Proceedings of the

Sixth York Doctoral Symposium on Computer Science & Electronics

Department of Computer Science & Department of Electronics
The University of York
York, UK

October 28th, 2013
Editor: Sam Simpson

Sponsoring Institutions

We are extremely grateful to the following institutions for generously sponsoring YDS 2013.

Academic Sponsorship

THE UNIVERSITY of York

The Department of Computer Science, the Department of Electronics & the Research and Enterprise Directorate.

Industrial Sponsorship

IBM

BT

DRIVING EMBEDDED EXCELLENCE | ETAS

RAPITA Systems Ltd.

Preface

YDS is an international post-graduate student conference, now in its sixth year. For the first time this year, YDS was organised jointly by post-graduate students from the Department of Electronics and the Department of Computer Science at the University of York. The primary aims of the symposium are to provide an opportunity for early-stage researchers to observe and take part in the process of academic publication through organising a full-scale, international academic event, and to allow doctoral students from the UK and Europe to experience a peer-reviewed conference where they can share and exchange their research and ideas in a supportive environment of their peers. Furthermore, by bringing together students from a wide range of areas, YDS hopes to also promote more interdisciplinary research.

YDS 2013 offered three categories for submission, full-length papers, extended abstracts and posters. We received 5 extended abstracts, 11 full-length papers and 16 posters. Full-length papers and poster submissions saw an increase in submission rates over previous years. Acceptance rate overall for long papers and extended abstracts was 56%. As well as papers from students at the University of York, we had papers submitted from 8 UK universities and 4 universities from mainland Europe, all of which bodes very well for the future of YDS as a leading venue for the dissemination of doctoral work. We also received a good mix of computer science and electronics submissions. It was important that the YDS Programme Committee ensured that all submitted work was reviewed with due anonymity and fairness, as the event is a serious training exercise for the hosts as well as for the authors. To this end, this year, short and long papers received three anonymous reviews each, with posters receiving one review each.

We are very grateful to our sponsors for allowing us this fantastic opportunity, the Department of Electronics, the Department of Computer Science and the Research and Training Directorate at the University of York, IBM, BT, ETAS, Rapita Systems and PurpleFrog Text Ltd. Among other things, their financial help enabled us to offer prizes for best long paper and extended abstract, best presentation, best poster judged by the industry panel and best poster voted for by attendees. Distinguished representatives from our sponsors also sat on a well-received panel question and answer session allowing students to find out what options await new PhD holders in industry. We were also honoured to host an invited keynote talk by Prof. Rashik Parmar of IBM.

I should like to express my warmest thanks and awed appreciation to the Organising Committee and the two co-Chairs, Pedro Ribeiro and Simos Gerasimou for their efficiency and professionalism in dealing with all aspects of the logistics of the event; to the wonderful Programme Committee for the time and effort

in giving complete, objective and impartial reviews; to our academic staff for their promotion and support of YDS, especially Mike Dodds, and with a special mention to Prof. John Clark who has always been very supportive but this year once again went the extra mile for us; to the Computer Science administration staff who were always helpful and cheerful whatever we asked them (which was a lot at times); to Camilla Danese in Electronics for all her help; and to the University for giving us the opportunity to run YDS.

YDS 2013 has been a fantastic, busy, occasionally stressful but hugely enjoyable experience which I am very happy to have been given the opportunity to be part of and I would encourage all post-graduate students to be part of this event.

Sam Simpson
Programme Chair, YDS 2013

Organisation

All members of the YDS 2013 Organising and Programme Committees are from the University of York, UK.

Organising Committee

Konstantinos Barmpis
Theodoros Georgiou
Simos Gerasimou (co-chair)
Usman Khan
Stuart Lacy
Nils Morozs
A. Imran Nordin
Jimi Ogunyomi
Pedro Ribeiro (co-chair)
Ibrahim Shiyam
Sam Simpson
Alexander Turner
Andy Turner

Programme Committee

Mustafa Aydin
Furqan Aziz
Chris Bak
Kostas Barmpis
Mohammed Baz
Pedro Burmester Campos
Prof. Alan Burns
Prof. Alister Burr
Ipek Caliskanelli
Gabriel Costa Silva
Ralph Eastwood
Christos Evripidou
Abimbola Fisusi
Luis Fuente
Prof. Edwin Hancock
Tiong Hoo Lim
Prof. David Howard
Horacio Hoyos Rodrguez
Sorachai Kamollimsakul
Ali Karami
Prof. Tim Kelly
Selahattin Kosunalp
David Lawson
Prof. Andy Marvin
Prof. John Mcdermid
Azizi Mohdsayuti
Nils Morozs
A. Imran Nordin
Ibrahim Ozturk
Prof. Richard Paige
Prof. Helen Petrie
Simon Poulding
Jo Pugh
Roger Rivitt
Prof. John Robinson
Prof. Colin Runciman
Sam Simpson (chair)
Frank Soboczenski
James Stovold
Bharath Sudev
Prof. Jon Timmis
Prof. Andy Tyrrell
Prof. Jim Woodcock

Contents

IV Poster Presentation Abstracts

Part I

Keynote Talks

A Glimpse into the Future of Technology

Rashik Parmar

IBM UK Ltd.

rashik_parmar@uk.ibm.com

Abstract. Technology has become an essential part of daily life and in many ways it is invisible to an individual. However, we are still a long way from some of the visions that Science Fiction authors have described in the past. The talk explores some of the ways in which we can expect to evolve rapidly and transform peoples lives. It also looks at some of the global challenges that will inspire innovation in technologies and business models.

Biography. Rashik Parmar is president of IBM's Academy of Technology and also a Distinguished Engineer. During his twenty eight years of practical experience in IBM, he has worked for financial, retail and manufacturing clients on IT projects of all sizes. Overall, he specialises in ensuring the technical success of complex IT projects. He currently leading projects related to IBM's Smarter Cities programme and development of techniques to drive industry level innovation.

Rashik is IBMs Partnership Executive for Imperial College, London. He is also an Adjunct Professor for Department of Innovation and Entrepreneurship at the Imperial College Business School and Visiting Professor to the Intelligent Systems and Networks Group at the Department of Electrical & Electronic Engineering.

Part II

Full-Length Papers

Examining eye-tracker scan paths for elderly people using smart phones

Suleyman Al-Showarah, Naseer Al-Jawad and Harin Sellahewa

Applied Computing Department, The University of Buckingham
Buckingham, MK18 1EG
{suleyman.al-showarah,naseer.aljawad,harin.sellahewa}@buckingham.ac.uk

Abstract. A scan path is a set of regions represented by letters given to the visited regions on a screen until a target is found; moreover a scan path is one of the most important metrics measured by eye tracking systems. This paper presents experiments conducted on three screen sizes of smart phones (3.2'', 7'', and 10.1''). String edit method was used to compare the sequence of scan paths for two different age groups (younger age group 20-39 and elderly people 60+). The aim of this study is to investigate; first: if scan paths are smart phone screen size-driven or stimulus driven; and second: to investigate if scan paths are age-driven. The results revealed that there is stimulus-driven in particular smart phone applications than screen size-driven, and also there is age-driven yielded for all the experiments results in favor of younger age group.

Keywords: Ageing effect, elderly people, human computer interaction, eye-tracking, usability of smart phone, scan paths, and mobile computing.

1 Introduction

Due to the increased population of the elderly people over the world in general and the availability of new smart phones technologies, it becomes important to address the elderly needs in using smart phone technology that rapidly changes and develop. The potential of this research for elderly people has been mentioned in different researches for the lack of data of elderly people regarding the effects of ageing on technologies [1], [2].

The way of using equipments to acquire human skills and experience is helpful; using eye movement is similar to the other methodological inputs (e.g., gesture, conversational speech [3]), also eye movement is considered to be a faster input media than others, and it is recommended for its high accuracy [4].

Eye tracker devices are important input media used widely in human computer interaction to collect and measure data related to search efficiency/inefficiency and cognitive processing ease/difficulty [5]. Eye-movements are recorded by eye tracker devices and can be used in various fields for measuring human's viewing behavior.

There are two groups of metrics derived from eye-tracking systems. The first group is known as static metrics and the second group is known as dynamic metrics.

Calculating the accumulated data or averaging the observed data is mentioned for static metrics, such as fixation durations, number of fixations, fixations rate and saccades rate. Dynamic metrics are mentioned for the data that obtained from time series data that contain information on the time axis such as the scan path [6], [7].

Age-driven studies for elderly on smart phone are not covered yet. Examining scan-paths is useful for understanding overall tendencies of viewing behavior of people. In this paper, we are investigating the scan paths on smart phones applications in two ways; first: investigate if scan paths are smart phone screen size -driven or stimulus-driven, second: to investigate if scan paths are age-driven on smart phone.

String-edit method; commonly known as Levenshtein distance method is widely used for calculating the distance between two pairs of scan paths to find the similarity between them [8], [9]. In string-edit, there are three operations for calculating the distance. Those operations are called deletion, insertion, and substitution. When using these operations, one string can be transformed into another string. The distance between the two strings is defined as the smallest cost when matching any two strings [8], [10], [11][9]. The Levenshtein algorithm is further explained in Section 4 under Experimental results and discussions .

The rest of the paper is organized as follows. In section 2, the related studies for scan paths are discussed. Section 3 explains the methodology and procedures relevant to the study. A discussion of the experimental results is presented in section 4, followed by our conclusion and future work in section 5.

2 Related work

There is in general, a shortage of studies that take into account the effects of ageing on technology. In particularly, there is limited research to investigate the effect of ageing on the use of smart phones. Moreover there is no study till now to our knowledge that has investigated if scan paths on smart phones are age-driven. One of the topics covered in scan path studies is if scan paths are stimulus-driven and participant-driven. Josephson and Holmes [12] analyzed scan paths of repeated views for participants using three web pages. The string edit method was used to measure the resemblance of eye path sequences. A study was conducted on students with an average age of 22.5 years; all participants were regular users of the Internet with an average usage of almost 9 hours a week. The participants were not given specific tasks in viewing the web pages, and they were told to view the page as he or she would in that situation for 15 seconds. A repeated measures analysis of variance revealed a statistically significant main effect for stimulus type, with a text-intensive news story page generating more similar sequences than a graphic-intensive advertising page.

Most similar sequences were from different participants, rather than from the same participant for the re-exposure view, suggesting strong stimulus influences. Stimulus-driven was revealed from their study by viewing web pages as they would view it in that situation, participants were not given specific tasks in viewing the web pages, but our study depend on finding the target on the screen. Whereas in their study; the participants were asked to revisit web page to find the similarity between first and

second view, but our study aimed to find the similarity between each age group for each smart phone screen sizes separately. Moreover they applied their study on a computer monitor, where as we applied the study on smart phones.

Another topic covered in scan path field conducted by Takeuchi and Habuchi in [7] was to describe a new method - uniform distance model - to be used in analysis for scan-paths based on string-edit method. They applied their study on 20 participants; 10 of them experienced in using internet for more than 10 hours a week, and other 10 participants experienced in using internet for less than 6 hours a week. Each participant was asked to browse three web pages, each for 20 seconds. The results revealed that uniform distance model is the smallest distance compared to city block distance model and Euclidian distance model. There was no specific task, just browsing the web pages until the eye tracker ends by its self for 20 seconds. In contrast, there are specific tasks in our experiments to find a target on the screen and the experiment ends for each task once the participant finds the target or the eye tracker ends by itself if the participants could not find the target in 30 seconds.

In the field of modifying the original sequence of the letters based on the fixation duration; the researchers in [13] have conducted study titled by "Scan path Analysis by the String-Edit Method Considering Fixation Duration". Their study aimed to compare the performance of the modified coding against the original one. They construct the modified sequence from the original sequence by repeating the character for the modified one based on the fixation duration by using the formula (n = [(t - 50) /100] +1) that says when the fixation duration is between 100ms and 150msec, a single letter (e.g., "A") is used to modified one, and a double character for the modified (e.g., "AA") when the fixation duration is between 150ms and 250ms. The results revealed that the modified scan path differs much from the original scan path, and they recommended for the need for further studies to select the suitable scan-path type, depending on the purpose of the analysis.

In the field of comparison for the difficulties in behaviors, the researchers in [1] have conducted a study between younger (13-29) and elderly (62-74) based on eye tracker metrics: scan path length for the length of saccades by degree in visual angle, fixation duration in ms, duration of task accomplishment per seconds for accomplishing the task on web pages. The result shows that there were difficulties for younger and elderly users caused by an inconvenient navigational structure and an inappropriate page design, and there are several important differences between the two age groups. It suggests the necessity of particular consideration for elderly users in Web design.

The authors in [14] have conducted a study to investigate if scan paths were participant-driven or stimulus-driven using seven participants (two dyslexics and five non-dyslexics, ages are between 22-40 years). All the participants have at least seven years experience in using the web. They conducted the study based on repeated views to find the similarity of scan paths between first and second view as done in [12]. By using string-edit method to compare scan paths in repeated exposure to web site home pages, the results revealed that there was variability in scan path for dyslexic people, and similarity between scan paths of different participants was higher than the similarity between scan paths of the same participants repeatedly exposed to the same

stimuli (stimulus driven). Although they used Levenshtein algorithim to measure similarity, their focus was on the attention images shown on web pages on a computer screen. We applied our study on smart phones for elderly by using Levenshtein algorithim as well.

Most of the papers in this field (analysis scan paths search) are interested in finding if scan paths are stimulus-driven or participant-driven, describing new methods to compare scan paths, and comparing between methods used for finding the distance between sequences of letters, and finding the distances between scan paths based on geographical position of letters (i.e., based x and y coordinates). In addition, re-exposure studies for scan paths have been mainly used in most researchers' work that used to find the similarity between pairs of scan paths, but there is no considaration yet on the age-driven for elderly on smart phones with scan path search. As examples of researchers who have conducted research in the field of scan paths as mentioned above are: [12], [7], [13], [14], and others. In this paper, we are investgating if scan paths are age-driven, and if scan paths are smart phone screen size-driven or stimulus-driven.

3 The Study Methodology and Procedure

3.1 Smart phone screen sizes

Our study consists of three representative smart phone screen sizes. Typical s*mall* smart phones have screen sizes between 3 and 5 inches. *Medium* size represents mini tablets screen sizes, which are typically are typically 7 inches. *Large* size represent full-size tablet devices which are typically 10.1 inches or larger. We chose 3.2 inches to represent small, 7 inches to represent medium screen sizes, and 10.1 inches to represent large screen sizes. Specific details of the selected devices are available in section 3.5.

3.2 Experiments

Each smart phone size has two experiments (EXP1 and EXP2) and each experiment is conducted using participants from two age groups (younger age group YG, and elderly age group EG). A participant will be involved in only one experiment to avoid any influence on the participant's performance. Fig. 1 illustrates the experiments and different groups of participants.

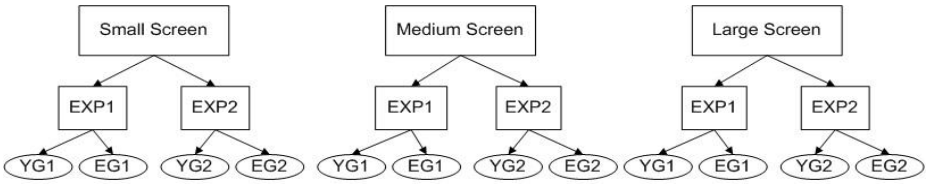

Fig. 1. Smart phone experiments on different screen sizes using participants of different age groups.

3.3 Participants

32 participants took part in this exploratory experiment. Group YG includes 16 younger participants aged between 20-39 years, and group EG includes 16 elderly (age 60+). All the participants have experience in using apps with smart phones (e.g. Skype, Facebook, and email systems) for less than one year and half.

3.4 Stimuli

The experiment is composed of nine smart phone applications; each of these applications has two experiment groups (EXP1 and EXP2) as illustrated in Fig. 1 above. The Table 1 below shows the questions/tasks posed to participants of each experiment group.

Table 1. Experiment questions and applications for EXP 1 and EXP 2.

App no	Apps	EXP1	EXP2
1	Skype contact list	Locate the user who is not online.	Locate the image of David Albert.
2	Skype Calling screen	Locate the Backspace button.	Locate the Numbers Field.
3	Skype Profile screen	Locate the button used to change your current status to be visible	Locate the account holder's Picture.
4	Facebook account holder profile	Locate the number of incoming messages	Locate the account holder's name?
5	Yahoo Email folder list	Locate the number of deleted messages.	Locate the number of new messages?
6	Gallery screen	Locate the delete image button?	Locate the Share photo button?
7	Alarm screen	Locate the active alarm?	Locate the button to add a new alarm?
8	Skype Main Screen	Locate the button that will show a list of contacts?	Locate the 'exit' button?
9	Setting screen	Locate the button to view Wireless and networks settings?	Locate the button that lets you change Sound settings.

3.5 Apparatus

Eye-link 1000 desktop device mounted with an IR illuminator on the right, which is used to track and record eye movements. This illuminator is used to generate reflection patterns on the user's cornea [15], [16]. A chinrest was used when conducting the experiments to fix the chin and forehead of participants at a distance of 50-55 cm.

