOVA*, Andrej ŽLNKA⁺

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

Visitors attending various exhibitions are forced to deal with different kinds of problems, most of which are directly related to information, information availability, discoverability, quality and quantity. Information about various exhibitions and related exhibits available for visitors have very limited scope. This is the consequence of the ways in which is this information distributed, as well as the media used for the distribution (prints, audio-video recordings, infographics etc.). Prints are restricted by their physical dimensions while audio-video recordings are mostly limited by the time it takes to present the selected set of information. Another problem is the ability of visitors to discover the existence and location of information they are interested in.

Localization of the exhibits and information about them can be very difficult in crowded conditions. These localization problems are mostly caused by bad placement of signs, information tables and other markings in the environment, or by selecting inappropriate formats for these navigation points. The quantity of information presented to the visitor can have a significant impact on his/her overall experience of the exhibition. Not all exhibits are interesting for individual visitors and can be viewed as unwanted and time wasting. This is the reason that information overflow is a big problem.

Many exhibition organizers at the moment do not have effective tools that can be used to collect feedback and other valuable information from exhibit visitors. Information is mostly collected through paper forms or through electronic forms (sent only by a fraction of visitors). Furthermore, the feedback collection process that involves information about the number of attendees at the exhibition or number of

attendees at the location of individual exhibits is complicated (e.g. by tracking through Wi-Fi, which has limited accuracy) – there is not a single solution that is able to provide all of these features in a single, simple and effective package.

Unlike solutions provided by other companies e.g. Infsoft¹ or Meridian², we provide the experience of real time information discovery and sharing, from which benefits both the visitors and exhibition organizers: The application Indoo.rs³ offers only basics about the exhibition without the information about individual exhibits. BeaCode tries to identify individual visitors and their needs and preferences and use this knowledge to present all exhibitions in a new and interesting way - multimedia.

The solution to all these problems is our mobile application BeaCode. The BeaCode app is focused on indoor navigation based on the Bluetooth Low Energy (BLE) Beacon technology. The actual product is a composite of two parts - the mobile application designed for visitors and the administrative web application designed for the organizers. Mobile applications are implemented for both most used mobile platforms - iOS and Android. The web application is based on modern front-end platform – AngularJS.

By utilizing BLE Beacon technology, it is able to pinpoint visitors' location on the map of the exhibition (present in our app), relative to other objects - different exhibits. With this utilization, it is able to navigate the visitor from one exhibit to another, filtering out unwanted stops, minimizing the distance the visitor need to overcome. Based on his/her movement relative to other points of interest around, the application will present entirely different

* Master degree study programme in field: Information Systems

⁺ Master degree study programme in field: Software Engineering

Supervisor: Dr. Alena Martonová, Institute of Informatics, Information Systems and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

¹ <https://www.infsoft.com/>

² <http://meridianapps.com/>

³ <https://indoo.rs>

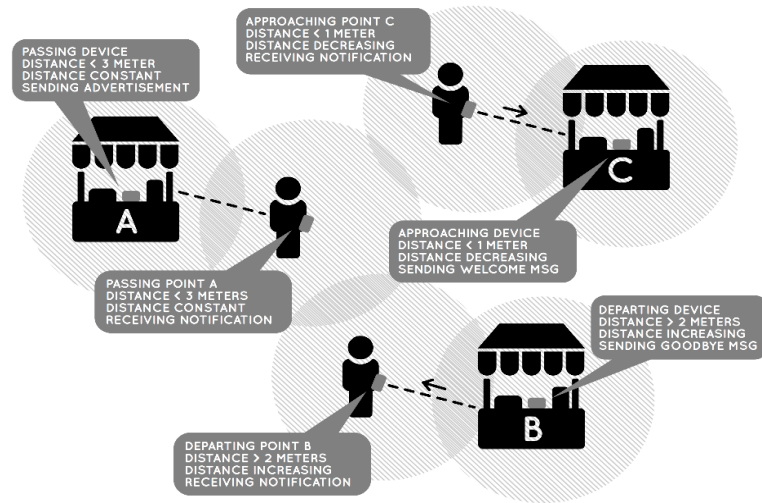


Figure 1. Illustrative situation of three different cases of position related information sharing using beacon enabled devices.

information. This approach of information providing is very dynamic and responsive and can effectively bridge the gap between the visitor and the exhibit he/she is interested in. Moreover, this information is fully customized by organizers and provides virtually endless information in many different formats. These Beacode features improve visitor experience without forcing him to scan any code, make photos or type any input to phone. BeaCode automatically shows actual information based on actual time and visitor's location. And this is done with high accuracy thanks to BLE (unlike to Wi-Fi, which have lower accuracy). Thanks to the position context which is constantly updated, our application can provide information not only when the visitor is near the product but also when the user is getting closer or moving away from the product (see Figure 1).

Mobile application provides a dashboard of all exhibitions that are scheduled to take place near the location where the user is located. This dashboard prefers exhibitions that are similar to the preferences of the user and filters out those that do not match these preferences. Preferences are expressed in form of hashtags keywords, and are attached to user profile. Other exhibitions can be searched for by a search engine built into the app that can look for any feature exhibition. All exhibitions are viewed in detail, exposing detail information about the exhibition itself as well as its exhibits. Exhibits can be also viewed in detail, including descriptions and photographs (later videos). In addition, this information is available for the user in advance before the exhibition. The second part of BeaCode is the web administration application designed for organizers. Organizers are able to create new exhibitions, add different exhibits and prepare guidelines for visitors. In the future, the web interface will be also used for statistical evaluation of collected

data from past exhibitions including summaries about the number of attendees at the exhibition and number of attendees at individual exhibits. Published and filled feedback forms will be also viewed here. Organizers are using this valuable information to make their exhibits more effective and more profitable.