Three types of smart phones have been used in conducting the experiments: HTC wildfire mobile phone, dimensions 106.8 x 60.4 x 12 mm, screen resolution: 240 x 320 pixels, with screen size of 3.2 inches; Samsung Galaxy Tab 2, dimensions 193.7x122.4x10.5 mm, Screen Resolution 1024 x 600 with screen size 7 inches. Samsung Galaxy Note 10.1 inches, dimensions 262x180x8.9 mm, screen resolution 1280 x 800 pixels.

3.6 Experimental design

Screens of the smart phones was divided into set of regions represented by letters starting from letter "A" depending on the number of regions of interest on the screen. Eye tracker generates the scan path by following the participant's eye movement that visits different regions to find the given target. "ABBCHEF" can be an example of a participant visiting these regions in sequence. Participants were asked to perform the search tasks based on the smart phone's interface. Each participant was asked to find the answer of the question (target) as shown in Table 1. The questions were designed carefully and has only a specific answer, as it is recommended by Broder in terms of finding specific information on a web page [17]. In addition; one of the ways used to collect data from the participants is to instruct the participant to find a given target according to a given instructions [6]. In our study, each participant will be asked to find a certain target. After finding the target, the participant needs to give an indication to move to the second task. All participants were tested individually in a specific lab set up for conducting the experiments. Before conducting each experiment, the aim of the study was explained to each participant as well as a full description for the smart phone applications, this will be followed by eye tracker calibration to start the experiment.

4 Experimental results and discussions

To investigate if scan paths are age-driven, and to investigate if scan paths are smart phone screen size-driven or stimulus driven, we examined the distances between all scan paths for each of the nine applications, for all age groups and for all three smart phone screen sizes separately. The algorithm used to calculate the distances between pairs of scan paths is Levenshtein algorithm which finds the minimum cost to transform one string into another. Where is the small distance value indicates to low dissimilarity and the large distance value indicates to high dissimilarity as explained in the example in Table 2 below [9]:

Table 2. Example of Levenshtein distance calculation.

	A	F	B	F	F	D	C	D	F
A	0	1	2	3	4	5	6	7	8
B	1	1	1	2	3	4	5	6	7
C	2	2	2	2	3	4	4	5	6
F	3	2	3	2	2	3	4	5	5
E	4	3	3	3	3	3	4	5	6
F	5	4	4	3	3	4	4	5	5
F	6	5	5	4	3	4	5	5	5
G	7	6	6	5	4	4	5	6	6
D	8	7	7	6	5	4	5	5	6
C	9	8	8	7	6	5	4	5	6

The formula of Levenshtein algorithm is defined in equations (1) and (2) below:

$$A[i][j] = \min \begin{cases} A[i-1][j\quad\ ] + 1 \\ A[i\quad\ ][j-1] + 1 \\ A[i-1][j-1] + C(i,j). \end{cases} \quad (1)$$

The first and second terms of equation (1) the minimization handles the costs of deletions and insertions, and the third term handles substitutions, with:

$$c(i,j) = \begin{cases} 0, s1\ [i-1] = s2[j-1] \\ 1, otherwise. \end{cases} \quad (2)$$

The Levenshtein distance, "LDIS", between the participants of each application is presented by "A" in equation (1).

As we have in our study; 3 screen sizes: (small, medium, large), 2 experiments: (EXP1 and EXP2), and 2 age groups: (younger, elderly). Where each of EXP 1 and EXP2: has 9 applications.

As example of calculating the mean and STD for all participants of YG1 in small screen size, we followed the following steps. First: equation (3) shows the way of calculating the distance between pairs of scan paths for one application. As an example of this; if we want to find the distance between pairs of scan paths of two participants in younger age group for one application, we will bring the scan path for first participant (i.e., *scanpath(YG1(first participant))*) and compare it with another from the second participant (i.e., *scanpath(YG1(second participant)))*:

$$LDIS_{i,j} = \left(scanpath\left(YG1(i)\right), scanpath\left(YG1(j)\right)\right) \quad (3)$$

Second; after finding the distances between all scan paths in each application of 9 applications between all participants. We find the mean of (LDIS's) for each application represented by (*meanApp*) as shown in equations (4):

$$meanApp = \left(\frac{1}{n}\right) * \left(\sum_{k=1}^{n} \mathrm{LDIS}_{i,j}(l,k)\right) \quad (4)$$

Where *i* and *j* are participants in age group YG1, *n* is the number of calculated distances for each application, and *l* is number of applications.

Then; third: we calculated the mean and STD for *meanApp* for 9 applications as shown in Table 3.

We based on mean and standard deviation in analysis our data, and in our future work we will improve our study from analysis side by using different type of statistical analysis. The analysis of this study was conducted in two ways. First, we calculated the mean and standard deviations of the distances for each age group as illustrated in Fig. 1. Table 3 shows the mean and STD values of distances, and Fig. 2, Fig. 3, and Fig. 4 below show the minimum and maximum mean distances of two age groups for three smart phones. The results reveal that the dissimilarities of scan paths of younger age group in both of experiments (i.e., EXP1 & EXP2) as well as in the three screen sizes are smaller than the dissimilarities associated with the elderly group. These results indicate that scan paths are age-driven in scanning smart phone applications. We

have not come across any other study that considered scan path similarity to investigate if scan paths on smart phones are age-driven.

Secondly, we considered each experiment (i.e. EXP 1, EXP 2) individually for the three smart phone screen sizes to investigate if scan paths are stimulus-driven or smart phone screen size-drive. By taking the two smallest distances of the nine applications for each screen size of each experiment, we see that scan paths of App 7 has the smallest mean distance for two out of the three smart phone sizes of EXP 1 (both age groups). In EXP 2, scan paths of App 3 has the smallest mean distance for two out of the three smart phone screen sizes. Table 4 shows the mean scan path distances of 9 apps for large screen size in EXP 1. Each of App 7 in EXP 1 and App 3 in EXP 2 was repeated 4 times to be one of two smallest distances for each that considered to be the larger number in EXP 1 and EXP 2

This observation indicates that there is no influence of smart phone screen sizes on the increasing of scan path dissimilarity. In other words, scan paths are stimulus-driven than smart phone screen size-driven.. However, there was influence by stimulus as it was for App 7 of EXP 1 and App 3 of EXP 2, shown in Table 1, for all three smart phone screen sizes that has the smallest scan path distance. These finding are in line with the previous work in [14] and [12] where they found scan paths to be stimulus-driven.

Table 3. Scan paths for all age group for three smart phone sizes

		EXP 1		EXP 2	
Screen size	Age group	Mean	(STD)	Mean	(STD)
Small	YG 20-39	4.41	(3.31)	3.00	(0.85)
	EG 60+	10.37	(6.45)	6.93	(4.14)
Medium	YG 20-39	8.33	(4.30)	3.89	(1.90)
	EG 60+	13.44	(5.61)	10.35	(4.59)
Large	YG 20-39	7.33	(4.44)	5.04	(2.84)
	EG 60+	11.81	(3.61)	10.44	(2.19)

Table 4. Mean scan path distances of 9 apps for large screen size in EXP 1.

App No	1	2	3	4	5	6	7	8	9
Mean of Distance (YG)	2.67	13.67	13.00	4.67	12.33	4.33	3.00	7.00	5.33
Mean of Distance (EG)	11.33	14.00	15.00	16.33	12.33	8.33	4.33	12.33	12.33

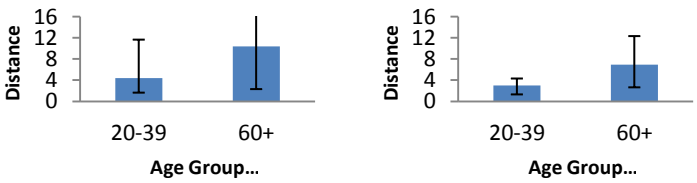

Fig. 2. Minimum and maximum mean distances for each group (high dissimilarity for elderly compared with younger) for small smart phone.

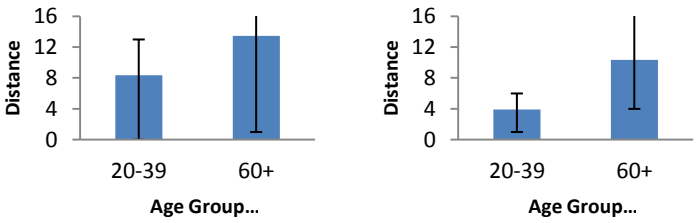

Fig. 3. Minimum and maximum mean distances for each group (high dissimilarity for elderly compared with younger) for medium smart phone.

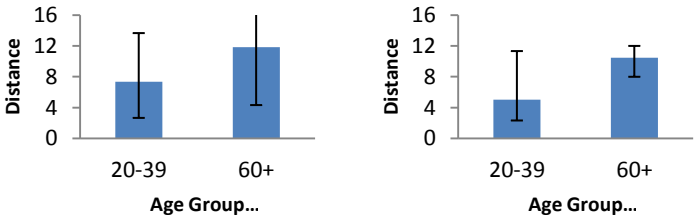

Fig. 4. Minimum and maximum mean distances for each group (high dissimilarity for elderly compared with younger) for large smart phone.

5 Conclusion and future work

This study presented an experiment that combines evidence of viewing behavior of participants using smart phones, with a comparison between the scan paths for younger and elderly people.

We analyzed scan-path data by Levenshtein algorithm. To the best of our knowledge, this is the first paper to adopt age-driven in smart phones for applying the string-edit method on eye-tracking data. In using string-edit analyses, we were able to demonstrate evidence of higher dissimilarity of scan paths for elderly age group than younger age group for each experiment of smart phone screen sizes.

Furthermore, there was a small distances (lower dissimilarity) of scan paths for the applications with popular contents such as active alarm application and account holder picture on the Skype application, which suggests that viewing patterns tend to be more stimulus-driven than smart phone screen size-driven. For future work we would like to focus on the geographical distance of fixations for the substitution cost.

Acknowledgments: Many thanks to Dr David Windridge from University of Surrey for providing us the eye tracker machine to collect data for our study.

References

1. R. Fukuda and H. Bubb, "Eye tracking study on Web-use: Comparison between younger and elderly users in case of search task with electronic timetable service," p. 27, December 2010.
2. Z. Dongfang and S. Qiang, "A discussion based on a design of cell phones usability by the elderly in China," 2009.
3. R. J. Jacob, "Eye Tracking in Advanced Interface Design," p. 53, 2006.
4. T. Miyoshi and A. Murata, "Usability of input device using eye tracker on button size, distance between targets and direction of movement," 2001.
5. L. Cooke, "Is Eye Tracking the Next Step in Usability Testing?," *0-7803-9778-9/06/$20.00 ©2006 IEEE.*, p. 7, 2006.
6. R. J. Jacob and K. S. Karn, "Eye tracking in human-computer interaction and usability research: Ready to deliver the promises," *Mind*, vol. 2, no. 3, p. 4, 2003.
7. T. Takeuchi and Y. Habuchi, "A Quantitative Method for Analyzing Scan Path Data Obtained by Eye Tracker," 2007.
8. F. Cristino, S. Matht, J. Theeuwes and I. D. Gilchrist, "ScanMatch: A novel method for comparing fixation sequences," *Behavior Research Methods*, vol. 42, no. 3, pp. 692-700, 2010.
9. A. T. Duchowski, J. Driver, S. Jolaoso, W. Tan, B. N. Ramey and A. Robbins, "Scanpath comparison revisited," in *ACM*, 2010.
10. R. Dewhurst, M. Nystrm, H. Jarodzka, T. Foulsham, R. Johansson and K. Holmqvist, "It depends on how you look at it: Scanpath comparison in multiple dimensions with MultiMatch, a vector-based approach," *Behavior research methods*, vol. 44, no. 4, pp. 1079-1100, 2012.
11. C. Privitera and L. Stark, "Algorithms for defining visual regions-of-interest: comparison with eye fixations," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 22, no. 9, pp. 970-982, 2000.
12. S. Josephson and M. E. Holmes, "Attention to repeated images on the World-Wide Web: Another look at scanpath theory," *Behavior Research Methods, Instruments, \& Computers*, vol. 34, no. 4, pp. 539-548, 2002.
13. H. Takeuchi and N. Matsuda, "Scan-path analysis by the string-edit method considering fixation duration," 2012.
14. A. Al-Wabil, P. Zaphiris and S. Wilson, "Examining visual attention of dyslexics on web navigation structures with eye tracking," 2008.
15. D. Nettleton and C. Gonzalez-Caro, "Analysis of User Behavior for Web Search Success Using Eye Tracker Data," 2012.
16. SR_Research.Ltd, "EyeLink? 1000 Installation Guide Tower, Desktop, LCD Arm, Primate, and Long Range Mounts Remote, 2000 Hz and Fiber Optic Camera Upgrades," 2010.
17. A. Broder, "A taxonomy of web search," in *ACM*, 2002.

A Candour-based Trust and Reputation Management System for Mobile Ad Hoc Networks

Eric Chiejina, Hannan Xiao and Bruce Christianson

School of Computer Science,University of Hertfordshire,Hatfield, Herts, UK

Email:{e.chiejina, h.xiao, b.christianson}@herts.ac.uk

Abstract. The decentralized administrative controlled-nature of mobile ad hoc networks (MANETs) presents security vulnerabilities which can lead to attacks such as malicious modification of packets. To enhance security in MANETs, Trust and Reputation Management systems (TRM) have been developed to serve as measures in mitigating threats arising from unusual behaviours of nodes. In this paper we propose a candour-based trust and reputation system which measures and models reputation and trust propagation in MANETs. In the proposed model Dirichlet Probability Distribution is employed in modelling the individual reputation of nodes and the trust of each node is computed based on the node's actual network performance and the quality of the recommendations it gives about other nodes. Cooperative nodes in our model will be rewarded for expanding their energy in forwarding packets for other nodes or for disseminating genuine recommendations. Uncooperative nodes are isolated and denied the available network resources. We employed the Ruffle algorithm which will ensure that cooperative nodes are allowed to activate sleep mode when their service is not required in forwarding packets for its neighbouring trustworthy nodes. The proposed TRM system enshrines fairness in its mode of operation as well as creating an enabling environment free from bias. It will also ensure a connected and capacity preserving network of trustworthy nodes.

1 Introduction

Unstructured networks are networks with a decentralized control of operations. Such networks lack centralized infrastructure and administration. Mobile ad hoc networks (MANETs) are unique examples of unstructured networks. MANETs are characterized by limited bandwidth and are less efficient unlike wireless networks with a centralized administration. Typically, a MANET is prone to eavesdropping, high security threats, rapid and continuous changes in network topologies due to nodes mobility [1]. Due to these distinguished features, all network nodes in a MANET must act as a router, server and client [2], mandating these nodes to collaborate for the effective and efficient operations of the network. Specialised network protocols have been employed in network layer of nodes in MANETs to ensure cooperation among nodes. Moreover, it is usually assumed that all the network nodes will act in accordance to the application and protocol specifications. However, due to limited resources or anomalous behaviours of some nodes, these assumptions are not always true. Network nodes sometimes make local decisions on whether to follow the network basic operations or not. These nodes may decide to act either selfishly by not forwarding packets or maliciously by advertising false routes [3]. Such an abrupt change in a node's behaviour may result in reduced network efficiency and

increased susceptibility to attacks. Therefore, a trust management system that ensures an effective and reliable collaboration of all network nodes in a MANET is essential. These systems would ensure that network nodes build a good reputation and attain a certain level of trust before such nodes can effectively operate in a network. As a result, there would be a significant reduction or elimination of malicious nodes trying to disrupt the operations of the network. Therefore, this would ensure that legitimate network nodes attain the required goals [3]. The rest of the paper is organised as follows: Section 2 introduces literature about trust and reputation systems in MANETs. In section 3, the concept of the proposed candour-based system is explained. Section 4 concludes the benefits of the proposed system and outlines future work.

2 Literature Survey

Over the past decade, a lot of research works have been proposed and carried out on TRM systems in mobile ad hoc networks which employed Price-based and Reputation-based schemes to enforce cooperation among nodes in the network. The Price-based schemes [4-9] treat packet forwarding as a service which can be paid for and they introduce a form virtual currency to regulate packet-forwarding collaboration among nodes. Most of the price-based schemes require tamperproof hardware [4], [5] or virtual banks that all the nodes in the network can trust [6], [7]. These price-based schemes use the virtual currency as a form of reward to nodes that participate in packet forwarding activities. In the case where a trust authority or virtual bank is required, it requires assistance from a fixed communication infrastructure to implement the reward schemes, which is not applicable for a pure ad hoc network.

On the other hand reputation-based schemes [10-23] employ different monitoring techniques in gathering data which are used in computing the reputation and trust of nodes in the networks. The monitored data can be derived from direct observations of nodes activities or from recommendations from other nodes. These reputation-based systems are geared towards punishing and isolating selfish or malicious nodes in the network by denying these uncooperative nodes the available network resources. The cooperative nodes are allowed to carry on with their normal network activities which are perceived as a reward by these systems as long as they continue to forward packets for other nodes. For example, He at al [15] proposed a secure and objective reputation-based incentive scheme for MANETs. The reputation of nodes in their proposed model is quantified by objective measures, and the propagation of reputation is efficiently secured by one-way-hash-chain based authentication. Their model uses punishment as a way of encouraging packet forwarding and discipline selfish nodes by probabilistically dropping packets that originates from those nodes.

Most of the existing reputation-based schemes suffer from lack of effective mechanisms to measure and propagate reputation and trust in the network. Secondly, the cooperative nodes in these reputation-based schemes are not truly rewarded for continuously expending energy in routing or forwarding packets. The continuous unrewarded cooperation results in low energy levels in these cooperative nodes. This may in turn have an adverse effect on their trust, reputation as well as individual network performance. As a

result, such nodes may end up being punished and isolated from the network when attempting to route or forward packets again. Therefore there is need for a reliable trust and reputation management system that would enforce cooperation by ensuring that collaborative nodes are rewarded for conducting favourable network operations, while selfish and malicious nodes are punished and isolated from the network. Hence, a Trust and Reputation Management system that incorporates punitive and incentive measures in its mechanism will ensure a fair platform for all the nodes in the network. In this paper we proposed a candour-based trust and reputation management system. This candour-based TRM system enshrines fairness in its mode of operation. Furthermore, it creates an unbiased enabling environment, which ensures that nodes are rewarded, isolated or punished based on the individual network behaviours of the nodes. Nodes in the proposed system are given incentive for expending their energy in forwarding packets for other nodes and for disseminating genuine second-hand reports. Our proposed TRM system considers that nodes have limited energy. Its functions cater for situations that will hamper an active nodes performance level due to low energy. It considers the fact that genuine nodes which are unable to forward packets due to low energy may still provide accurate recommendations. These recommendations usually require low amount of energy to action.

3 The Candour-Based TRM System

Fig. 1 shows the overview of the proposed candour-based TRM system. The following subsections explain the various module of the system.

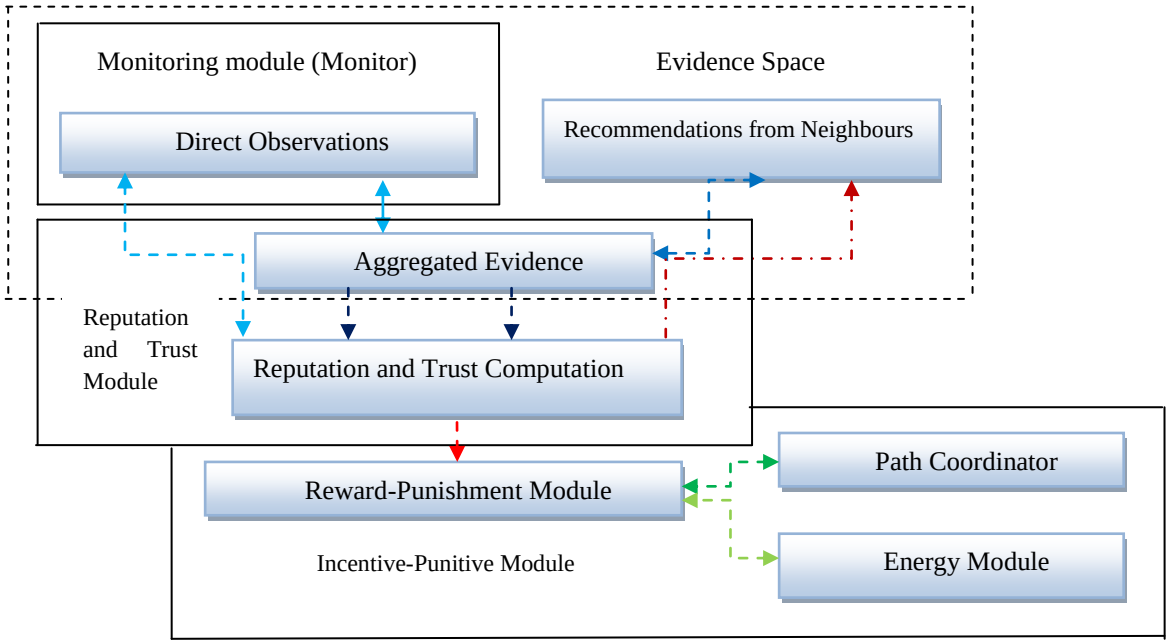

Fig. 1. The schematic diagram of the candour-based TRM system

3.1 The Monitor Module

The monitoring module comprises entirely of the monitor which is an essential part of the proposed TRM system. It specializes in detecting and reporting successful and unsuccessful packet forwarding activities and malicious modification of packets. It also ensures that nodes are not unfairly penalized for unintentionally dropping packets whereas the actual cause may be due to packet collision. To ensure the viability of the proposed monitoring process that will be carried out by the monitor, the monitor will only observe the activities of nodes that are 1-hop away, and each node will have the ability to carry out Omni-directional transmission. The monitor incorporates the packet acknowledgements and packet precision techniques in its mode of operations. It captures packets through listening of transmissions in promiscuous mode. Through monitoring of passive acknowledgements and the packet precision method, a node will be able to determine if its next hop neighbour is exhibiting any of the following behaviours;

i. Carrying out a packet modification attack if the data contents have been dishonestly modified
ii. Effectuating latency delays by retarding the retransmission of packets
iii. Displaying a selfish behaviour by not forwarding a packet
iv. Carrying out a prevarication attack if a self-induced fallacious packet is transmitted
v. Acting like a black hole if the packet intended for forwarding is not retransmitted or dumped.
vi. Launching an impersonation attack if the IP addresses or the MAC addresses have been spoofed.

Fig. 2. The Internal structure of the Monitor

The monitor detects, investigates and registers abnormal behaviours of nodes and passes the direct observations and recommendations to the reputation and trust modules for evaluation and computation. The next section will explain how this will be carried out.

3.2 Reputation Computation

Nodes continuously observe the behaviours of their neighbours that are 1-hop away and compute a reputation value for the successful observations carried out. The reputation of nodes in the network is computed using the Dirichlet Probability Distribu-

tion. The Dirichlet Probability Distribution was chosen over other distributions because it provides a sound and flexible platform suitable for designing a practical reputation system [24]. Dirichlet Distribution is able to differentiate a very large amount of negative reports from large positive reports. It is also useful in implementing reputation reports with grade levels, i.e. very bad – bad – uncertain – good – very good. This will enable nodes in the proposed TRM system to evaluate and decides which recommendation to integrate in computing the total aggregated reputation of a node in the network. The Dirichlet Probability Density Function (*PDF*) for a set of possible outcomes and for positive real parameters can be defined as [24, 25]:

$$f(\tilde{p} \mid \tilde{\alpha}) = \frac{\Gamma[\sum_{i=1}^{k} \alpha_i]}{\prod_{i=1}^{k} \Gamma[\alpha_i]} \prod_{i=1}^{k} p(\theta_i)^{\alpha_i - 1} \quad (1)$$

Where $\tilde{p}$ represents the set of possible outcomes given by

$$\tilde{p} = \{p(\theta_i) | 1 \leq i \leq 3\} \quad (2)$$

such that $p(\theta_1)$ may represents the probability of forwarding packets, and $p(\theta_2)$ may represents the probability of dropped packets and $p(\theta_3)$ may represents the probability of maliciously modified packets.
$\tilde{\alpha}$ represents a set of positive real parameters such that

$$\tilde{\alpha} = \{ \alpha_i | 1 \leq \ i \ \leq 3 \} \quad (3)$$

The parameter α_i can be interpreted as the prior observation counts of the possible outcomes such that α_1 may represents the number of successfully observed packet forwarding, α_2 may represents the number of successfully observed packet dropping and α_3 may represents the number of successfully observed malicious modification of packets.

The reputation of a node in the network $\Re$, can be determined by the probability expectation of the Dirichlet Distribution given by the equation below [24, 25]:

$$\Re = \mathcal{E}(p(\theta_i) | z_i, q_k) = \frac{z_i + Cq_k}{C + \sum_{i=1}^{k} z_i} \quad (4)$$

$$\alpha_i = Cq_k + z_i \quad (5)$$

α_i can be interpreted as prior observation counts for the possible outcomes of the observed events out of a *k* possible events. *C* is the cardinality of the state space over which a uniform distribution is assumed. $z_i = (z_1 \ldots\ldots\ldots z_k)$ represents the accumulated evidence over the observed elements of the state space and q_k is the base rate parameter over the state space. In case where no evidence is available, the base rate alone determines the probability distribution of the events. (e.g. the case of new nodes in the network) [24, 25]. As more evidence becomes available as a result of

observations, the influence of the base rate diminishes; it reaches a point where the evidence alone determines the probability distribution of the events.

3.2.1 Evaluation of Recommendations from Neighbouring Nodes

Nodes rely on the recommendations from its neighbours in evaluating the total reputation value of a node. To avoid the effect of false second-hand reports affecting the reputation value, a deviation test will be carried out to determine the validity of the recommendation. The result of the deviation test will affect the reputation and trust value of the recommending node positively or negatively. This is similar to the work carried out in [17, 23]. For instance, if the reputation value of a node ***A*** on a subject node ***B*** is given as $\Re_1$ and the recommendations of node ***B*** from node ***C*** is given as $\Re_2$, the deviation test can be evaluated using difference in the expectation value of the Dirichlet Probability Distribution. Let ϑ be the deviation for the test;

$$\Longrightarrow |\{\mathcal{E}(p(\theta_i)|\, z_i, q_k)_{\,2}\} - \{\mathcal{E}(p(\theta_i)|\, z_i, q_k)_{\,1}\}| \geq \vartheta \tag{6}$$

Where $\mathcal{E}(p(\theta_i)|\, z_i, q_k)_1$ and $\mathcal{E}(p(\theta_i)|\, z_i, q_k)_2$ are the expectation values of $\Re_1$ and $\Re_2$ respectively. ϑ is always positive and acts as the threshold validating recommendations from other nodes.

3.2.2 Aggregating Direct Observations and Recommendations

To compute total reputation of node ***A*** about a subject node ***B*** after a certain period i.e. t + 1, the reputation derived from direct observations and the recommendations from other nodes are aggregated to give a final reputation value. This implies that the total aggregated reputation of node ***A*** about node ***B*** is given by;

$$\Re_{ab(t+1)} = \delta\, \Re_{ab(t)} + \varphi\, \vec{r}_{b(t+1)}\ , \quad 0 \leq \delta \leq 1 \tag{7}$$

Where $\vec{r}_{b(t+1)}$ is the sum of all the recommendations from node ***A*** 1-hop neighbours about node ***B*** during a given period t+1. $\Re_{ab(t)}$ is the current reputation value. δ is the decaying factor which controls the rate at which old reputation value decays after a given period, and it's such that $\delta \in [0, 1]$. φ is a small positive weight which acts as a discount factor. After n periods of time, the total aggregated reputation of node ***A*** about node ***B*** can be given as;

$$\Re_{ab(t+n)} = \delta^n\, \Re_{ab(t)} + \varphi\, \vec{r}_{b(t+n)}\ ,\ 0 \leq \delta \leq 1 \tag{8}$$

3.3 Trust Evaluation

The Trust evaluation of a node in the network is a combination of the aggregated reputation value and the accuracy of the node's recommendations about other nodes. It is denoted as $T\langle \Re, \omega \rangle$ which is a combination of two factors. $\Re$ denotes the trustworthiness of the node based on is the reputation as calculated in 8 which represents

the node's actual network operations, while ω will denote the trustworthiness based on the accuracy of the recommendations a node makes about other nodes.

3.3.1 Computation of Accurate Recommendations

The accuracy value of the recommendations made by a node denoted as χ, can be defined as follows

$$\chi \triangleq \frac{\eta}{\eta + \gamma}; \quad \text{for } 0 \leq \chi \leq 1 \tag{9}$$

Where η is the cumulative number of recommendations that are correct and γ is the cumulative number of recommendations that are incorrect. A value of χ = 1 indicates absolute accuracy, and a value of χ close to zero indicates low accuracy.
The confidence value, ϱ associated with the accuracy value χ is defined as [20];

$$\varrho = 1 - \sqrt{\frac{12\eta\gamma}{(\eta+\gamma)^2\,(\eta+\gamma+1)}}, \text{ where } 0 \leq \varrho \leq 1 \tag{10}$$

A value of ϱ close to 1 indicates high confidence in the preciseness of the computed accuracy value, while a value of ϱ close to 0 indicates low confidence in the computed accuracy value. The trustworthiness of a node based on the accuracy of its recommendations about other nodes can be given as a pair of the accuracy value, χ and the confidence value, ϱ, which is similar to the notion of trust evaluation of nodes based on packet forwarding activities applied in [19,20].
The trustworthiness evaluation associated with the pair (χ, ϱ) can be defined as;

$$\omega(\chi,\varrho) \triangleq 1 - \frac{\sqrt{\frac{(\chi-1)^2}{m^2} + \frac{(\varrho-1)^2}{n^2}}}{\sqrt{\frac{1}{m^2} + \frac{1}{n^2}}} \tag{11}$$

where m and n are parameters that determine the relative importance of the accuracy value and the confidence value.

3.3.2 Total Trustworthiness of a Node

The total trustworthiness of a node is computed by combining the trustworthiness based on the reputation of the node in terms of what it does .i.e. forwarding packets, and in terms of the accuracy of its recommendation as defined in equation (10).
This implies that the total trustworthiness, $\mathbf{T}\,\langle \Re, \omega(\chi,\varrho)\rangle$ can be given as;

$$\mathrm{T}\,\langle \Re, \omega(\chi,\varrho)\rangle \;=\; [\Re,\ \omega(\chi,\varrho)\,] \tag{12}$$

3.4 The Reward and Punishment Scheme

The reward and punishment scheme ensures that collaborating nodes are rewarded for effectively carrying out network operations, while selfish and malicious node will be denied network resources and isolated from the network. After the reputation and

trust of a node in the network has been computed, nodes that are found to have trust and reputation values below the given threshold value are classified to be untrustworthy, while the nodes that have trust and reputation values that are above the threshold value are classified as trustworthy nodes in terms of their actual network activities and in terms of their recommendations about other nodes. The computed trust and reputation values of nodes are stored in a trust table. These values are periodically updated when new values are computed from newly accumulated observations. The Path Coordinator is responsible for isolating and denying misbehaving (untrustworthy) nodes the available network resources. It accesses the trust table before making a routing decision to ensure that untrustworthy nodes are eliminated from the routing paths. It also ensures that any packets that originate from those untrustworthy nodes are rejected. This ensures that only paths with trustworthy nodes are used for routing or forwarding packets. On the other hand, the reward scheme ensures that nodes that are found trustworthy are able to activate an idle period. It works with the observation that when a region of the network has a sufficient density of trustworthy nodes, only a small number of the nodes needs to be on at any time to forward traffic for active connections. The reward scheme decision is based on an estimation of how many of its trusted neighbours will benefit from it being awake, and the amount of energy available to it. The scheme employs a Ruffle algorithm [26] which ensures minimum power assignment for each trustworthy node such that symmetric connectivity is preserved.

For instance given a region of trustworthy nodes T_n = (W,E,c) with maximum power assigned to each node. The Ruffle algorithm aims to find a minimum power assignment for each of the trustworthy nodes in the network such that the symmetric connectivity in T_n is preserved while packet forwarding and routing remains effective and efficient. The algorithm is as follows;

- *Assign to each of the trustworthy node an ID based on the energy level. This information is gotten from the energy module.*
- *The trustworthy nodes are then sorted by their ID.*
- *For each trustworthy node W_i(i = 1, n), find the number of connected trustworthy nodes in its neighbourhood. This information can be derived from the path coordinator.*
- *Find the distance to the closest trustworthy node that has an ID great than the ID of W_i for each connected trustworthy nodes to W_i.*
- *Find the distance S_d = distance of the furthest of all the closest trustworthy nodes.*
- *Reduce the range of W_i to S_d.*

The Ruffle algorithm also aims to reduce energy consumption on sending packets for trustworthy nodes which are wake up. With the successful implementation of the Ruffle algorithm, the reward scheme will ensure a connected and capacity preserving network of trustworthy nodes.

4 Conclusion and Future Work

This paper proposes a candour-based trust and reputation management system for mobile ad hoc network which will ensure that selfish and malicious nodes are eliminated and denied network resources while the trustworthy nodes are rewarded for forwarding packets. The proposed system will employ a reward scheme that allows trustworthy nodes to randomly activate idle time when their service is not required. This will preserve their energy and in turn prolong the life span of the network of trustworthy nodes. Future work comprises of the full implementation of the proposed model using C++ and NS 2.34, evaluating the effectiveness of the implemented model in detecting misbehaving nodes and rewarding trustworthy nodes.

REFERENCES

[1] J.H. Cho, A. Swami and I.R. Chen, "Mission-Dependent Trust Management in Heterogeneous Military Mobile Ad Hoc Networks," 15th International Command and Control Research and Technology Symposium, Santa Monica, California, June 2010.

[2] P. B. Velloso, R. P. Laufer, D. O. Cunha, O. Carlos, M. B. Duarte, and G. Pujolle, "Trust Management in Mobile Ad Hoc Networks Using a Scalable Maturity-Based Model", in Proc. IEEE Transaction on Network and Service Management, vol.7, no.7, September 2007.

[3] G. Theodorakopoulos and J. S. Baras. " Malicious Users in Unstructured Networks", In Proceedings of the IEEE International Conference on Computer Communications, A S. Ganeriwal and M. B. Srivastava, "Reputation-based framework for high integrity sensor networks" Anchorage, AK, USA, May 2007, p. 884 – 891

[4] L. Buttyan and J. Hubaux, "Enforcing service availability in mobile adhoc WANs," IEEE/ACM Workshop on Mobile Ad Hoc Networking and Computing (MobiHoc), Boston, MA, USA, August 2000.