The ultimate vision of our product is to create a useful ecosystem for visitors as well as exhibition organizers, crucial for securing a flawless experience for visitors and also the highest possible ratings and profits for exhibition organizers.

Acknowledgement: This work was partially supported by the Science and Technology Assistance Agency under the contract No. VEGA 1/0752/14.

References

- [1] Deng, Z., Yu, Y., Yuan, X., Wan, N., Yang, L.: *Situation and development tendency of indoor positioning*. In: China Communications, vol. 10, no. 3, (2013), pp. 42-55.
- [2] Sharhan, S. M. H., Zickau, S.: *Indoor mapping for location-based policy tooling using Bluetooth Low Energy beacons*. In: 2015 IEEE 11th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), Abu Dhabi, (2015), pp. 28-36.
- [3] Gast M. S.: *Building applications with iBeacons*. O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472. (2014)
- [4] Wetchakorn, T., Prompoon, N.: *Method for mobile user interface design patterns creation for iOS platform*. In: 2015 12th International Joint Conference on Computer Science and Software Engineering (JCSSE), Songkhla, (2015), pp. 150-155.

SmartStore - Automated Optimization of Warehouse

Matúš SALÁT*, Mária DRAGÚŇOVÁ*, Matej ČERVENKA*,
Martin ŠIDLO*, Ondrej KUDLÁČ*, Lukáš ŠIMEK*

*Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
tim3fiit@googlegroups.com*

Ordering products via internet has become a natural thing of our everyday lives. Therefore, there are many e-shops, which compete for customers. Delivery time is one of important factors, which can determine, if a customer orders a product from desired e-shop, or rather finds it in an e-shop of competitor. In our work, we focus on shortening of the delivery time. This should lead to minimizing the chance of choosing another e-shop, because of shorter delivery time and an improvement of customers' satisfaction.

Optimization of a warehouse capacity and orders' processing is almost a necessity for large e-shops – an example is well-known company Amazon, which is also an innovator of human-machine collaboration in the warehouse [2]. However, there are warehouses of medium-sized e-shops in Slovakia, which are still organized and managed in an old-fashioned way using only human experiences and simple computations.

We propose a system named SmartStore, that can be used as a third-party software for an e-shop warehouse. The shortening of products' delivery time is executed on two levels. Firstly, we minimize absence of a products in the warehouse by sales prediction using module SmartOrder. Secondly, we minimize time of order processing with SmartCollect module, which is used to find the shortest path for products collection in the warehouse. To evaluate our solution, we collaborate with the biggest Slovak e-shop - Martinus¹.

The main goal of SmartOrder is to optimize warehouse stock count, because warehouse capacity is limited and products should be managed wisely. Based on data from the real sales, we have created a method that predicts an amount of future orders. We ensure an optimal stock count of every product by covering the demand between deliveries of the new supplies.

We developed our own algorithm based on frequency of sales per a time period. We have

experimented with time periods of various length. Currently, we use the time period of 3 months. We use regression to predict sales of each product for the period until next delivery. A delivery time of each product is taken into consideration when prediction is calculated. Prediction of a product includes:

- *Date of order* – Employee should order products at the date specified in this parameter.
- *Number of pieces* – This parameter specifies how many pieces should be ordered at *Date of order*.
- *Date of running out of stock* – Predicted date of sellout of all specified products.

Module SmartOrder serves as a recommendation for an employee of a warehouse, not as a tool for automatic product ordering. Owing this functionality, an employee has actual information about product ordering every day. All the predictions are accessible through user interface of the SmartStore web application.

After successful ordering, customers' order must be packed and send from the warehouse. A few companies have their own robots in their warehouses to collect and prepare the shipment. But a vast majority of warehouses employ humans to do it manually. These workers are given a list of ordered items with coordinates of their positions in the warehouse and they have to find and collect each item in the list. Usually workers choose their collect path by themselves which is often unnecessarily long and more time-consuming [1]. With bigger list of items, problem of optimal path searching is more difficult for worker. This is causing not just longer collecting path, but it also wastes money for unnecessary work.

This kind of problem is specific for the graph theory. There are many algorithms, that can be used to find the best possible path. They can be categorized as follows [1]:

* Master degree study programme in field: Information Systems

Supervisor: Peter Krátky, Institute of Informatics, Information Systems and Software Engineering Faculty of Informatics and Information Technologies STU in Bratislava

¹ <https://www.martinus.sk/>

- TSP (traveling salesperson problem)
- The shortest path problem

Both approaches bring several algorithms, that have their specific features. The shortest path algorithms offer the ability to find the best possible path. These approaches deliver the best path under any circumstances, but with more complex graph, it would take too long, to bring the desired results. The TSP presents algorithms with sufficient accuracy in polynomial time [1].

For the best results, we have designed our product SmartCollect in partnership with Martinus. We use information about physical layout in the warehouse to simulate the real conditions in our application. We have proposed various methods of solving the shortest path problem with some modifications for better performance.

SmartCollect provides a service for finding collect path using our algorithm. A partner e-shop system sends a request with an unordered list of items and our service responds with ordered list with the best possible path for worker. For the average size of our partner's order, we are able to deliver the best path in less than 50 ms.

Workers can collect items without noticing that they are the most effective. Our algorithm prefers to create collect path beginning with the furthest items. In this way, the items, which are close to the final location point, are not unnecessarily carried all the time. Workers can also use mobile app, to visualize the collect path in the warehouse. The only thing they have to do, is scan the barcode of the order. For better user experience, visualized collect path on store map contains colorized shelves with desired books.

We have also developed a passive functionality that suggests reorganization of warehouse based on product sales. The feature monitors the frequency of sales pursuant to which suggests better placement for items. Products with high frequency of sales have a high priority, while products with low frequency of sales have a low priority. Warehouse is divided into three sections. First, books with high-priority, then books with medium priority and the last one is for books with low priority. This feature also saves time and effort of employees in order processing.

We evaluate our methods' results using one year old dataset of real sales to achieve optimal use of stock. Before final deployment, we test all our improvements to prevent the deterioration of the performance.

To evaluate our improvement, we compare stock counts of products with and without our predictions. In the Figure 1, we depict results of SmartOrder prediction, reaching 55% of precisely stocked products and only 7% of overloaded products.

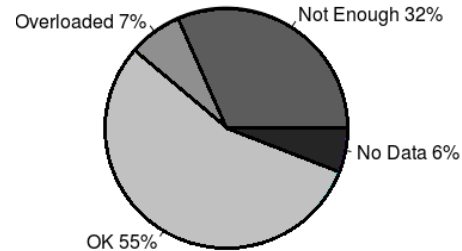


Figure 1. Evaluation of product orders using SmartOrder prediction (data from Martinus).

SmartCollect evaluation is in progress now. We have designed multiple tests, where one group of the workers choose the collection path on their own and second one use collection path generated by our service. Metric of this test is collecting time of the order processing. We plan to use step counters to see whether the workers do less physical effort.

Our solution consists of two modules - SmarOrder and SmartCollect, which are designed to improve effectiveness and decrease delivery time of ordered products. First goal is to make products available in the warehouse, which is handled by SmartOrder. Subsequently, computing of the optimal collect path is executed by module SmartCollect, which saves the time of the warehouse worker and an order can be sent to the customer in the shortest possible time.

Together, these solutions gain an advantage to an e-shop compared to other e-shops, which are not using similar technology. To evaluate our improvement, we compare our achieved results with existing results from partner e-shop and we take into consideration feedback of workers in its warehouse to make our solution as smart and usable as possible.

References

- [1] Key, R.; Dasgupta, A. Warehouse Pick Path Optimization Algorithm Analysis. In: *Proceedings of the International Conference on Foundations of Computer Science (FCS)*. The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp), 2015. p. 63.
- [2] Knight, W. 2015. *MIT Technology Review*. Inside Amazon's Warehouse, Human-Robot Symbiosis. <https://www.technologyreview.com/s/538601/inside-amazons-warehouse-human-robot-symbiosis/>

Tasks, Code Reviews, Activities, Source Code, and Knowledge about Software

Peter BOBOVSKÝ*, Michal KRÁL⁺, Peter KUČERA*, Marek MURA*,
Miriam POMFFYOVÁ*, Lukáš RAČKO⁺, Michal SLOVÍK*

*Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
skupina8fiitteam@googlegroups.com*

Nowadays, it is common for software developers to rely on a quantum of support tools to aid them in their work by handling source control, task management or bug tracking. Data provided by different tools is poorly interconnected and therefore offers little perspective on the activity of individual developers or the project's overall development progress.

The goal of our project is to collect data from selected repositories such as StackOverflow, Git, Gerrit or Bugzilla, analyze it, find connections and link relevant data from different sources together. By achieving our goal we create a common repository for all commonly used software development tools, where one can easily keep track of every aspect of their software project's development and discover before unseen connections between different repositories.

Development of new software consists of multiple steps. During development, developers use various environments that help them with collaboration and communication. These environments include control systems, integrated development environments, discussion boards etc. Data stored in these environments by software developers form tracks of software development. These tracks help us understand how to make development more efficient. However, these tracking tools are independent. Data stored in them is barely linked and typically the only connection one can find are users with the same name or identification number.

We propose a system that connects data from these various environments together under one roof - a system called TRACKS.

There are many solutions offering similar services as our own. Another well-known example is the Team Foundation Server (TFS), which connects data about tasks, people responsible for each task and

their code. Another system, Gerrit, contains data about versions of code and the person responsible for writing and reviewing the code.

The advantage of our solution is that it eliminates the need for a direct connection between the IDE and each individual repository. We create a centralized repository where data from various sources such as Stack Overflow, Bugzilla or Github are stored and linked together. Support for downloading data from new repositories can be added at any time.

The main difference between existing solutions and our own is that they only collect and link data of a specific type from a single repository, while our solution gathers and connects data related to all aspects of software development - tasks, code reviews, activities, source code, and knowledge about software.

An important feature of our solution is that the preprocessed data stored in our TripleStore database can be used further by researchers for the purposes of data mining and knowledge discovery, without the necessity of repeated downloading.

Developers use various tools while working in software development. These tools range from source code in a Git repository, through bugs in Bugzilla, all the way to finding solutions on Stack Overflow to problems marked in Microsoft Team Foundation Server.

These tools mostly lack any connection between each other. As such, data about a singular product is spread across several places. And while it is possible to keep track of all these places, it creates unnecessary work for the developers and slows down their development process by a big margin.

In order to connect these repositories, we have to analyze them, analyze the data they contain and figure out how that data can be connected.

⁺ Master degree study programme in field: Software Engineering

^{*} Master degree study programme in field: Information Systems

Supervisor: Martin Konôpka, Institute of Informatics, Information Systems and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

We have decided to download and connect data from these repositories first:

- Stack Overflow - used mainly for problem solutions. Every topic is a question. A question can have answers and comments, which helps find important and relevant data more easily.
- Bugzilla - contains reports of bugs in a product and the steps required for reproduction of these bugs.
- GitHub - contains all source code for open source projects. Also frequently contains discussion and bug reports, which can sometimes be linked to Bugzilla.

These repositories offer application programming interfaces to get the data they have. With only what is offered to us, we are able to download large volume of data from these sources from the past, as well as downloading any new being created.