[5] L.Buttyan and J. P. Hubaux, Stimulating Cooperation in Self-Organizing Mobile Ad Hoc Networks? in Technical Report No. DSC/2001/046, August 2001.

[6] M. Jakobsson, J. Hubaux, and L. Buttyan, "A micro-payment scheme encouraging collaboration in multi-hop cellular networks," Proceedings of Financial Crypto 2003, Gosier, Guadeloupe, January 2003.

[7] S. Zhong, J. Chen, and Y. Yang, "Sprite: a simple, cheat-proof, creditbased system for mobile ad-hoc networks," IEEE INFOCOM 2003, San Francisco, CA, USA, April 2003

[8] M. Felegyhazi, L. Buttyan, and J. Hubaux, "Equilibrium analysis of packet forwarding strategies in wireless ad hoc networks - the static case", in Proc. Personal Wireless Communication (PWC), Venice, Italy, September 2003.

[9] J. Cai, U. Pooch, "Allocate fair payoff for cooperation in wireless ad hoc networks using Shapley Value", Parallel and Distributed Processing Symposium, 2004. Proceedings. 18th International , 26-30 April 2004

[10] S. Marti, T.J. Giuli, Kevin Lai, and Mary Baker, "Mitigating routing misbehaviour in mobile ad hoc networks". In Proceedings of MOBICOM 2000, pages 255–265, 2000

[11] P. Michiardi and R. Molva, CORE: A COllaborative REputation mechanism to enforce node cooperation in mobile ad hoc networks, in Proc. 6th Int. Conf. Commun. Multimedia Security, 2002, pp. 107–121

[12] S. Buchegger and J. Y. Le Boudec, “Node Bearing Grudges: Towards Routing Security, Fairness, and Robustness in Mobile Ad Hoc Networks,” Proc. IEEE 10th Euromicro Workshop on Parallel, Distributed, and Network-based Processing , pp. 403-410, Canary Islands, Spain, January 2002.

[13] S. Bansal and M. Baker, “Observation-Based Cooperation Enforcement in Ad-Hoc Networks”, Technical Report, Computer Science Department, Stanford University, CA, July 2003.

[14] M. Virendra, M. Jadliwala, M. Chandrasekaran, and S. Upadhyaya, “Quantifying trust in mobile ad-hoc networks," in Proc. IEEE International Conf. Integration Knowledge Intensive Multi-Agent Syst., Waltham, USA, April 2005

[15] Q. He, D. Wu, and P. Khosla, “A secure incentive architecture for ad hoc networks," Wireless Commun. Mobile Comput., vol. 6, no. 3, pp. 333- 346, May 2006.

[16] J. Li, R. Li, and J. Kato, “Future Trust Management Framework for Mobile Ad Hoc Networks: Security in Mobile Ad Hoc Networks,” IEEE Community Magazine, vol. 46, no. 4, April 2008, pp. 108-114.

[17] S. Buchegger and J.-Y. L. Boudec, ”Performance analysis of the CONFIDANT Protocol”, in Proc. 3rd ACM Int. Symposium on Mobile Ad Hoc Network Computing, Lausanne, Switzerland, 2002, pp. 226–236

[18] J. Hu, “Cooperation in Mobile Ad Hoc Networks”, Technical report (TR-050111), Computer Science Department, Florida State University, Tallahassee. January 2005.

[19] C. Zouridaki, B. L. Mark, M. Hejmo, and R. K. Thomas, “A Quantitative Trust Establishment Framework for Reliable Data Packet Delivery in MANETs,” in In Proceedings of the 3rd ACM Workshop on Security of Ad Hoc and Sensor Networks (SASN’05), pp. 1–10, November 2005

[20] C. Zouridaki, B. L. Mark, M. Hejmo, and R. K. Thomas, “Robust Cooperative Trust Establishment for MANETs,” Proc. 4th ACM Workshop on Security of Ad Hoc and Sensor Networks, Alexandria, VA, 30 October 2006, pp. 23-34.

[21] A. A. Pirzada and C. McDonald, “Trust establishment in pure ad-hoc networks,” Wireless Personal Communications: An International Journal, vol. 37, no. 1-2, April 2006, pp. 139–168

[22] A. Boukerche and Y. Ren, “A Security Management Scheme using a Novel Computational Reputation Model for Wireless and Mobile AdHoc Networks,” Proc. Int’l Workshop on Modelling Analysis and Simulation of Wireless and Mobile Systems, Vancouver, British Columbia, Canada, pp. 88-95, 2008.

[23] S. Buchegger and J. Y. Le Boudec. "A Robust Reputation System for Mobile Ad-hoc Networks". In Proceedings of IEEE/ACM Symposium on Mobile Ad Hoc Networking and Computing (MobiHOC), Lausanne, CH, June 2005.

[24] Josang, Audun; Haller, Jochen; "Dirichlet Reputation Systems," Availability, Reliability and Security, 2007. ARES 2007. The Second International Conference on, vol., no., pp.112-119, 10-13 April 2007

[25] Aboulwafa, S.; Bahgat, R.; "DiReCT: Dirichlet-based Reputation and Credential Trust management," Informatics and Systems (INFOS), 2010 The 7th International Conference on , vol., no., pp.1-8, 28-30 March 2010

[26] Brinza, D., Derado, G., Aznita, R., Li, Y. and Zelikovsky, A. (2004) `Energy efficient protocols for ad-hoc networks', Georgia Electronic Design Center Industry Advisory Board (GEDC IAB'04), ppt. Retrieved from http://cs.gsu.edu/~dima/dima/cv/index.html

Domain-Specific Languages for Web Application Development Processes

Athanasios Zolotas

Department of Computer Science, University of York, UK
and
Digital Lightspeed Solutions Ltd.
amz502@york.ac.uk

Abstract. The continuous evolution of the Web has led to increased ambition, and hence the construction of more complex web applications. Many web applications that are developed do not fulfil the actual needs of the clients. This is a result of poor requirements elicitation, poorly defined requirements and incomplete or inadequate testing. Many solutions have been proposed to help the better elicitation and management of the requirements, and better testing of web applications, but these approaches are not widely used in practice. We argue for the development of a set of Domain-Specific Languages (DSLs) that will be deployed across the different phases of the web application development lifecycle to facilitate the production of correct web applications that deliver on the *contract* defined between clients and developers.

Keywords: Web Applications, Project Failures, Model-Driven Engineering, Domain Specific Languages, Requirements Engineering

1 Introduction

Web applications are evolving. Simple websites with static pages have transformed into more complex and dynamic web applications. A typical web application handles the data given by the users to produce relevant content. This increasing complexity of web applications, the introduction of HTML 5[1] in 2011 and the fact that users have higher demands, changed the way that the requirements of an application are elicited. The requirements engineering processes used in developing web applications are similar to those used in developing conventional software. Usually, a basic set of requirements of an application can be written in plain text and given by the client to the developers. These documents, known as *tendered contracts*, are commonly used when clients want to make 'first contact' with candidate companies and describe to them the desired *workflows* and *business processes* that the application should fulfil. The candidate companies propose a strategy for developing the solution, along with a time and cost estimation for the described product, and the client selects the one that looks the most appropriate.

[1] http://dev.w3.org/html5/spec/

Model-Driven Engineering (MDE) promotes models as the basic elements in the software engineering process. Models are artefacts that can be used to describe a system. The semantics and the syntax of a model conform to a *metamodel*; metamodels are a special kind of model that is used in the definition of languages. [1].

Programming languages can be divided into two categories; the general-purpose programming languages and the domain-specific programming languages (DSLs). The difference between these two is that the latter is built to be used in a specific domain. For instance, JAVA can be used to describe systems from different domains while SQL can be only used for the interaction with a database. A proposed advantage of using a DSL is that it allows better understanding of the written code, especially by experts of the domain [2]. Another claimed advantage is that DSLs promote better communication between the experts in the domain [2]. Every programming language, and so every DSL, is an instantiation of a specific metamodel. That metamodel describes the syntax and the semantics (the rules) of the programming language.

This position paper presents work in progress on developing a family of DSLs that will facilitate the web application development processes, focusing on producing the correct software that the clients asked for. It is structured as follows. In Section 2, the problem and the scope of the project are presented. In Section 3, the related work and the current available solutions to the problem are given. Section 4 describes the ideas on a potential solution while Section 5 summarises the work carried out so far. Section 6 contains the plan that will be followed in validating the proposed solution. Finally, the expected contributions of the project are discussed in Section 7.

2 Problem

Different studies show that a proportion of the software that is developed fails to completely fulfil client needs, because of inadequate requirements elicitation or ambiguity in the way that the requirements are reported. According to Abbott [3], 40-60% of projects fail due to poorly elicited requirements. Marasco [4] identifies this failure to 50% of the projects claiming that a significant amount of money is spent to develop the wrong software [4]. Regarding web applications, Epner [5] claims that 53% of the web applications built do not include the required functionality while only 16% of them meet the business needs. A survey carried out by Lowe [6] revealed that more than 3 out of 4 web developers indicated that eliciting correctly the requirements is very important. Finally, in [7] 43% of the web-design companies which participated in the survey mentioned that the clarity of the requirements is a problematic issue. In addition to the requirements engineering phase, the testing phase of any software development lifecycle is a milestone where incorrectly implemented features can be identified; acceptance testing can reveal such problems. A set of tools and techniques that can be deployed to these two phases of the software lifecycle will help in the process of building correct software. The objective of this project is to create

a methodology that will allow the development of web applications that fulfil the actual needs of the clients, particularly, as *contracted* between client and developer.

The general scope of this research project is restricted to the development processes of web applications and not of the conventional software. The set of possible requirements is unlimited in classic desktop software and can't be included in a DSL unless this DSL ends up being a new natural or a new programming language. In contrast, web applications and specifically data-driven and static ones, are bounded by the elements that are described in the HTML standard.

3 Related Work

A number of approaches have been proposed to face the problems of ambiguity in requirements and the need of producing software that matches the client's needs. These approaches are described in this section.

The first two techniques identified during the literature review were the Program Design Language (PDL) [8], and the Requirements Statement Language (RSL) [9]. Their idea is to create DSLs that combine natural language with technology terms to describe the *design* of a system which can be translated to working code. Thus both these solutions can be thought of as a form of high level coding of the application. The intended audience is programmers and designers of military applications[8][9]. Other solutions were also proposed to facilitate the requirements engineering phase like the High Ordered Software (HOS) [10], the Structured Analysis and Design Technique (SADT) [11] and the Problem Statement Language (PSL) [12]. They were built to be used by developers to store and manage the already elicited requirements, and not by the clients to express the requirements.

In 2007, Kaindl et al. [13], identifying that the problem of ambiguous requirements still exists and that the available solutions are not used in the industry, proposed a tool that allows the specification of requirements in plain English, but each of the words used should be defined using an embedded dictionary. Their tool was equipped with more functionality for defining user groups and general management of the stored requirements [13].

Different MDE approaches have been proposed to represent elicited requirements in models. The Navigational Development Techniques (NDT) [14] focus on web applications as the authors identified a gap in the available methodologies [14]. Their goal was to provide the techniques to capture, define, analyse and validate the requirements of Web applications. The Web Requirements Metamodel (WRM) [15] was proposed to store the requirements of web applications into models using use case scenarios while word definitions from a dictionary can be used to remove ambiguity. WebRatio [16], based on the Web Modelling Language (WebML) [17], is a suite that is used to design and code (data-driven) web applications. It supports the use of the Business Process Model and Nota-

tion (BPMN) [18] to let clients, helped by business analysts, express the business processes of the application and automatically build them.

Finally, the Interaction Flow Modelling Language (IFML) [19] is an Object Management Group standard adopted in 2013 that is used to model workflows of different types of applications (including web applications). IFML can represent workflows omitting the low level programming details and can be used as part of the WebRatio process described above, substituting the BPMN.

4 Proposed Solution

MDE can be used in addressing the problem described in Section 2. The proposed solution consists of a set of Domain-Specific Languages that will allow clients to express the requirements in *tendered contracts* and define *test cases* that will either confirm that the produced solution fulfils their needs, or will identify where the solution falls short. This process will be aided by different model management techniques like *Model-to-Model* and *Model-to-Text Transformations*.

4.1 A Requirements DSL

Different requirement elicitation techniques are used to produce the requirements specification document. The requirements should be written in a clear and unambiguous way so the developers are sure that the document represents the actual clients' needs. The document is written in plain English language. The challenge in using natural language is that it can lead to ambiguous statements. An artefact which can guarantee that the developers understand the requirements that the clients expressed and that the clients expressed the correct requirements, should be created. Although the solutions described in Section 3 were aimed at that direction, they were not adopted. The problem of producing the wrong software due to poorly elicited requirements still exists. In addition, these solutions were not designed for non-technical users and so the possibility to be used by the clients is in question. Moreover, the majority of these solutions focus on the design and coding of web applications and not in the requirements engineering phase [20].

These gaps lead to the need for a new approach to express the requirements included in tendered contracts. Programming languages can be a solution for that problem as their concrete syntax avoids any ambiguities. A statement written in a programming language can be always translated in the same way. A challenge in this idea is that non-technical clients are not familiar with programming languages. A user friendly version of a programming language, that omits low-level implementation details, could be used to let users describe their requirements. Clients are normally familiar with *concepts* that appear in web applications. Some examples of these concepts are: *User Management*, *User Registration*, *User Login*, *Search in Database*, etc. A DSL based on those concepts can be a potential solution to the usability of the proposed DSL. Fowler [2] claims that DSLs offer a better communication channel between the developers

and the customers as they can be easily used by non-programmers. The communication between developers and customers is a bottleneck in the development of applications and DSLs help to solve that problem [2]. In addition, in order to avoid forcing the users write code using the syntax of the proposed DSL, a drag and drop methodology - arguably one of the easiest ways to interact with a computer - can be used. The clients will also be able to create *workflows* between these concepts to describe the requirements and the business processes of the web application.

Finally, another challenge is to create a DSL that will be able to capture the majority of the user requirements but should fulfil three strong criteria: It should be abstract enough, with well-defined and controlled semantics, in order for it not to become a new programming language, concrete enough to avoid being a new human natural language, and sufficiently different from existing requirements specification languages to justify the effort of building a new DSL.

4.2 A Testing DSL

If the requirements of the web application are expressed using the proposed Requirements DSL, it would be possible to use model transformation techniques to generate test cases that will be later feed into a test execution suite (e.g. Selenium - Webdriver which is specifically built for web applications) to check the correct functionality of the final product. This will first help in the testing process by reducing the time needed to write the test cases. Secondly, it will guarantee that all the requirements are tested and it will let the developers know if they have not fulfilled any of them. Finally, as the definition of the requirements will be the first step in the process, the test cases can be generated immediately letting the developers follow a Test-Driven Development approach.

In addition to the automated test cases generation, clients will be able to refine the test cases using a Testing DSL. This DSL will be used, with the help of Business Analysts, when the automatic generation of test cases is not feasible or is restricted. For instance, in some cases, web applications follow a multi-step process (e.g. when ordering from an online shop). All the steps that belong to the same activity (in this example the activity of completing an order) should be tested in accordance to other (previous or next) steps of the activity. Another example is that in some cases the restrictions that should be tested by the test cases will not be directly written in the requirements. For instance, if one of the requirements declare that an input field will accept only 2-digit integers, this should be translated into 3 test cases, for testing what happens when a 1-digit, a 2-digit or a 3-digit integer is provided. The Testing DSL will provide the functionality needed to aid such cases, where automated generation of test cases is not feasible.

4.3 Applications for the DSLs

The following are some additional applications for the DSLs that will be examined for the doctoral project. The proposed solution is sketched in Figure 1.

6 Athanasios Zolotas

Estimate Approximate Implementation Cost It would be of interest to have an accurate cost and time estimation for the application after the requirements were defined. This can be done by having predefined costs for different concepts (i.e. the cost for implementing a simple database query, the creation of a static page, etc.). These costs can be either saved into a database or can be created using another simple graphical DSL. In any case the requirements stored in the model will be analysed to produce an approximation of the cost and time needed to build the web application based on these values.

Code Generation As discussed in Section 3 WebRatio and WebML can generate code for the web applications described using them. In order to have a working result, extended parametrization is needed. It would be of interest to generate parts of the code for the web application, based on the requirements that were declared using the requirements DSL (e.g. the UI elements). A feasible solution will be the transformation of the requirements models to IFML models so they can be fed to the WebRatio suite. However, the primary scenario in this project is that the code will be fully written by the company's developers.

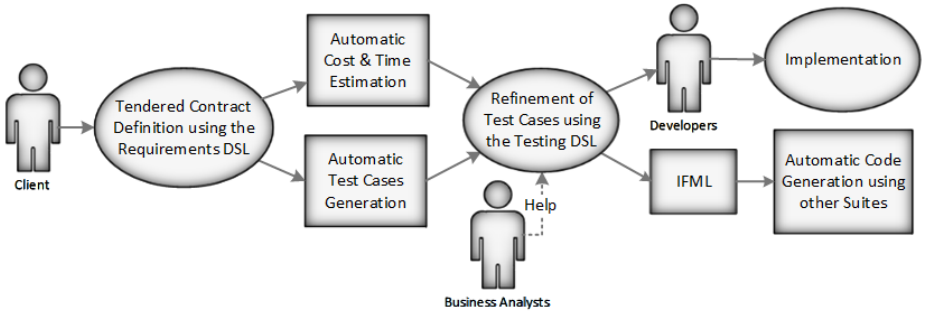

Fig. 1. Diagram of the idea.