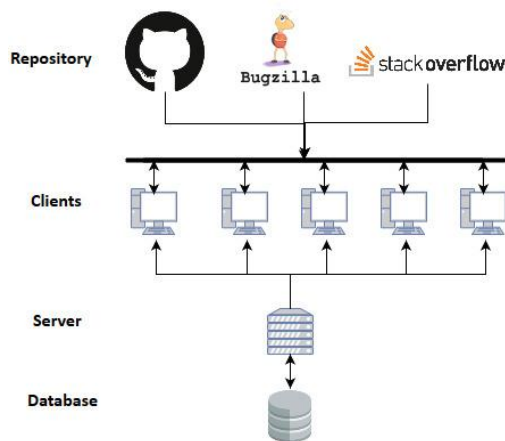


Figure 1. Communication overview.

In this section we describe how and where we gather all the necessary data and why we have decided for this solution. Instead of downloading all the data from a single device, we have decided to use a botnet network of client devices, and it was for a great cause – we prevent servers that host repositories from getting overloaded and additionally, our system does not appear to them as a DDoS attacker. A considerable setback of this solution is the possible negative impact of the client's network activity on the local machine's performance.

The process of downloading data from a specific repository consists of several steps. Firstly, the server assigns a new ID to the communicating client. After receiving the ID the client requests a new JobExecutions. JobExecutions is a file that contains job details and settings. Client checks this file and if the corresponding JobExecutable is unavailable or

outdated on the local machine, it requests an update from the server. The JobExecutable file is responsible for direct communication with the target repository.

JobExecutable contains data from the repository which is then transformed into RDF graph form. These RDF graphs are stored in a remote apache-jena-fuseki database. The data stored in RDF graphs allows us to connect them, analyze or do general data mining research on top of it (Valerie Bönström, 2003)[1]. An overview of components participating in data downloading and storage can be seen in the Figure 1.

TRACKS is a software development tool that helps all people in software related professions from small teams to big companies to see yet undiscovered tracks of work done on projects. By analyzing data on several levels we search for tasks, code and people involved. This allows us to connect information from every analyzed repository on the same project. With the results of this process we can help clarify connections and mutual relationships between repositories, that were unseen before. Teams can switch to a new project effortlessly, due to easily discoverable relations between multiple project repositories. Project managers can easily track productivity of their coworkers and track which problems they were stuck on, and how they solved them. In general, this software development tool can be used in projects which lack the contextual connection between its own multiple underlying systems.

References

- [1] Bonstrom, V., Hinze, A., & Schweppe, H. (2003, November). Storing RDF as a graph. In *Web Congress, 2003. Proceedings. First Latin American* (pp. 27-36). IEEE.

Segmenting Walls and Objects from 3D Hierarchical Data

Jakub GINTER*, Miroslav HAŠČIČ*, Mário HUNKA*,
Viktor KOŠŤAN⁺, Richard PINTÉR⁺

*Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
teamfiit11@gmail.com*

In this paper we present a system for reconstruction and automatic segmentation of hierarchical 3D data. Output can be used by CAD software mostly used by architects and designers. We created a pipeline consisting of a certain combination of commonly used segmentation methods. We discuss structure of our system by modules, presenting its advantages that brings potential business value, but also identifying flaws that need to be handled in our future work.

Work of an architect often requires tedious measurements of building interior, which is followed by yet another tedious work of putting measurement data into software like AutoCAD. Our main motivation for this project is to automatize these activities and ease work for our end users by that.

Building interior can be scanned by depth cameras, which provides us measurements and disposition of interior surface in form of point cloud. Open source libraries (PCL) provides many segmentation methods that can be used to recognize building foundations and individual objects [1].

Available software products have many drawbacks - incompatibility of input data, too general purpose software (there is no specialized software for building plans), financial demands. We present a system that provides a wide variety of data input formats, requires no interaction from user and specializes on buildings.

We designed the system to be able to cover entire process from acquiring, reconstructing and evaluating data with the possibility to export results to formats compatible with software that is used by our users. Therefore is divided into 4 main modules.

Interior3DRecon module handles core functionality of our software - segmentation. It contains class *ClosedSpace* which serves as its interface and represents a building interior to segment. This class contains methods for loading

point cloud data from external file, segmentation of the loaded data and export to DXF and PCD file.

Visualization module is responsible for visualisation of *Interior3DRecon* results mainly for visual validation of the results by developers, but also for end user to view the results before exporting. It requires *CloseSpace* class from which it gathers segmentation results. It also provides interface for viewing visualizations represented by *VisualizerWindow*.

Kinect_GUI module provides tools for hand held scanners. Main functionalities are initialization of scan and exporting point cloud. GUI also offer testing utility for scanner drivers and for information purposes shows log data about scan progress. After initialization of scan, data grabber tries to load connected device and dedicated graphics card because of performance boost. Reconstructed point cloud is visualized during the scan.

QT_GUI module provides GUI for users that lets them call methods from *Interior3DRecon* and *Visualization*. It initializes *ClosedSpace* instance after input data files are specified and then provides means to execute the segmentation methods. In addition to this it creates *VisualizerWindow* instance when a visualisation is demanded and passes it *ClosedSpace* instance.

Our goal is to develop system that is not limited by dataset quality or format. Therefore it can process variety of data types like: cvs, ply, obj, etc. These types contain information about scanned points and their position in space. These data sources can contain inconsistencies in density. We handle noisy dataset with *Outlier Detection*. We try to identify outliers, remove them and create more accurate dataset by that. Sparse dataset is matter of our future work [2].

Our main source of raw data are point clouds created by scanning environment by Microsoft Kinect

* Master degree study programme in field: Software Engineering

⁺ Master degree study programme in field: Information Systems

Supervisor: Assoc. Professor Vanda Benešová, Institute of Computer Engineering and Applied Informatics, Faculty of Informatics and Information Technologies STU in Bratislava

one. We use SLAM algorithms to create complete scans of interior spaces. Our system offers capabilities to scan with different types of sensors too. User can see real time depth data coming from sensor that are being reconstructed.

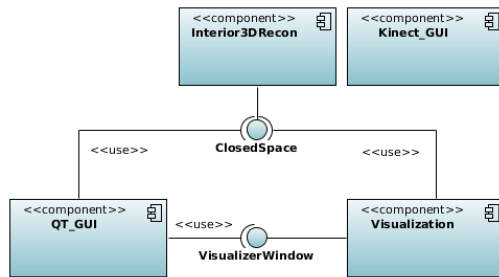


Figure 1. Model of main modules.

Original point cloud data goes through processing pipeline that segments planar wall objects from the point cloud. Firstly, point cloud is segmented based on dominant normal of points. Then these segments are segmented using RegionGrowing algorithm to differentiate between points that do not belong to the same object.

Next is plane fitting using RANSAC algorithm associated with filtering of the results based on inliers and outliers ratio. Further plane edges are computed. These wall objects are part of building foundations, so there is emphasis on accuracy. Therefore, recognized planes are further refined by removing outliers and repeated plane refitting [3].

For development of the most suitable segmentation methods we have implemented own visualizer. It serves only for quality verification for developers. In the future, when the system will be on the public market, it is highly possible that our end user will need this feature as well.

Our target users are architects, interior designers and all who needs to scan some scene and work with its model in CAD software. That's why is our final output in DXF format, which can be processed by every CAD tool, i.e. AutoCAD. This 3D model can be further edited and exact measures of various objects can be computed. And that's primary usage of our system. User can easily scan a building or room (or insert some dataset), process it with our system and open new 3D model in his favourite CAD tool.

In this paper, we gave an insight to our project that primary concerns with scene segmentation. This topic itself is matter of research nowadays. There is no fully functional system that can replace all the work of an architect.

In our research, we try to come up with a method that reflects reality at its best (size, distances, angles) - walls are perpendicular to each other, the measurements of building/room correspond with reality, rooms contains objects with exact position and size.

We think of this software as an utility for architects/designers that unburden them from tedious work. In the future, we will focus on object recognition and identification. In other words, we want to know that this object is a chair and that one is a sofa. With growing complexity of computations it is necessary to consider parallelism or another optimization to minimize computing time.

References

- [1] R3D Team, "Rekonštrukcia 3D scény Projektová dokumentácia -inžinierske dielo,," 2015.
- [2] S. Sotoodeh, "OUTLIER DETECTION IN LASER SCANNER POINT CLOUDS."
- [3] R. Schnabel, R. Wahl, and R. Klein, "Efficient RANSAC for Point-Cloud Shape Detection," vol. 0, pp. 1–12, 1981.

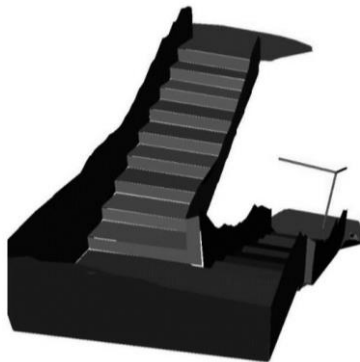


Figure 2. Preview of segmented output.

Scalable Architecture and Implementation of Media on Demand System in Vehicles

Tomáš BARÁNEK*, Lukáš MASTILAK*, Matúš POHANČENÍK*, Matúš SLOVÍK*,
Tomáš SOKOLÍK*, Jaroslav TÓTH*, Radoslav ZÁPACH*

*Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
team17-2016@googlegroups.com*

In the domain of public transportation, each transport company makes effort to achieve clients thanks to different additional services during traveling. One of these services is offering a multimedia content to passengers. Actual solution offers them this service through the seat-back screens in buses and trains. The aim of this project is to enable passengers to transfer this multimedia content from the built-in screens to their personal devices, either to mobile phones or to notebooks and enable portability between different devices and operation systems. Based on this, the system is built on a web portal solution that offers users the multimedia content on a customized web page, which can adapt to different devices by a responsive design.

The developed solution of the web portal delivered to passengers with help from built-in screens and web portal is aimed to be the replacement of a previous solution. The previous solution was developed by Funtoro company – its server-based environment is proprietary, lacks an opportunity to add new features, open documentation, and flexible administrative configuration. These factors led to the development of the server design with the similar services (centered around the provision of multimedia content to users, remote configuration, centralized reporting and refreshing of multimedia content), but with scalable (modular), and well-documented architecture.

The Figure 1 shows three-layers architecture that is applied on the application server (component diagram). The Client layer sets presentation elements in the form of the web page that is shown on the screens of users at the built-in screens and is also available through the 801.11n connection for other devices. Firstly, web page is used to provide access to selected multimedia files that are located on Media on Demand (MOD) server. Secondly, clients at screens

have also option to connect external storage devices to a screen and browse multimedia files on them. Reload component can be accessed to control state of screens. Multimedia services include streaming of movies, music albums, picture slideshows, maps with GPS functionality, web games, TV multicasting, and Internet access. This functionality can be easily upgraded with another functionality that is available through the web portal or another Android apps.

Services and events that are invoked on the provided web page are processed in the Web Portal layer. The communication between Web Portal and Client is implemented using HTTP protocol and REST API. This layer handles listing of multimedia files, creating of streaming link to the desired file, collecting of statistics from displaying and usage of items and interactive maps. These services can be associated with the user console.