5 Preliminary Work and Current Status

The sponsoring organisation has supplied us with a number of case studies that they worked on during the last years. These requirement specification documents and tendered contracts were studied to identify the common concepts that appear in documents like this. These concepts, examples of which are given in Section 4, will be the keywords of the new Requirements DSL. Studying more and more requirements documents will let us identify all the possible concepts that can appear in web applications.

The proposed solution will be facilitated with a tool that will allow the clients "draw" the workflows using the proposed Requirements DSL. A tool that will

look like a prototyping tool and that can be used online through the companies' websites should be developed. During the first period of the research a variety of already existing prototyping tools were evaluated in order to find if there is the possibility of using one of them as the base and extend it to cover the needs of the new DSL. The only tool that covered this need is the Mozilla Pencil[2] which is open-source and can be extended to include the elements of the new DSL. However, it can be deployed and used online by the clients and the format of the file that is exported by the tool omits some details that are necessary for project. As an open-source tool the export format can be re-coded to fulfil our needs. As a second choice, a new editor can be created from scratch based on the Eclipse GMF[3] widely used in MDE.

6 Plan for Evaluation and Validation

The plan for evaluation and validation of the proposed solution will be done following 3 different approaches. In this Section these approaches are described.

First of all, the correctness of the proposed solution should be checked. The proposed solution is built to allow the description of the requirements and the workflows of web applications. In order to verify this, a set of already built web applications will be selected and the application of the requirements DSL on these cases will be tested. During this process we will try identify if the proposed solution is comprehensive enough to express the requirements and workflows of existing applications that belong to the scope set for this project. Examples are complex web applications that are used for online shopping (e.g. Amazon), social networks (e.g. Twitter) or simple data-driven web applications (e.g. Wikipedia).

The second approach in evaluating the proposed solution will involve the developers of the sponsoring organisation. Requirement specifications documents will be given to them written in natural language. The proposed solution will be applied to these case studies and the developers will be asked to evaluate how the new approach helped in better understanding what is the intended behaviour of the web application. This will let us extract direct evidence that the proposed solution fulfils or not the purpose of its existence.

The third approach will involve the clients. They will be asked to use the DSLs to express the requirements of the web application that they need to be developed for them. One of the challenges proposed solution is that it should be used by the clients who are not technology experts. Thus we need to guarantee that it is user-friendly enough and can be used by this category of users.

7 Expected Contributions

As described in Section 2 the problem of producing wrong software that does not match the client's needs still exists although different approaches were proposed

[2] `http://pencil.evolus.vn/`
[3] `http://www.eclipse.org/modeling/gmp/`

to counter it. The proposed approach tries to solve this problem offering a set of DSLs and MDE techniques, some of which are inspired by past solutions, that will be applied though the software development lifecycle.

The sponsoring organisation's experience shows that the production of unacceptable software due to ambiguity of requirements leads to long disputes between the clients and the developers. Following the proposed approach will support the generation of a set of test cases which will confirm if the software satisfies requirements. If these test cases are signed by the client, they can be seen as a contract that includes a set of unambiguous checkpoints that confirm that the produced solution behaves the way that the client asked.

The test cases will be generated automatically just after the specification of the requirements, allowing the company's developers to start implementing the solution following a Test-Driven Development (TDD) approach.

This methodology will allow companies came up with a quick and accurate time and cost estimation as described in Section 4 reducing the time needed to study the tendered contracts and extract, sometimes hidden, requirements.

Last but not least, there is significant potential in automatic code generation. The introduction of IFML, and its eventual conformance with toolsets that can generate code, i.e. the WebRatio suite, affords an opportunity to translate stored requirements into IFML models using Model-to-Model transformations and generate, full or partial, implementations of the web application. The possibility of generating at least the code for the User Interface components is, based on our initial experiments, feasible.

Acknowledgments. This project is funded by the Engineering and Physical Sciences Research Council (EPSRC) as part of the Large Scale Complex IT Systems (LSCITS) initiative. It is carried out in cooperation with Digital Lightspeed Solutions ltd. and is supervised by Prof. Richard Paige from the University of York and Chris Wedgwood from Digital Lightspeed Solutions ltd.

References

1. Warmer, J., Kleppe, A.: The object constraint language: getting your models ready for MDA. Addison-Wesley Longman Publishing Co., Inc. (2003)
2. Fowler, M., Parsons, R.: Domain-specific languages. Addison-Wesley (2010)
3. Abbott, B.: Requirements set the mark. Infoworld (2001) 45–46
4. Marasco, J.: Software development productivity and project success rates: Are we attacking the right problem? `http://www.ibm.com/developerworks/rational/library/feb06/marasco/` (February 2006) [Accessed: 05/07/2013].
5. Epner, M.: Poor project management number-one problem of outsourced e-projects. Research Briefs, Cutter Consortium **7** (2000)
6. Lowe, D.: Web system requirements: an overview. Requirements Engineering **8**(2) (2003) 102–113
7. Lang, M., Fitzgerald, B.: Web-based systems design: a study of contemporary practices and an explanatory framework based on method-in-action. Requirements Engineering **12**(4) (2007) 203–220

8. Caine, S.H., Gordon, E.K.: PDL: a tool for software design. In: Proceedings of the May 19-22, 1975, national computer conference and exposition. AFIPS '75, New York, NY, USA, ACM (1975) 271–276
9. Bell, T., Bixler, D., Dyer, M.: An extendable approach to computer-aided software requirements engineering. Software Engineering, IEEE Transactions on **SE-3**(1) (1977) 49–60
10. Hamilton, M., Zeldin, S.: Higher order software: A methodology for defining software. Software Engineering, IEEE Transactions on **SE-2**(1) (march 1976) 9–32
11. Ross, D., Schoman, K.E., J.: Structured analysis for requirements definition. Software Engineering, IEEE Transactions on **SE-3**(1) (jan. 1977) 6 – 15
12. Teichroew, D., Hershey III, E.: PSL/PSA: A computer-aided technique for structured documentation and analysis of information processing systems. Software Engineering, IEEE Transactions on (1) (1977) 41–48
13. Kaindl, H., Smialek, M., Svetinovic, D., Ambroziewicz, A., Bojarski, J., Nowakowski, W., Straszak, T., Schwarz, H., Bildhauer, D., Brogan, J., et al.: Requirements specification language definition: Defining the ReDSeeDS languages. Institute of Computer Technology, Vienna University of Technology (2007)
14. Jose Escalona, M., Aragon, G.: NDT. a model-driven approach for web requirements. Software Engineering, IEEE Transactions on **34**(3) (May 2008) 377–390
15. Molina, F., Pardillo, J., Toval, A.: Modelling web-based systems requirements using WRM. In Hartmann, S., Zhou, X., Kirchberg, M., eds.: Web Information Systems Engineering WISE 2008 Workshops. Volume 5176 of Lecture Notes in Computer Science. Springer Berlin / Heidelberg (2008) 122–131
16. Acerbis, R., Bongio, A., Butti, S., Ceri, S., Ciapessoni, F., Conserva, C., Fraternali, P., Carughi, G.: Webratio, an innovative technology for web application development. Web Engineering (2004) 779–779
17. Ceri, S., Fraternali, P., Bongio, A.: Web modeling language (webml): a modeling language for designing web sites. Computer Networks **33**(1) (2000) 137–157
18. Object Management Group: Business Process Model and Notation Standard (formal/2011-01-03). `http://www.omg.org/spec/BPMN/2.0/` (January 2011) [Accessed: 05/07/2013].
19. Object Management Group: Interaction Flow Modeling Language. `http://www.ifml.org/` (March 2013) [Accessed: 05/07/2013].
20. Escalona, M., Koch, N.: Requirements engineering for web applications-a comparative study. Journal of Web Engineering **2** (2004) 193–212

Direct system integration of hardware models – a case study

Johannes Hiltscher and Wolfgang Rehm

Chemnitz University of Technology
Department of Computer Science
Computer Architecture Group
{jhil,rehm}@informatik.tu-chemnitz.de

Abstract. As the complexity of newly developed hardware grows, evaluating its correct operation becomes more and more challenging. Classic development and evaluation paradigms have been replaced by advanced techniques like Hardware/Software Co-Design, and simulation has become a vital part in the hardware development process. This creates new challenges like the generation of meaningful input signals to the device under test and the early provision of prototypes. In this paper a simulation and evaluation environment is presented which integrates hardware simulation models into the Linux kernel's device sub system. It will be shown that the proposed environment is capable of satisfying the two aforementioned requirements. This novel approach allows for prototyping extension hardware for general purpose computers with a low overhead and no need for additional hard- and software.

1 Introduction

Testing the correct operation of a newly developed IC (Integrated Circuit) is a necessary while demanding process. Hardware simulation is commonly employed to reduce both costs and time required for discovering and correcting errors. Simple hardware with only few internal states can be tested using static input patters, but with growing device complexity this approach quickly becomes infeasible. Only a subset of the device's internal states can be assessed with an affordable effort and a device working correctly in simulation might still fail in every day operation. More complex test patterns can be generated with low effort, either using formal methods if applicable or real-world stimuli encountered during system operation. Using appropriate techniques, a larger subset of states can be covered during tests, increasing the chance to discover subtle errors.

In this paper, a novel simulation and evaluation environment is proposed which integrates simulated hardware models into the Linux kernel directly. Using this approach, hardware designers can provide evaluation versions of a newly developed hardware early in the design process. The presented approach integrates the hardware model into the operating system environment almost transparently. Therefore software designers can start working with the developed device very early in the design process, providing true Hardware/Software Co-Design. The

proposed environment allows for running real applications on the simulated hardware and can thereby conveniently generate complex test patterns. In contrast to other environments, the presented one requires no extra hard- or software and can reduce the simulation overhead.

The remainder of this paper is structured as follows. First an overview of work in the field of hardware simulation environments is given in Section 2 which are contrasted to our novel approach. In Section 3, the core concepts of the proposed simulation environment are presented. Experiences gained from using the environment for prototyping an extension to an existing NIC (Network Interface Controller) are presented in Section 4. Section 5 summarises and reflects those experiences and presents performance figures encountered in the case study. This paper is concluded by a reflection of the performed work and an outlook to possible future extensions and optimisations in Section 6.

2 Related Work and Proposal of a Novel Simulation Framework

Besides formal or statistical methods for test signal generation, engineers can also use real-world data for evaluation. Three simulation approaches, differing in test pattern generation and input to the evaluated system, are introduced in this section. They are analysed in the following concerning the degree as to which the target system is modelled, the method of generating input signals and distributing them to the design under test as well as the typical structure of an evaluation environment. This is meant to give an overview only, therefore details like hardware modelling approach or simulator used are omitted. Afterwards a conclusion is given to point out the motivation for our novel simulation approach.

2.1 Trace-Driven Simulation

Trace-driven simulation [6] has long been used for the evaluation of hardware systems. In the first step signals are recorded from some existing system which are then used as inputs to the hardware under test in a second step. Multiple approaches for trace recording exist [13], differing in recording granularity, from which designers choose the most appropriate for their hardware. Simulation accuracy can be trade for lower run time and trace file size this way.

For trace-driven simulation, the hardware model can be constrained to the design under test. Literature enquiries reveal that this approach has been applied for a wide range of hardware components. To perform trace-driven simulation, a recording tool is required which may be either hard- or software, a good overview is given in [13]. In this paper only software based approaches similar to those described in Section 2.2 are considered which can be used to record traces at transaction level. Transaction level here means that data exchange is modelled using abstract units like bus packets. Such traces can easily be transformed to logic level traces where transactions are represented by sets of signals as defined by the respective bus protocol. Hardware trace recording is usually very costly, especially for modern interconnects like PCIe, QPI or HT (HyperTransport).

2.2 Partial System Model

Partial system models simulate the device under test and some auxiliary components to interface with it, keeping simulation overhead low. System software and hardware simulation are run concurrently, providing a true Hardware/Software Co-Design environment. Instead of using recorded signal traces, interaction of software with the evaluated device is propagated on the fly. This is commonly achieved by running the target OS (Operating System) in a VM (Virtual Machine) in parallel with the simulation. The simulated hardware is mapped to the VM's I/O (Input/Output) address space and the simulator is notified whenever it is accessed. A bus model is employed to convert accesses to the evaluated hardware to the transaction level and vice versa. For the example of QEMU [2], this is achieved using a software MMU (Memory Management Unit) as detailed in [14].

This approach allows designers to concentrate on the design under test without requiring a simulatable model of the target CPU (Central Processing Unit). Several simulators employing a partial system model have been proposed in recent years [10],[8]. On the downside, this approach can not be used for Hardware/Software partitioning, i.e. the allocation of functionality to either a software application or a dedicated hardware circuit. VMs are incapable of providing a comparable timing measure as they perform emulation for performance reasons.

2.3 Full System Model

Using a full system model for evaluation is the most elaborate approach of hardware simulation. A simulation model of the whole hardware environment is required, including complex components like the CPU. While this approach provides most insight into the overall performance and operation of the whole system, it is also the most demanding one. Several environments, like gem5 [4], MPARM [3], MARSS [11] or the commercial Simics [9], exist. They are complex and powerful tools able to accurately model and analyse a complete system and run an OS. As high simulation accuracy implies low performance, designers can usually trade off between them [4]. Most simulators support different degrees as to which they resemble original system behaviour. While being the most complex approach, it can ideally be used for Hardware/Software partitioning [3]. This is possible as CPU and hardware simulation can easily be synchronised and a realistic, timing-accurate model of the different components is available. Resource requirements and performance are hence directly comparable.

2.4 Synopsis

The previously discussed approaches are well suited for many use cases, with specific strengths and weaknesses depending on the needs of the designer. Depending on the specific case they can be overly complex, not applicable or infeasible for small groups due to imposed costs. Trace-driven approaches commonly use software recording measures to avoid high costs. Similar approaches as introduced

for partial system models can be used, making trace recording versus instantaneous trace forwarding the only difference. Partial system models have been shown to be well suited when designing a SoC (System-on-Chip) (c.f. [10]) which often integrates a different CPU than the machine used for hardware simulation. The main benefit of this approach is the completely transparent integration of simulated hardware at the cost resources for each VM. Due to a soft MMU being used, all memory and I/O accesses are slowed down. In [10] this slowdown is quantified by a factor of 7 for the i386 architecture but must be considered highly application dependent. In many cases a full system model will be unnecessarily complex for a given task. When designing extension hardware, usually there is no need to consider aspects like CPU caches. This renders the imposed overhead useless, which can be quite high depending on the accuracy of the model (e.g. [11] gives an average of 200k instructions per second for MARSS). In most cases it is hence preferable to simulate no more components than necessary.

Our novel approach aims at limiting the overhead of hardware simulation. It is capable of integrating simulated hardware into the Linux OS kernel without an abstraction like a VM. When developing extension hardware, especially for multi-node systems, this approach eliminates the overhead of the VMs. By removing the need to pass all memory and I/O operations through a soft MMU, only those targeting the simulated hardware are emulated. This imposes slowdowns only when necessary while other operations run native. The presented approach is primarily analysed from a work flow perspective in a case study and some preliminary performance figures are presented.

3 Core Components and Concepts of the Proposed Environment

Figure 1 gives an overview of the proposed simulation environment's components, annotated with their execution domain. A number of Linux kernel modules form its core which handle the communication with and the management of the simulated hardware, referred to as virtual devices or hardware. The concept of kernel module stacking is used extensively, low level modules provide basic functionality which others utilise to implement specific services. The lowest level module is the SVB (Simulation Virtual Bus) which provides protocol-agnostic communication facilities used by higher level protocol modules to model an interconnect at transaction level. Such interconnect models are referred to as protocol model kernel modules, they define transaction data structures and translate read and write accesses to the appropriate transactions. Device drivers communicate with the virtual hardware using functionality provided by a protocol model kernel module. The device drivers are in charge of handling the communication and data exchange with user space applications.

The SVB implementation strives to impose as few limitations as possible on designers. The intent is to provide simple, though flexible facilities for modelling different data exchange paradigms. The SVB implements the communication with the hardware simulator using Linux device files, hence any simulator provid-

ing file I/O can be employed. While SystemC is used in the case study, up-to-date Verilog or VHDL simulators can also interface with the SVB and hardware designers can develop using their preferred HDL (Hardware Description Language). Independent up- and downstream channels are maintained for each virtual device which can concurrently be used by multiple applications. No communication protocol is defined at this level as the SVB is implemented to be protocol agnostic. It can therefore be used as a foundation for modelling a wide range of interconnects by implementing different protocol kernel modules. Besides communication facilities, the SVB provides life cycle management for dynamic device (dis-)connection. A Linux bus structure is integrated into the kernel's device sub system supporting driver registration and automatic probing to support dynamic insertion and removal of virtual devices. Device drivers can therefore be developed at an early stage in the design process, only minimal changes are required when moving to physical hardware. To simplify the development of protocol kernel modules, the SVB supplies a number of convenient functions for blocking and non-blocking data exchange with virtual hardware. Dedicated helper files containing macros are supplied which aid driver writers by minimising the need for changes in driver code when moving from virtual to real hardware. When compiling a driver for virtual hardware, they evaluate to data exchange functions of the protocol model kernel module and to direct accesses to the memory mapped device otherwise.