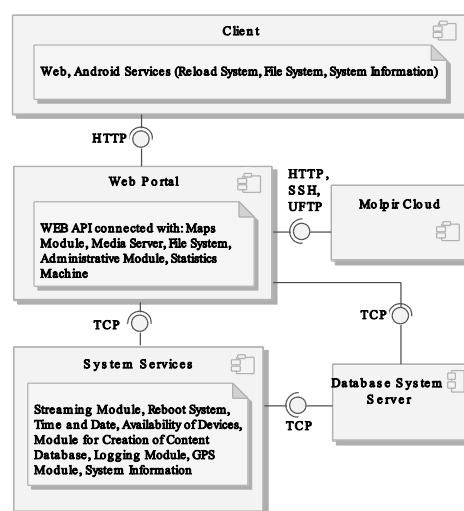


Figure 1. High-level architecture of MOD server.

* Master degree study programme in field: Software Engineering
Supervisor: Dr. Peter Pišteň, Institute of Computer Engineering and Applied Informatics, Faculty of Informatics and Information Technologies STU in Bratislava

On the other hand, web page can be accessed in the form of administrative, install, or steward console that provides additional services for administration of content, initial installation, and controlling of devices by steward (Administrative Module). This layer keeps connection with Database System which preserves information about actual multimedia content on the disk (including its physical location on the disk), collected statistics, and logging information. The database System is locally available on the server; the communication is maintained by TCP socket. Next, there is external Molpir Cloud module connected to this layer – its role is the remote administration of MOD servers, controlling of data flows between vehicles and transport servers and storage for logs and statistics.

The last layer interacts with the operating system and devices that are connected to the server. There are services that are responsible for transcoding and streaming, providing distributed time and date, rebooting of built-in screens, logging of critical system modules, GPS communication, monitoring of connected screens and monitoring of the connected SSD disk with the creation of content database. The communication with the upper layer is performed with TCP sockets, too.

Figure 2 shows deployment model of our new system. This model is not so different from the previous version. The main change is the system which is running on MOD server. MOD server represents a special server which is adapted for using in demanding conditions (buses, trains, airplanes, etc.). Peripheral devices (IP camera, built-in screens, etc.) are connected by the hubs to the server. The function of Gateway router is to connect User Devices through 802.11n and the arrangement of communication with Cloud services (for example

Centralised Management). New software design of MOD appliance brings many improvements against actual solution and thanks to that fact it interferes problems with an old solution from the company Molpir [1]. The system is portable between Linux platforms. Our solution is unique and it should bring more comfortable travelling to the passengers.

Summary of the main improvements over the old solutions is depicted in the following list:

- The designed system is well-configurable so customers and administrators can quickly adjust behaviour and user-design of the system without any programming skills.
- All user services can be accessed from the web interface – they are unified and accessible from any device.
- The architecture is implemented in the modular way where the components don't rely on each other so MOD system can be enhanced in the future and offer next services for passengers.
- The operators can remotely track system state and reactively repair them. The administration is also supported by centralised logging system that is configurable.
- The statistics from multimedia content and advertisements makes this solution also interesting for marketing purposes.

Acknowledgement: This work was assigned and partially supported by company Molpir s. r. o.

References

- [1] Molpir s. r. o.: MEDIA ON DEMAND SYSTEM. In: User Manual, (2013). [Online; accessed February 16, 2017]. Available at: http://shop.molpir.sk/Prilohy/254130002110_0.pdf

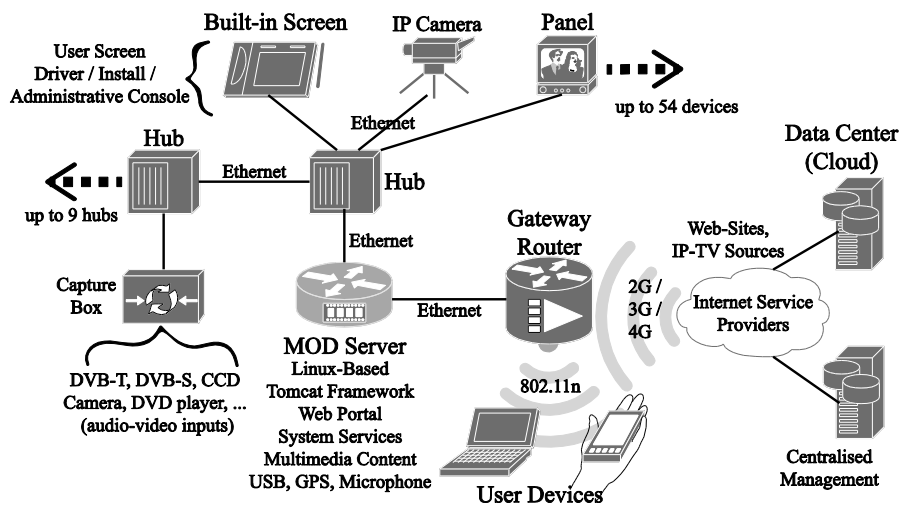


Figure 2. Deployment model.

Semantic Search in Old Press

Adam RAFAJDUS*, Martina REDAJOVÁ*, Tomáš REPISKÝ*,
Jakub HAGARA*, Jozef SITARČÍK*, Martin VAŠKO⁺

*Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
tp.team19.2016@gmail.com*

Searching and retrieving requested answers of users' questions become more important and interesting with increasing amount of information available on the Internet, therefore traditional search engines nowadays are not appropriate and not suitable anymore [3]. After initial searching it is often necessary to expand searchers' demands and repeat searching so the result is sufficient.

Semantic search is interesting and highly current topic for computer science community that solves the problem of finding answers to users' questions thanks to the search in natural language, e.g. using graphs. These are parsed and mapped to multiple languages and ontologies. Based on that the SPARQL queries are generated, which are enriched, so the semantics are kept and result of search is sufficient. The main idea of the semantic search is to improve search accuracy by understanding searchers natural language and automatic reformulation of requests to needed syntax of search engine.

Semantic search focuses on collecting information about words and by that, it is able to detect meaning of these words, lines or even block of text they are part of and to recognize concept matching, relations and synonyms. However, for semantic search it is necessary to have digitized sources, while it requires machine processing.

Despite the fact, that most of sources are in digitized form, it is not necessarily the only field for applying semantic search. One of the fields where such sources are very common is press archiving, which is important, because old press issues contain useful and important information, which cannot be found anywhere else. It is common practice to archive press sources as one unit per issue. For the correct application of the semantic search it is important to parse these issues into individual articles, while each article is a separate unit, whose meaning and content can differ from the others. Whereas the structure of these resources is often inconsistent, especially when

the resources are from different country or time period, it is necessary to design a tool for parsing these resources between the descriptive parts and the reworking of periodicals to correct bibliographic records for individual articles.

Semantic search has been focus of many research papers, even specific definition we could apply for our use case such as Named Entity Recognition and Merging Inflected Forms of Named Entities [1], which helps create named entities in Slovak language improving semantic search as a tool, or Real-World Dependency Parsing of Slovak Text [2], which further increases level of quality for natural language processing of Slovak language. The challenge of our work is applying such methods of semantic search on our dataset, which consists of old digitized press issues for archiving purposes, meaning they come in not consistent state appropriate for information extraction right away.

Data in old press can be very important and can contain a lot of interesting information which may be used by digital humanities [4], scholars, etc. However, it is very difficult to effectively obtain and process this data apart from reading it by eyes. The data created by optical character scanner were irregularly assigned to blocks and paragraphs, not always corresponding to the image.

Thanks to this, our focus slightly shifts to propose a preprocessing method to extract articles from digitized data, defining elements of articles, such as key words, to be used in semantic search and create an environment for people to help automated process to be more precise in defining entities from raw digitized data.

Notable part of our contribution is preprocessing digitized data, which were captured by ABBYY - optical recognition software which provides data in .xml, .pdf, .str formats. We focused on preprocessing of XML files, as they consists of all useful information we can gather from digitized data.

* Master degree study programme in fields: Information Systems

⁺ Master degree study programme in fields: Software Engineering

Supervisor: Dr. Nadežda Andrejčíková, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

1. Firstly, we remove redundant information. It is mainly to reduce memory usage and to increase readability of the digitized data to improve analysis process.
2. Secondly, we identified and distinguished blocks of text as headings or fulltexts. To achieve that, techniques like most commonly used font size, and median font size are used.
3. Separators are identified as well, as our analysis shown separators to be helpful for further use, which is about matching headings to appropriate fulltexts, creating groups of headings and fulltexts - articles.
4. To obtain this, we suggest and implement an algorithm, based on our analysis. As a next step, we manually analyze finite number of digitized press numbers to find patterns in printing techniques, distribution of texts to heading, to find algorithm to automate article extraction process.

Our algorithm considers separators and position of the text blocks on the page, as the key indicator for matching headings to fulltexts. To consider article as correctly identified, all blocks considered to be part of article must be identified and merged. As it is shown in Table 1, our manual evaluation of our proposed algorithm shows that it has very high success rate at detecting/extracting articles.

Table 1. Results of manual evaluation.

| Number of articles | Correctly identified articles | Success rate |
|--------------------|-------------------------------|--------------|
| 573 | 557 | 97,21% |

After preprocessing input data, we apply syntactic analysis to determine entities and relations between them. Syntactic analysis has more phases. First we use statistical analysis to get keywords for each article. To eliminate synonyms, we use lemmatization for Slovak language. Lemmatization selects words in inflected forms and transforms them to lemmas. This helps us group words in inflected forms. Eliminating synonyms is needed to determine words with same meaning so we analyze them as same entities. The entities represent real world objects that are connected through various kinds of relations creating important associations between them.

All of these functions like preprocessing of input data, syntactic analysis, entities and relations are vital to implemented search engine which is type of graph-based search. User can do fulltext search through natural language with help of Elasticsearch in all saved journal issues and articles saved on our server database, where search engine uses all previous obtained key words, entities and related entities to improve results as much as possible.

Based on proposed preprocessing and analyzing digitized data we developed a web application designed to allow users work with old press issues, precisely with parts that were extracted. Application consists of two modules, each developed to fit the needs of their primary user.

- One part is dedicated to librarians, who are able to see how our automatic extraction tool processed files and created separate articles. They are also able to customize this division in case of improper preprocessing that may occur in the automatic processes.
- Purpose of the second part, dedicated to an ordinary user, is to provide a search tool that can be used to find articles not only based on title or author. The search tool is based on interactive graph search using all data acquired by syntactic analysis, keywords, entities and relations, basically every information we extracted.

In this paper we describe our project on semantic search. We created tool for extracting articles from each archived journal issue and assigning needed bibliographic records. Part of the work was the design of a custom algorithm for detecting and connecting elements in the articles and the application of lemmatization within them. The result of our work is the basis for a tool for semantic search, of which the main part is the mechanism for the processing of source documents.

This project serves as a basis for the development of a tool for semantic search. In the future we expect adding visualization of the search result and named entities, and the ability to search through charts. It is also possible to extend the tool on the part of the users' personalization, whether with the possibility of registering users, creating collections or setting up their own libraries.