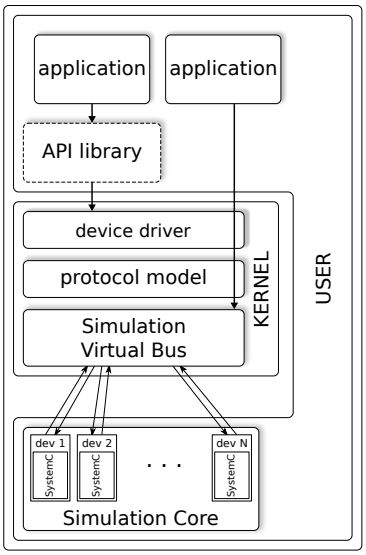

Fig. 1. Layer model of the proposed simulation environment

The hardware model is shown at the bottom of Figure 1. Multiple instances can be encapsulated in an application referred to as the Simulation Core which provides synchronisation between several virtual devices. Besides the hardware model, the Simulation Core also needs to contain functionality to probe the SVB for new transactions and translate these to inputs for the virtual hardware. The reverse translation has to be done for transactions originating from the virtual hardware which can then be passed to the protocol model kernel module.

Although recommended, using an intermediary protocol model kernel module is not strictly necessary. Applications can also directly communicate with the virtual hardware via the SVB but need utilise the appropriate transaction data structures. This allows for quickly testing communication protocols or basic features without the need to implement a protocol model kernel module. Although the protocol model kernel modules provide a good abstraction of the fact that virtual hardware is used, complete transparency is not achieved. Mem-

ory mapping of and transparent access to device memory can currently not be supported by the simulation environment. This has to be addressed during the design process of device drivers and user space applications, depicted topmost in Figure 1. To hide the different handling of virtual hardware from applications, an API (Application Programming Interface) library should be used.

4 Case Study: Extension of a NIC

The proposed simulation environment has been used to develop and evaluate an extension module to a NIC designed to provide high level local and remote DMA (Direct Memory Access) functionality to a connected hardware module. For this project, a protocol model kernel module for HT [7] was developed which extends the SVB functionality to simulate the behaviour an application would encounter for a real device. Two functions implement read and write operations based on the SVB data exchange functions. They automatically create transaction data structures resembling abstract HT packets for arbitrarily large amounts of data which are transferred to the virtual hardware via the SVB. If a response is expected, the caller is interrupted until its receipt. The HT protocol model kernel module employs a polling task which regularly checks the upstream SVB buffers for new messages using the non-blocking access functions (c.f. Section 3). Responses are forwarded to the process which triggered their generation and it may resume operation. The HT protocol model kernel module allows drivers to register callback functions which are used to handle requests from virtual devices, providing an option to make host memory accessible to virtual devices. The polling task calls the registered callback functions for each received transaction which is not a response and drops it in case none accepts it. The driver module of the network engine shown in Figure 2 makes its memory receive buffers available to the virtual hardware this way.

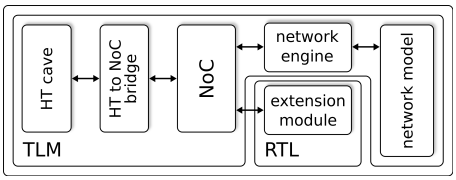

Fig. 2. Structural view of the simulated hardware

The original driver of the NIC consists of several different kernel modules. Modifications to them were necessary to make them work with the virtual hardware. These covered changing the allocation of memory buffers, adding callback functions for the protocol model kernel module and replacing memory mapped device accesses. Changes to the network engine driver were required as it maps parts of the device's I/O memory region to user space applications. This was replaced using device files which applications read from or write to for accessing the device. An API library is used by applications to interact with the NIC's network engine. Accesses to mapped I/O memory were replaced with file operations on the device files supplied by the driver using the standard Unix `pwrite` function to transfer an address and data. The library was also enhanced to read an environment variable to determine the virtual device to use on initialisation. However, these changes were minimal and applications

can be run without further modification on virtual and real hardware using the modified API library.

As the drivers already existed for the NIC hardware, the number of changes was not negligible. The impact on newly created drivers could be studied when implementing the extension module driver. It showed that, using the supplied helper files, in this case there was no re-engineering work required for the physical hardware. Therefore software developers can almost completely avoid re-engineering when making consistent use of the helper functions and macros. They automatically perform SVB-mediated device accesses in the simulation driver and memory mapped accesses in the physical hardware driver. Both drivers are hence functionally equivalent, although device accesses are performed differently.

The NIC as well as the developed extension were modelled using SystemC [5]. It was chosen for the hardware description as it provides higher level modelling capabilities than classic HDLs. Besides the availability of the C and C++ libraries, using TLM (Transaction-Level Modeling) [1] it is possible to implement very fast simulation models. This allows hardware designers to speed up their simulation when a precise logic level model is not required. Furthermore, abstract TLM based hardware models can be made available very early in the design process, allowing true Hardware/Software Co-Design. These can then be refined iteratively from a coarse-grained data flow model to a synthesisable RTL (Register Transfer Level) description. All TLM components were implemented according to the TLM 2.0 specification. For the case study RTL and TLM models were mixed, the NIC hardware and network model were implemented using TLM, the extension module is a synthesisable RTL description. This approach assures a high simulation performance while allowing fine-grained simulation where necessary. Figure 2 gives an overview of the simulated device's structure, annotated with the modelling paradigms used for the different components. The existing hardware is modelled according to the TLM approximately timed model (c.f. [1]) where transactions are asynchronously passed through the design with an annotated latency. They are exchanged between different units using defined interface functions and each traversed unit only alters latencies to reflect its timing behaviour. Simulator events are only generated when a unit explicitly needs to synchronise a transaction to the simulation time. This can significantly reduce the total amount of simulator events compared to a less abstract, signal based model where each signal at each connection point of two units can generate an event. While providing accurate results concerning transaction timing, both simulation run time and model complexity can be reduced.

In contrast, the developed extension module is modelled as a detailed RTL hardware description as the goal was to synthesise it for an FPGA (Field Programmable Gate Array). The design is quite complex, consisting of six different units and a NoC (Network-on-Chip) for their interconnection. TLM transactions are converted to logic signals by a dedicated module, a conceptual view of which is shown in Figure 3. The bridging from the approximately timed TLM to the RTL model is a crucial point in the simulation model. Specific care had to be taken here to correctly model the pipelining behaviour of the hardware. Hence

a distinction was made between *static latency* which is the forwarding latency of the traversed units and *dynamic latency* which depends on the amount of data transferred and the width of interconnects. Dynamic latency is only considered when a transaction reaches the end of a communication chain, e.g. a buffer. The TLM to RTL bridge implicitly adds the dynamic latency by serialising each transaction to a number of interconnect words.

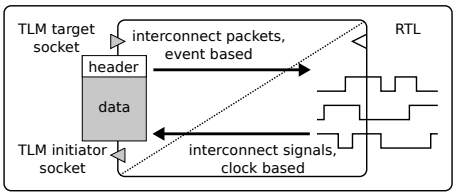

Fig. 3. Conceptual view of the TLM to RTL bridge

The network interconnecting the individual NIC instances, shown rightmost in Figure 2, is also modelled using the approximately timed TLM approach. Together with all simulated NIC instances it is run in one Simulation Core.

5 Evaluation

The conducted case study revealed that the presented simulation environment is capable of generating meaningful signal-level inputs to a device under test which are generated from real accesses. Furthermore, it allows hardware designers to model the timing behaviour of system components to approximate performance figures at early design stages. Software developers can be supplied with a development platform early in the design process. Applications and drivers can hence be implemented in a very early phase of the project. Testing and performance estimations can be done without the physical hardware available, so hardware vendors could supply interested customers with virtual evaluation versions of their products. Characteristics like latencies, flow control or buffering capabilities can be modelled to closely resemble real hardware behaviour, thus permitting comparison with other solutions.

The proposed environment mainly competes with the VM-based approaches presented in Section 2.2. In contrast to such frameworks no extra software is required as it could be shown that direct integration of simulated hardware into the Linux OS is feasible. Current shortcomings of the presented approach were discussed and methods for coping with them proposed. Further considerations on this point are presented in the following Section 6 as future work.

The environment was first used for developing a complex extension to a NIC which occupies about half of the slices of a Xilinx Virtex-IV VFX100 FPGA. The performance figures given in the following are therefore only hints guiding the way for further optimisation. All figures were collected on an AMD Phenom II X4 940 with 8 GiB of main memory running Ubuntu 10.04, with two simulated NICs. Simulation Core and kernel modules were compiled using GCC 4.4.3 with optimisation `-O3` and the SystemC 2.3 library. The simulation is run in steps of 1000 ns for which the duration is calculated using the processor's TSC (Timestamp Counter). Due to restrictions of the simulator, simulation of SystemC models always runs on only one core, trace files were not generated.

Running only the TLM model of the NIC yields an operating frequency for the simulated hardware of about 1.4 MHz. This is a slowdown by a factor of 111 compared to the physical hardware's 156 MHz frequency when implemented on a Xilinx Virtex-IV VFX100 FPGA. Simulating the hardware model described in Section 4, frequency dropped to about 14.3 kHz, which is a slowdown of about 10100 compared to the original device frequency. Comparing these results to other findings is hard as other papers usually don't give performance figures. For a rough comparison, figures found in [14] can be used as simulation run time and the number of executed instructions are given. The virtual platform is based on the ARM926 core which yields about 1 million instructions per second. Assuming the same 156 MHz as for the case study device, a slowdown of about 185 can be calculated. Although not explicitly stated, it can be assumed that a TLM hardware model was employed concerning the data throughput rates of TLM and cycle accurate AMBA (Advanced Microcontroller Bus Architecture) interfaces presented in [12]. It can therefore be assumed that the presented simulation environment achieves similar or better performance than VM-based solutions.

The presented case study revealed that quite some work was required to adapt the existing NIC driver for the simulation environment. Many changes were required to deal with shortcomings of the original driver. Much of the conducted work can however be considered non-recurring as it covered fundamental aspects like the protocol model kernel module and the SVB which can be reused in other projects. As the case study also revealed, a newly written driver could be created without any need for subsequent changes for physical hardware.

6 Outlook and Future Work

The presented simulation environment provides a robust foundation for evaluation and development of a wide range of devices. Performance measures, although promising, are likely to improve when re-considering some design choices. Mapping the SVB queues, along with the corresponding read and write pointers, to the running simulation to reduce the amount of system calls was considered. The existing code could also benefit from further review, possibly reducing the number of simulator events. However, this only affects the performance of the TLM parts of the simulation environment. As shown in Section 5, complex RTL hardware models degrade performance significantly. Fine grained simulation at logic level however causes high overhead, rendering simulation environment overhead almost negligible. As the presented approach for integrating simulated hardware mainly competes with VM-based frameworks, a comparison between these two could be conducted. This comparison should consider aspects like development overhead and analysing the amount of non-recurring vs. recurring work for both. Furthermore performance comparisons should be conducted for different applications to study the impact of virtualisation using a soft MMU on overall application performance.

A major drawback of the current implementation is the inability to transparently map memory regions of simulated devices to an application. Using a

custom page fault handler and mapping inaccessible dummy pages which generate a page fault when accessed was considered for its emulation. The handler would evaluate the faulting instruction, generate an appropriate operation and return after its completion. As the feasibility of this approach was not yet analysed, it has been scheduled for future work.

References

1. Aynsley, J.: OSCI TLM-2.0 language reference manual. Open SystemC Initiative (2009)
2. Bellard, F.: QEMU, a fast and portable emulator. In: Proceedings of USENIX Annual Technical Conference. pp. 41–46 (Jun 2005)
3. Benini, L., Bertozzi, D., Bogliolo, A., Menichelli, F., Olivieri, M.: MPARM: Exploring the multi-processor SoC design space with SystemC. J. VLSI Signal Process. Syst. 41(2), 169–182 (Sep 2005)
4. Binkert, N., Beckmann, B., Black, G., Reinhardt, S.K., Saidi, A., Basu, A., Hestness, J., Hower, D.R., Krishna, T., Sardashti, S., Sen, R., Sewell, K., Shoaib, M., Vaish, N., Hill, M.D., Wood, D.A.: The gem5 simulator. SIGARCH Comput. Archit. News 39(2), 1–7 (Aug 2011)
5. Board, I.S.S.: IEEE Standard for Standard SystemC© Language Reference Manual. The Institute of Electrical and Electronics Engineers, Inc., 3 Park Avenue, New York, NY 10016-5997, USA (2012)
6. Cheng, P.S.: Trace-driven system modeling. IBM Systems Journal 8(4), 280 –289 (1969)
7. HyperTransport Technology Consortium: HyperTransport^{TM}I/O Link Specification (Aug 2003), revision 1.10, available at www.hypertransport.org
8. Lin, J.W., Wang, C.C., Chang, C.Y., Chen, C.H., Lee, K.J., Chu, Y.H., Yeh, J.C., Hsiao, Y.C.: Full system simulation and verification framework. In: Information Assurance and Security, 2009. IAS '09. Fifth International Conference on. vol. 1, pp. 165–168 (Aug 2009)
9. Magnusson, P.S., Christensson, M., Eskilson, J., Forsgren, D., Hallberg, G., Hogberg, J., Larsson, F., Moestedt, A., Werner, B.: Simics: A full system simulation platform. Computer 35(2), 50–58 (Feb 2002)
10. Monton, M., Portero, A., Moreno, M., Martinez, B., Carrabina, J.: Mixed SW/SystemC SoC emulation framework. In: Industrial Electronics, 2007. ISIE 2007. IEEE International Symposium on. pp. 2338–2341 (June 2007)
11. Patel, A., Afram, F., Chen, S., Ghose, K.: MARSS: a full system simulator for multicore x86 CPUs. In: Proceedings of the 48th Design Automation Conference. pp. 1050–1055. DAC '11, ACM, New York, NY, USA (2011)
12. Schirner, G., Dömer, R.: Quantitative analysis of transaction level models for the amba bus. In: Proceedings of the conference on Design, automation and test in Europe: Proceedings. pp. 230–235. DATE '06, European Design and Automation Association, 3001 Leuven, Belgium, Belgium (2006)
13. Uhlig, R.A., Mudge, T.N.: Trace-driven memory simulation: a survey. ACM Comput. Surv. 29(2), 128–170 (Jun 1997)
14. Yeh, T.C., Chiang, M.C.: On the interface between QEMU and SystemC for hardware modeling. In: Next-Generation Electronics (ISNE), 2010 International Symposium on. pp. 73–76 (Nov 2010)

Automatic Application of Visitors to Evolving Domain-Specific Languages

Adolfo Sánchez-Barbudo Herrera[1], Ed Willink[2], Richard F. Paige[1], Louis M. Rose[1] and Dimitrios S. Kolovos[1]

[1] Department of Computer Science, University of York, UK.
{asbh500, richard.paige, louis.rose, dimitris.kolovos}@york.ac.uk
[2] Willink Transformations Ltd. ed@willink.me.uk

Abstract. Domain-specific languages (DSLs) invariably need to evolve, particularly when they are developed in an iterative and incremental fashion. This paper argues that the evolutionary process can be better controlled and managed by separation of DSL descriptions into those of structure and behaviour, and thereafter can be effectively supported by tools that automatically allow *visitors* to be applied to DSLs. A novel prototype tool is described that supports visitor applications to DSLs developed with EMF[3]. The advantages and disadvantages of such an approach to DSL evolution are highlighted.

1 Introduction

Domain-specific languages (DSLs) evolve, whether as the result of requirements for new language constructs or applications of the languages, or as a result of the engineering process that is to be followed. When DSLs evolve, the models that are expressed in these languages must be migrated to conform to a new version of the DSL. There are numerous approaches available to support migration, DSL evolution, and co-evolution problems [1, 2]. An open question is: how can DSL designers best structure their language descriptions so as to make the DSL evolution (and hence model migration) process more manageable? Are there language design patterns that can be applied that make DSL evolution, or model migration, or model-metamodel co-evolution, cheaper or more automatable?

This paper contributes to the discussion on this question by exploring the use of the *visitor* design pattern [3] in the construction of underlying DSL metamodels. Broadly, we are interested in investigating whether the direct application of visitors – which decouple structure from behaviour – can have benefits in terms of DSL design and evolution. As part of this, in this paper we present a novel automated approach to applying visitors directly to arbitrary metamodels. We implement the approach using automated model management languages (including QVT [4]) and assess the advantages and disadvantages of the approach, both in terms of application of patterns and DSL design and evolution.

[3] Eclipse Modeling Framework. http://www.eclipse.org/modeling/emf/

2 Background and Related Work

2.1 Visitors

The visitor pattern [3] is used to represent operations (behaviour) that are to be applied to components (classes, objects) of an object-oriented class structure. An application of the pattern is separated into a *visitor* (which represents the operation behaviour) and a *visitable* (the structures to which behaviour is applied). In particular, the visitor pattern decouples structures from behaviours that apply to these structures; it is useful in situations where the behaviours may be changing frequently. When the structure changes frequently, the visitor makes change more difficult to manage: changes to class structures require changes to all visitors. Typical uses for visitors [3] include: encapsulating structure (graph, tree) traversal algorithms; encoding pretty-printers or visualisation/rendering algorithms; expressing analytics to be applied to structures; and defining validators, simulators or execution engines applied to structures. Many authors have used visitors [5, 6], studied their implementations [7, 8], and considered the flexibility-performance tradeoff [9, 10]. However, to the best of our knowledge most visitors related work are focused on manual implementations of the essence of the design pattern.

In the context of Model-Driven Engineering, we can mention some work benefiting from visitors like the *MontiCore* framework [11], in which they apply the design pattern for the same objectives presented here. Nevertheless, they are manual implementations relying on Java reflection, which is a flexible but a slow solution [9, 10]. Instead, our approach leans towards an automated solution which generate visitors relying on the faster original design pattern, at the cost of having to prepare the DSL to work with them.

2.2 Model Evolution

Many approaches have been proposed for addressing model evolution and model-metamodel co-evolution. These include manual approaches (e.g., editing XMI directly), operator-based (e.g., COPE/Edapt [12]), and inference-based approaches (such as [1, 2]). Manual and inference approaches are dependent: in manual approaches, migration strategies are specified by hand in response to metamodel evolution, while in the latter the original and evolved metamodels are compared, and their differences used to generate a migration strategy. Operator-based approaches are interdependent, and metamodel developers use operators to apply co-evolution patterns. In general, fully automatic co-evolution of models and metamodels is impossible, due to the existence of breaking and unresolvable changes [13]. These are changes for which the correct migration action can only be determined by consulting the metamodel user, and include changes that require domain expertise to resolve. We focus on model behaviour, which at the best of our knowledge hasn't been widely addressed when considering DSL evolution. Particularly, we propose the usage of visitors in an automated way to address the evolution of the behaviour for the DSL models.

3 DSL development

DSLs are developed in a number of ways: by tailoring existing general-purpose languages, by reusing DSLs, or from first principles, e.g., by identifying a set of concepts, constructing an abstract syntax, and then building language processing tools (e.g., simulators, editors). In the last case, the evolution of the DSL can follow three steps, as depicted by Figure 1. In the first step, the set of concepts in the DSL is populated (e.g., by domain analysis) and initial structures are determined. In the second step, the structures and concepts in the DSL are restructured and refactored (for example, to enhance expressiveness or remove redundancy). In the third step, behaviour is attached to the DSL, e.g., via model transformation.

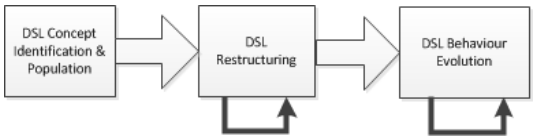

Fig. 1. A DSL design process

Existing approaches to model migration and metamodel evolution are particularly well suited to evolutionary tasks in the first two steps (e.g., adding new meta-concepts, moving meta-concepts), but are not yet fully robust or sufficiently expressive to even semi-automatically support the evolution of behaviours of DSLs. For example, suppose one type of behaviour attached to the DSL is in the form of a static type-checker. If the structure or concepts in the DSL change, then existing model evolution tools will not easily allow the type checker to be suitably evolved – significant manual rework will be required, and perhaps even new model traversal algorithms constructed.

Though existing model migration/evolution tools are not currently well suited to DSL behaviour evolution, DSLs could potentially exploit the visitor pattern to make it easier to manage the evolutionary process. By applying visitors to DSLs and effectively decoupling the structure of the DSL from their behaviour, it should make it easier to provide automated support to evolve behaviour. In particular, consider the following two use cases.

- A DSL has 100 concepts in its metamodel. If new behaviour needs to be added to the DSL (e.g., a simulator), then it is possible that each of the 100 meta-concepts will need to be modified to take into account the behaviour of the new language construct. However, if a visitor pattern has been applied, then a new operation can be encapsulated as a visitor, and the existing metamodel can be used unchanged.
- Consider the same DSL, but now suppose we have five visitors (operations) currently applied to the DSL, and that a new meta-concept is to be added.

This may invalidate the existing five visitors, but if we have automated tool support for visitor application, then we may be able to automatically update the visitors (modulo detailed behaviour that differs from any default behaviour), thus providing a potentially significant saving.

The use of the visitor pattern for DSL design has the potential to provide benefits to language evolution, both in terms of metamodel changes and in support of new behaviours associated with languages. However, if we want to apply and rely on the use of visitors in DSL design, it is beneficial to have automated tools that allow us to apply visitors to arbitrary DSLs.

4 Automated Visitor Application

In this section we present an MDE-based solution to automatically apply the visitor pattern to arbitrary EMF[14] based languages. The automated solution we present requires minimal manual configuration effort, allowing the user to focus on defining the fine-grained behaviour of custom visitors.

The MDE tooling is based on model transformations. The model transformations involved in the application are:

1. A model-to-model (M2M) in-place transformation, responsible for mainly making each concrete (non-abstract, instantiable) class *Visitable* via inheritance, and creating its corresponding *accept* operation.
2. A model-to-text (M2T) transformation so that the *Visitor* and *Visitable* Java interfaces are automatically generated, as well as default *Visitor* implementations (which may be overridden).

A MDE workflow is responsible for chaining the M2M transformation, the EMF Generator and finally the M2T transformation; in our prototype, we implemented this using MWE2[4] (but do not discuss this here due to space limitations). We explain the M2M and M2T steps in the workflow in following subsections.

4.1 M2M in-place transformation

The first activity is the execution of a M2M in-place transformation, to make every class *Visitable*, and weave the corresponding *accept* operations. To accomplish this, we have used QVT Operational Mappings language[4] supported by the Eclipse QVTo project technology[5].

The first step in the QVTo transformation involves querying the EMF based model of our target language for *Visitor* and *Visitable* classes, in case these concepts already exist. If they are not found, corresponding interfaces are added to the target package.

Once we have the corresponding *Visitor* and *Visitable* interfaces, we can weave the latter into the package classes; this requires executing the mapping

[4] `http://www.eclipse.org/Xtext/documentation.html#MWE2`
[5] `http://projects.eclipse.org/projects/modeling.mmt.qvt-oml`

```
1  mapping inout EClass::weaveVisitableInterface(in visitableClass: EClass)
2    when { self.eSuperTypes->isEmpty() } {
3    self.eSuperTypes += visitableClass;
4  }
```

Listing 1.1. QVTo mapping to weave the *Visitable* interface into the language.

```
1  mapping inout EClass::weaveAcceptMethod(in visitorClass: EClass)
2    when { self.eOperations->select(name = 'accept')->isEmpty()
3       and not self._abstract
4       and not self.interface; } {
5    var acceptOp := createAcceptMethod(visitorClass);
6    acceptOp.eAnnotations += createGenModelAnnotation("return v.visit" + self.
         name + "(this);");
7    eOperations += acceptOp;
8  }
```

Listing 1.2. QVTo mapping to weave the *accept* methods.

of Listing 1.1. By means of a mapping's `when` clause, we ensure to execute the mapping for each class that doesn't have a supertype. Once they have been made *Visitable*, the remaining classes will automatically be visitable as well, by inheritance.

Listing 1.2 depicts the last step which adds the `accept` operation into every instantiable class. The mapping's `when` clause ensures that only instantiable classes are considered when we create an `accept` operation, which also includes an Ecore[6] annotation to guide the EMF generator to produce the corresponding `accept` method implementation.

4.2 M2T templates

Another activity in the MDE workflow involves the execution of M2T transformations, using the Acceleo technology[7], in order to:

1. Create metamodel-specific *Visitor* and *Visitable* external interfaces, so that, e.g., our *Visitor* interface will have one *visit**ConcreteElement*** for every concrete (instantiable) concept of the language.
2. Create metamodel-specific *abstract* visitor implementations, providing a variety of default implementations, e.g., abstract contextful visitor, abstract visitor that returns *null*, a delegator, etc.

Listing 1.3 provides a template excerpt of a specific concrete visitor depicting how the visit methods are implemented for a visitor counting the number of instances of a model element. Each visitor simply increments the count of the corresponding concept type. This particular visitor implementation comprises not only the increment action, but also the traversal strategy.

[6] Ecore is the name that receives the language used to define any EMF based modeling language.

[7] `http://www.eclipse.org/acceleo`

```
1  [for (eClass : EClass | instantiableEClasses)]
2   public @Nullable [countClassName/] visit[eClass.name/](@NonNull [eClass.name
        /] object) {
3    context.inc[eClass.name/]Count();
4    [for (eStructuralFeature : EReference | eClass.eAllReferences->select(
        containment))]
5     [if (eStructuralFeature.upperBound > 1)]
6    for ([eStructuralFeature.eType.name/] [eStructuralFeature.eType.name.
        toLowerFirst()/]Obj : object.get[eStructuralFeature.name.toUpperFirst
        ()/]()) {
7     [eStructuralFeature.eType.name.toLowerFirst()/]Obj.accept(this);
8    }
9     [else]
10   if (object.get[eStructuralFeature.name.toUpperFirst()/]() != null) {
11    object.get[eStructuralFeature.name.toUpperFirst()/]().accept(this);
12   }
13    [/if]
14   [/for]
15   return context;
16  }
17 [/for]
```

Listing 1.3. Acceleo template for visit methods of an *ElementsCounter*.

5 Discussion

We previously explained the motivation for introducing the Visitor pattern in our DSLs. On one hand, separating structure from behaviour and allowing easy mechanisms to add, remove, or modify the behaviour of models has the potential to make it easier to modify the behaviour of a DSL. The main shortcoming attributed to this pattern also applies to DSLs: the pattern is effective as long as the structure remains static, otherwise existing visitors might not comply with the intended behaviour because they need to react to structural changes.

The proposed approach not only helps us apply the Visitor pattern to DSLs that need to evolve, but is also useful in the DSLs evolutionary process by leveraging the shortcomings of the own visitor pattern. We will discuss this with an illustrative example.

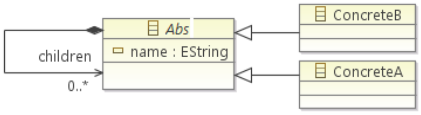

Fig. 2. A simple DSL example

Consider the Ecore diagram in Figure 2. Suppose we want a visitor to count the number of model elements from any model conforming to our DSL. The M2T template we showed before produces an *ExampleElementsCounter* visitor (Listing 1.4). Suppose that we need to evolve our DSL by introducing a new

```
1  public class ExampleElementsCounter
2    extends AbstractVisitor<ExampleElementsCount,Void> implements Visitor<Void>{
3    public ExampleElementsCounter(ExampleElementsCount context) {
4      super(context);
5    }
6    public @Nullable Void visitConcreteA(@NonNull ConcreteA object) {
7      context.incConcreteACount();
8      for (Abs absObj : object.getChildren()) {
9        absObj.accept(this);
10     }
11     return null;
12   }
13   public @Nullable Void visitConcreteB(@NonNull ConcreteB object) {
14     context.incConcreteBCount();
15     for (Abs absObj : object.getChildren()) {
16       absObj.accept(this);
17     }
18     return null;
19   }
20 }
```

Listing 1.4. Generated *ExampleElementsCounter* visitor.

```
1  public class ExampleElementsCounter
2    ...
3    public @Nullable Void visitConcreteB(@NonNull ConcreteB object) {
4      context.incConcreteBCount();
5      for (Abs absObj : object.getChildren()) {
6        absObj.accept(this);
7      }
8      for (ConcreteC concreteCObj : object.getNewChildren()) {
9        concreteCObj.accept(this);
10     }
11     return null;
12   }
13   public @Nullable Void visitConcreteC(@NonNull ConcreteC object) {
14     context.incConcreteCCount();
15     return null;
16   }
17 }
```

Listing 1.5. Regenerated and consistent *ExampleElementsCounter* visitor.

ConcreteC which may be contained by an *Abs*, via a *newChildren* reference. Unfortunately, this DSL evolution would turn our current *ExampleElementsCounter* visitor into a deficient one: *ConcreteC* elements wouldn't be counted.

However, our MDE approach can help to make the *ExampleElementsCounter* visitor consistent. If we execute the MWE workflow again, we address this issue by obtaining new consistent visitors (Listing 1.5). Arguably, this is a contrived example that is well suited for the solution we have produced: the behaviour of a model element counter doesn't require more information beyond that contained in the metamodel. More complex behaviours would not so easily be handled as a result of changing the DSL structure.

6 Conclusions and Future Work

We have discussed DSL evolution and suggested that a separation of concerns, supported by the visitor pattern, may help in addressing certain DSL evolution scenarios. We have described a novel prototype for automated application of visitors to EMF[14] based DSLs. The prototype contributes:

- A minimum of automation to build a DSL-specific visitor framework to address DSL evolution in terms of behaviour.
- The foundation for a Visitor Generation Framework (VGF) capable of generating more complex visitors by the means of additional DSLs describing the desired behaviour.

Our approach is targeted at evolution (addition, removal, modification) of DSL behaviours. The approach is not immune to traditional model evolution issues: whilst some DSL-independent behaviours might be fully automated via custom efficient generated visitors when structural changes occur, there are scenarios in which existing behaviours may be impacted. Depending on the nature of the metamodel change (e.g. removing a concept) and the nature of the concrete intended behaviour, the proposed approach might not be able to correctly regenerate functional visitors. In future work, we need to explore these scenarios and possible solutions.

Other considerations we are taking into account is providing support for a family of DSLs rather than a single one: the DSLs must be a self-contained family of languages (i.e. with no relationships to external DSLs), because the visitor pattern needs to be applied to every DSL in the family. Depending on relationships between concepts of different DSLs (inheritance, aggregation, etc) the framework may be able to generate adequate visitors exploiting the features and design patterns for a target programming language.

Separating structure from behaviour in modelling languages is not a new idea [15]; DSLs may be complemented with additional artefacts (e.g., OCL [16]), which are responsible for integrating behaviour with our models. As future work we also want to study and compare pros and cons of these approaches, although we envision that generated metamodel-specific visitors may have performance advantages over a general-purpose engine. It may be that visitors can be generated from OCL constraints, or by means of more specific behavioural DSLs.

One of the shortcomings of applying the visitor pattern is that visitors can break when adding/removing/modifying concepts to our DSLs; our MDE-approach could also introduce conventions derived from the API Evolution[17] field when regenerating visitors after having an structural change. We plan to investigate how these techniques could apply to an entire MDE workflow.

Acknowledgement. We gratefully acknowledge the support of the EPSRC via the LSCITS initiative.

References

1. A. Cicchetti, D. Di Ruscio, R. Eramo, and A. Pierantonio. Automating co-evolution in model-driven engineering. In *Proc. EDOC*, pages 222–231. IEEE, 2008.
2. L.M. Rose, R.F. Paige, D.S. Kolovos, and F.A.C. Polack. An analysis of approaches to model migration. In *Proc. MoDSE-MCCM Workshop*, 2009.
3. Erich Gamma, Richard Helm, John Vlissides, and Ralph Johnson. *Design Patterns*. Addison-Wesley, 1994.
4. Object Management Group. Meta object facility (MOF) 2.0 query/view/transformation v1.1, January 2011.
5. Martin Nordberg. Variations on the visitor pattern. Quintessoft Engineering, 1996.
6. Bruno CdS Oliveira, Meng Wang, and Jeremy Gibbons. The visitor pattern as a reusable, generic, type-safe component. In *ACM Sigplan Notices*, volume 43, pages 439–456. ACM, 2008.
7. Judith Bishop and R Nigel Horspool. On the efficiency of design patterns implemented in C# 3.0. In *Proc. TOOLS 2008*. Springer.
8. Jan Hannemann and Gregor Kiczales. Design pattern implementation in java and aspectj. In *ACM Sigplan Notices*, volume 37, pages 161–173. ACM, 2002.
9. Jens Palsberg and C Barry Jay. The essence of the visitor pattern. In *Proc. COMPSAC'98*. IEEE.
10. Bertrand Meyer and Karine Arnout. Componentization: the visitor example. *Computer*, 39(7):23–30, 2006.
11. Holger Krahn, Bernhard Rumpe, and Steven Völkel. Monticore: a framework for compositional development of domain specific languages. *International journal on software tools for technology transfer*, 12(5):353–372, 2010.
12. M. Herrmannsdoerfer, S. Benz, and E. Juergens. COPE - automating coupled evolution of metamodels and models. In *Proc. ECOOP'09, LNCS 5663*. Springer.
13. B. Gruschko, D.S. Kolovos, and R.F. Paige. Towards synchronizing models with evolving metamodels. In *Proc. Workshop on Model-Driven Software Evolution*, 2007.
14. Dave Steinberg, Frank Budinsky, Marcelo Paternostro, and Ed Merks. *EMF: Eclipse Modeling Framework*. Addison-Wesley, 2008.
15. Lars Hamann, Oliver Hofrichter, and Martin Gogolla. On integrating structure and behavior modeling with ocl. In *MDE Languages and Systems*. 2012.
16. Object Management Group. Object constraint language v2.3.1, January 2012.
17. Martin Fowler. *Refactoring: improving the design of existing code*. Addison-Wesley Professional, 1999.