References

- [1] Blanárik, M. Named Entity Recognition and Merging Inflected Forms of Named Entities. 2016. Master degree Research work. Supervisor doc. RNDr. Michal Laclavík, PhD.
- [2] Loeb, J., Šimko, M. Real-World Dependency Parsing of Slovak Text. In: Proc. Of The Tenth Int. Conf. On Natural Language Processing Dubrovnik, September 29 - October 1, 2016. Springer, 2016.
- [3] Sudeepthi, G. Anuradha, G. Babu, M. Surendra P. A survey on semantic web search engine. In: IJCSI International Journal of Computer Science Issues (www.IJCSI.org), March, 2012. volume 9, issue 2. pages 241-245. [cit. March 20 2017]. ISSN 1694-0814
- [4] Sabharwal, A. Digital Curation in the Digital Humanities: Preserving and promoting archival and special collections. Chandos Publishing, 2015. ISBN 9780081001783

My Food Court

Martin PREKALA*, Miroslav RÁC*, Jozef BALÚN*,
Igor ŠIMKO*, Štefan SCHWARTZ*, Matej SCHWARTZ*

*Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
cely.tim@pokopec.sk*

To choose the right food for a lunch from a number of different types, tastes and cuisines is not always the simplest task of the day. The process of choosing a lunch is specific for everyone. Our product, represented in a form of the mobile and the web application, brings new perspective and way of solving this problem. Because we are focusing on a man, his personal profile and his preferences in eating habits, we are changing the view centred on the selection of restaurants for a whole new perspective which is focused on the food.

The following example illustrates the everyday routine of choosing the food for a lunch.

For Milan, who works at the city centre, it is comfortable to eat his everyday lunch at restaurants. Around his location, there are several restaurants and canteens, which offer various types of cuisines and a lot of different foods. Moreover, these offers are being changed every day. To choose the right food is not as simple task as it seems to be.

Everyday, before the lunch break, he visits websites of the nearest restaurants, searches for the lunch menu lists and looks for the right food for him. After few months, it becomes a necessary routine consuming the precious time. He also noticed, that menu is composed out of the limited collection of foods, which are repeatedly shown in the lists. His favourite one is tomato soup and he also likes the Asian cuisine. Whenever he sees tomato soup or Kung-Pao in the offer, the choice is made right away.

Milan is aware of the reasons why healthy eating is important for him. In addition, he is lactose intolerant. That is why he focuses on his diet. He is also really playful, so he appreciates any assistance to his everyday routine. He will spread the word about it among his colleagues and friends.

There are plenty of web and mobile applications, which aggregate daily menus. One of the biggest worldwide used application is Zomato. Zomato comes from India and is located in more than twenty

countries worldwide, from the United states to Canada to Italy. In all countries combined, it has 1.2 million of restaurants. In Slovakia, Zomato is the number one of food application.

Zomato and other applications like ham.sk mostly provide a list of restaurants, which can be filtered throughout user's location or restaurant type and emphasizes the user's restaurant preferences. Applications provide data from different type of catering facilities like restaurants, canteens, bistros, buffets, etc. and they try to show a lot of information at once, therefore the result is not simple. Easy survey is not presented and users mostly do not notice the important things. The advantage of these applications is, that data is obtained directly from catering facilities.

In comparison to other similar applications, the main functionality of our application is personalized food recommendation. My Food Court can predict what food do you prefer the most, and even raw materials you like. On the other hand, we can take into account your health condition. Our application is able to filter out foods with dangerous raw materials for you health. UX of our application in comparison to e.g. Zomato is focus on important things like user profile, distance, rating. All this information's are marked as important in user research.

In the first place, we deliver the list of meals to users in different way of representation. We give to them the list of meals sorted by their preference to food. There is a probability, the user will choose their meal out of the first places in the list. There is no need to read the entire list of meals, which can contain tens of items, if not necessary. Thanks to knowledge of their diet preferences, we can recommend meals also in the areas where our users are for the first time ever.

We retrieve data from users as they like or dislike food or they mark some as their favourite. They express their opinion about food just by clicking on its detail. These are all inputs for our recommendation machine.

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

Accompanying Events - TP CUP

In current version we suggest three types of recommendations.

1. *Popularity recommender* - in this recommender is based on food views. Food has better rate for view thier detail.
2. *Like recommender* - this is basic approach based on users interaction and rating system. Users can like or dislike food and affect food in food list.
3. *Distance recommender* - for some people as we find out are very important distance to food. We introduced distance recommender, closest daily menus are higher rated for user.

In the future we will prepare recommendation method is based on the profile similarity for each user. We create user based models containing their history of clicking and liking foods. Afterwards by comparing similar models we can predict what some user wants or does not. Based on the total amount of matched profiles, we can say with a certain probability that the list of preferred foods is the same.

User can always see the list of all available foods. List items are ordered by actual context of user preferences. These preferences can be controlled by user. They can simply switch recommendation system, which causes reordering of displayed food list. For example, one type of preference may be the number of likes for the food.

Every user has the option to adjust their preferences, respectively. It is possible to filter the list of searched foods. The filter includes:

- *Geolocation* - In case a user wants to schedule lunch outside his current position, he can adjust the position of the pin.
- *Radius* - restaurant radius search changing
- *Type of food* - to choose between various kinds of food (sweet, fleshy, vegetarian, gluten-free, ...)
- *Date* - the relevant date of serving time of daily menu

Nevertheless, the main value of My Food Court is not in the usage of mobile application and people browsing and looking for their next meal once a day, everyday. Obviously, without them, application could not function, however, it is just a mean to the most valuable aspect. It is hidden in the data about these user.

Real value of this product is in data collection. My Food Court application aggregates data related to locations, which may vary from large areas like neighbourhood settlements or just one functional building. Instead of just helping users to choose foods for a lunch, we can use these data for recommending to restaurants owners what foods should they include

in their menu to improve their sales. Moreover we can give valuable information to developers about missing restaurants in area.

In the near future, maybe you will not have to ask yourself what do you want to eat for a lunch. My Food Court will recommend you something suitable for you. As a next step we plan to implement artificial intelligence, which will learn your eating habits based on your history. Maybe you enjoy your dessert every Wednesday or hot tomato soup always enhances your mood during rainy days - our application will recognize that. My Food Court will see, that your nutritional intake was not good last week. Our recommender will suggest you some healthy food. It will ensure that you have balanced diet.