Part III

Extended Abstracts

A Media Access Control Protocol for Wireless Ad Hoc Networks with Misbehaviour Avoidance

Chaminda Alocious, Hannan Xiao and Bruce Christianson

School of Computer Science, University of Hertfordshire, Hatfield, Herts, UK
{c.alocious,h.xiao,b.christianson}@herts.ac.uk

Abstract. The most common wireless Medium Access Control (MAC) protocol is IEEE 802.11. Currently IEEE 802.11 standard protocol is not resilient for many identified MAC layer attacks, because the protocol is designed without intention for providing security and with the assumption that all the nodes in the wireless network adhere to the protocol. However, nodes may purposefully show misbehaviours at the MAC layer in order to obtain extra bandwidth conserve resources and degrade or disrupt the network performance. This research proposes a secure MAC protocol for MAC layer which has integrated with a novel misbehaviour detection and avoidance mechanism for Mobile Ad Hoc Networks (MANETs). The proposed secure MAC protocol the sender and receiver work collaboratively together to handshakes prior to deciding the back-off values. Common neighbours of the sender and receiver contributes effectively to misbehaviours detection and avoidance process at MAC layer. In addition the proposed solution introduces a new trust distribution model in the network by assuming none of the wireless nodes need to trust each other. The secure MAC protocol also assumes that misbehaving nodes have significant levels of intelligence to avoid the detection.

Keywords: MANETs, Medium Access Control, Misbehaviour Avoidance

1 Introduction

Computer network security is one of the most important elements in computer systems. Wireless networks have been widely used in banking, military, medical and in many other sectors [3]. Wireless network security is becoming increasingly important due to the dramatic enhancement of the wireless devices (e.g. PCs, tablets, mobile phones). There are two types of wireless networks, infrastructure based networks (WLANs) and wireless Mobile Ad-hoc Networks. Infrastructure based networks are controlled by a centralized base station which is the receiver of the network for all the connected nodes [4]. In contrast MANETs are self-organized, dynamically changing the topology without a centralized base station. Wireless nodes in the MANETs communicate by forwarding packets on behalf of each other by working as router.

MANET is an autonomous collection of mobile nodes that communicate over the bandwidth constrained wireless network environment [10] [8]. MANETs need to contain the basic security requirements such as availability, fairness, authorisation, data confidentiality and data integrity [6]. MAC layer nodes misbehaviour has been a

problematic scenario for MANETs and infrastructure based networks. Some selfish mobile stations do not follow the IEEE 802.11 protocol rules in sharing medium. IEEE 802.11 protocol assumes nodes in the wireless network fully cooperate to the protocol [10]. However, due to vast enhancement of the programmability of network devices, changing these MAC layer protocol parameters has become easier. Distributed Coordinates Function (DCF) uses the Binary Exponential Back-off (BEB) mechanism to assign back-off for wireless stations, but unfortunately due to vulnerability, this mechanism can be exploited easily. Rest of this paper organized as research background in the next section. Next sections organized as, the proposed secure MAC protocol design and the conclusion.

2 Research Background

The research motivates to provide solutions for following MAC layer selfish misbehaviours.

- **Back-off value Manipulation:** In 802.11 MAC protocol selfish nodes use smaller back-off values than they should and also use fixed back-off values instead of random values. Back-off value Manipulation also includes nodes doesn't double the congestion window size after a collision.
- **Adaptive Cheating / Adaptive Misbehaviour:** Some nodes are smart to adapt their misbehaviour strategy to prevent them from being caught by regular detection methods. Intelligent nodes are aware of the detection scheme and adapt to mislead the detection.
- **Colluding Nodes:** Sender and receiver can negotiate to misbehave as a pair, in schemes that trusts the sender or receiver or both [6]. Detection of such misbehaviours can be complicated.

MAC layer misbehaviours have been studied from different perspectives using different methodologies. Many researches have focused on solving the MAC layer misbehaviour problems by modifying the existing IEEE 802.11 protocol [1] [2]. These procedures include changing the Binary Exponential Back-off (BEB) algorithm, properties of the CSMA/CA control packets and the authority of back-off value allocation to the receiver [4]. In contrast, statistical inference based detection techniques do not modifying the underlying protocol architecture. Instead these techniques gather the protocol transaction data to analyse misbehaviour [8].

Research done in [4] [5] has identified out many problems, such as receiver misbehaviour, colluding nodes and adaptive misbehaviour. In their approach [4] the receiver assigns back-off value to the sender and monitors the sender's behaviours. The research carried out by Radosavac et al. in [6] presented their work based on the previous study in [4]. Their protocol has addressed the major drawback in previous proposal in [4] which was assumed the receiver is trusted. Their approach [6] was that sender and receiver agrees through a public discussion on random back-off value. Protocol always ensures that honest party agreed value is truly random. But it is assumed in [6] that one of the parties has to be trustworthy and the honest receiver can monitor the behaviour of the sender and identify the deviation [6]. However, this approach failed to detect colluding nodes. Smart selfish misbehaviour detection method

has presented in [9] with a predictable random back off algorithm that can mitigates the effect of the smart MAC layer misbehaviour. Research done by Rong et al. in [8] have explained how statistical and probability models can be utilized to detect cheating stations. This approach has used Bianchi Stochastic Model to build a probability distribution model for packet inter-arrival times.

3 Proposed Secure MAC Protocol and Detection Mechanism

This project aims to propose a secure MAC protocol design, which can be integrated with a novel MAC layer misbehaviour detection and avoidance mechanism. Secure MAC protocol is a novel approach as firstly, the sender and receiver handshake prior to deciding the back-off value but the receiver has the authority to decide the final value. This negotiation requires a mechanism to stop each sender and receiver from generating small back-off values. The secure MAC protocol consists of a statistical analyser as a first line of defence to detect generation of small back-off values. Secondly, secure MAC protocol effectively uses common neighbours to detect misbehaviours at the MAC layer. Common neighbours (CNs) actively work with the sender and receiver in the process of monitoring, detecting and penalizing. In addition the mechanism introduces a new trust distribution in the network by assuming none of the wireless nodes need to trust any other. This trust model is a process of involving CNs to construct a trust distribution to the network. This research also assumes that misbehaving nodes are having significant levels of intelligence to avoid the detection. This research addresses sender misbehaviour, receiver misbehaviour and also colluding node misbehaviour.

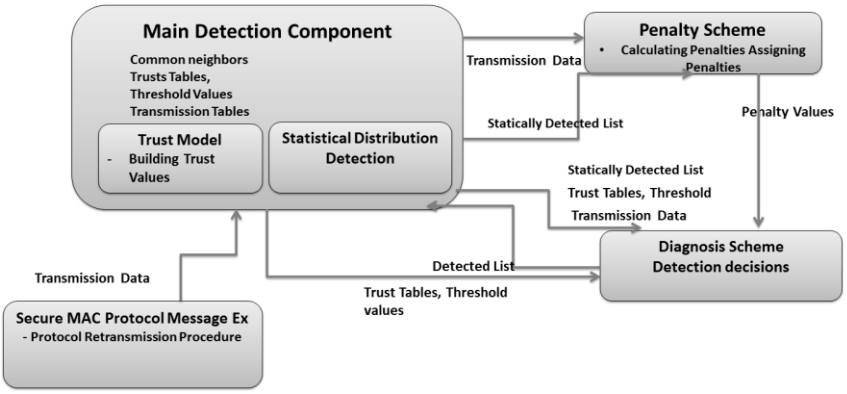

Fig. 1. Secure MAC Protocol and Detection Framework

Fig. 1 demonstrates the proposed Secure MAC protocol. Firstly, the **main detection component** does all the analysis of network traffic and manipulation of CNs data. This main module communicates with other components such as penalty scheme, diagnosis scheme and secure MAC protocol message exchange. This component con-

tains the trust model and statistical data analyser which act as the lowest level of detection technique. Secondly, **penalty scheme** is the module that assigns penalties for deviating senders in each transmission. The third component is **diagnosis module** which performs the detection operation based on the data received by the main detection component, penalty scheme and Trust model.

3.1 Common Neighbours (CNs) and Trust Model

Wireless nodes in the transmission range of both sender and receiver are considered as CNs. The main rational behind using CNs is to build a trust model for the network, monitoring and reporting on node behaviour. For example in **Fig. 2**, nodes 2, 3, 6, and 8 are CNs of nodes 5 and 7. CNs monitoring eliminates most of the unwanted communication overhead. CNs keeps records of the transactions of different sender receiver pairs for a period of time. **Table 1**contains all the transmission details recorded in a neighbour node (node 3 in **Fig. 2**) such as communication ID, expected back-off value (BOV), average sender access times and deviation factor.

According to the Table 1, there are communication entries that were recorded from multiple sender receiver pairs. Ex: Node 3 records communications between sender receiver pairs such as (5-7), (8-7) and (5-3). In **Fig. 2** wireless node 3 records all the transactions of sender receiver pairs of which node 3 is a CN.

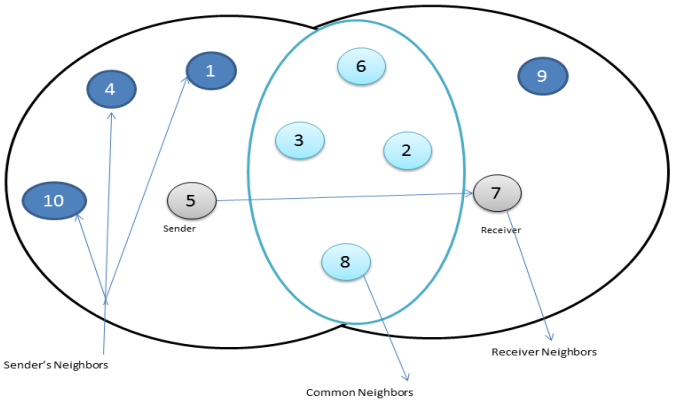

Fig. 2. Distributed Cooperate Detection Mechanism Common Neighbours

Communication ID	Expected BOV	Deviation Factor (∂)	Sender Access
5 → 7	5	4	5
8 → 7	10	4	7
5 → 3	11	10	10

Table 1. Node 3 records during the transmission period

Trust model is the process in which each wireless node builds trust among each other. This trust model assumes that no one has to trust anybody but trustworthiness gradually changes when nodes start to follow the protocol. The trust model is designed in such way that if a node collaborates to detect adversary nodes (observing and reporting) then the node is able to maintain a fair share. Deviation factor repre-

sents the selfishness of the given transmission. If the transmission fails to adhere to the expected back-off value then the deviation factor will be increased by 1 if not it will be decreased by 1. The following equation (**Eq. 1**) calculates trust value based on data from **Table: 1**. Node 5 appears twice in the **Table: 1**, let's define that as "**n**" deviation factors (λ) of each of these appearances. The maximum value of (λ) is 10 so each value divided by 10 to get a probability for the node deviation**.** Table: 2 also maintain the nodes status in the network which has defined as normal state (NML) and misbehaving state (MSB).

$$\text{Trust Value (\%)} = \frac{\sum_{k=0}^{n} \lambda_k / 10}{n} \tag{1}$$

Node Id	Trust Value (%)	Status
8	50	NML
5	70	MSB

Table 2. Common Neighbour (node 3) Trust value table

3.2 Secure MAC Protocol Message Exchange

The standard IEEE 802.11 protocol message exchange has been modified and re-designed as a security embedded protocol. Provided that DCF control packets such as RTS, CTS, DATA, and ACK been modified to add more data fields. The major inspiration behind modifying the protocol is to change the operation of back-off value calculation and allocation authority, while allowing CNs to be involved in the protocol operations.

3.3 Penalty Scheme (Misbehaviour Avoidance) and Voting Policy(VP)

Penalty scheme is important to discourage wireless node misbehaviours and forces adversary nodes to follow the standard protocol once they have violated it. A penalty value needs to be assigned whenever a node deviates from the protocol. In this proposed protocol, receiver assigns the penalty value to the sender. Voting Policy (VP) is introduced to the penalty scheme to detect colluding neighbours and also to minimize misdiagnosis. VP operates at the receiver end by broadcasting a request to obtain the trust value of the relevant sender in each CN. If sender expected back-off value (B_exp) and the actual back-off value observed by receiver and CNs is B_exp. Following equation (Eq. 2) calculates the penalty value of node id "i". The variable α is a value (0…1) that minimizes the hidden terminal effect in monitoring actual back-off values by the receivers and CNs. Numbers of CNs are defined by "N" and TRV is the trust value of nodes that appear in Table.1. Then the receiver calculates new back-off value by adding penalty to the next transmission.

$$Penalty = (\text{B\_act} - \alpha \text{B\_exp}) * \frac{\sum_{j=1}^{N} (\text{TRVj})}{N} \tag{2}$$

$$\text{NewBackoffValue} = \text{BOV} + Penalty \tag{3}$$

3.4 Diagnosis Mechanism

In this research there are two levels of detection to avoid selfish misbehaviours at MAC Layer. Firstly, detection at random numbers generation level**.** This is a statistical inference detection approach to prevent the nodes to generate small back-off values out of well-known function's distribution. If a node generates such back-off

values, then there is a high probability that statistical detection module will name this node as a misbehaving node or increase node's deviation factor. Secondly, detection based on CNs's behaviour monitoring and trust model. After a certain time period and certain number of transmissions the trust values must have saved for each node at each CN, CNs then report their recorded trust values to the diagnosis module for the decision. Diagnosis mechanism calculates the average trust value and makes a decision of node's behaviours.

4 Conclusion

This research has proposed a novel design for a secure MAC layer protocol which is aimed to be resilient for MAC layer misbehaviours. The proposed protocol also assumes that wireless nodes can follow some adaptive misbehaviour strategies to avoid the detection. This research has proposed a novel detection approach which includes the CNs monitoring, trust model and penalty scheme and diagnosis scheme. In the future work, performance of the proposed protocol will be evaluated using ns2 network simulation.

5 References

1. Lolla, V.N., Law, L.K., Krishnamurthy, S.V., Ravishankar, C., Manjunath, D.: Detecting MAC Layer Back-off Timer Violations in Mobile Ad Hoc Networks. In: 26th IEEE International Conference on Distributed Computing Systems, pp.63-63, ICDCS(2006)
2. Jabbehdari, S., Sanandaji, A., Modiri, N.: Evaluating and Mitigating the Effects of Selfish MAC layer Misbehaviour in MANETs. In: IEEE Wireless Communications Letters, vol. 4, 2012
3. Rong, Y.: Detecting Mac Layer Misbehaviour and Rate Adaptation in IEEE 802.11 Networks: Modeling and Sprt Algorithms. In: Ph.D. Dissertation. George Washington Univ., Washington, Choi H., DC, 2008
4. Kyasanur, P.; Vaidya, N.F.: Selfish MAC layer Misbehaviour in Wireless Networks, In: IEEE Transactions on Mobile Computing, vol.4, no.5, pp.502-516, Sept.-Oct. (2005)
5. Gunasekaran, R., Uthariaraj, V.R., Sudharsan, R., Sujitha Priyadarshini, S., Yamini, U.: A distributed mechanism for handling of adaptive/intelligent selfish misbehaviour at mac layer in mobile ad hoc networks. In: Journal of Computer Science and Technology, vol. 24, pp. 472 - 481, 2009
6. Radosavac, S., Cardenas, A. A., Baras, J. S., and Moustakides, G. V.: Detecting IEEE 802.11 MAC layer Misbehaviour in Ad hoc Networks: Robust strategies against individual and colluding attackers. In: Journal of Computer Security, vol. 15(1), pp. 103-128, 2007
7. Rong, Y., Lee, S.K., Choi, H.A.: Overview and Challenges of Routing Protocol and Mac layer in Mobile Ad-hoc Network. In: Journal of Theoretical and Applied Information Technology, 2009
8. Rong, Y., Lee, S.K., Choi, H.A.: Detecting stations cheating on Backoff rules in 802.11 Networks using Sequential Analysis. In: INFOCOM. IEEE, (2006).
9. Guang, L., Assi, C., Benslimane, A.: MAC layer Misbehaviour in Wireless Networks: Challenges and Solutions. In: Journal of Computer Security, vol. 14(4), pp. 6-14, 2008
10. Qin, L., Kunz, T,: Survey on Mobile Ad hoc Network Routing Protocols and Cross-layer Design. In: Carleton University, Systems and Computer Engineering, Technical Report, SCE 04-14, 2004

To Build Light Gamification upon Social Interactions: Requirement Analysis for the Next Version of Topolor

Lei Shi, Dana Al Qudah, Alaa Qaffas and Alexandra I. Cristea

Department of Computer Science, University of Warwick
CV47AL Coventry, United Kingdom
{lei.shi, d.al-qudah, aqaffas, acristea}@dcs.warwick.ac.uk

Abstract. The introduction of social dimension enables traditional adaptive educational hypermedia systems to provide more versatile personalized services. Topolor has been developed to investigate the impacts of social interaction and feasible engagement strategies in such a system. We have evaluated Topolor's social features from the perspectives of *usefulness* and *ease of use*. We intend to develop the next version of Topolor, starting with enhancing relatively lower rated social features. This paper presents our plan of building light gamification upon the evaluated social interaction features with relatively lower rating.

Keywords: adaptive educational hypermedia, social e-learning, gamification.

1 Background

Adaptive educational hypermedia systems can provide rich learning experience with educational hypertext and hypermedia adapted to learners' personal needs [1]. By introducing a social dimension, it becomes possible to offer more personalized services, e.g., social interaction tool and learning peer recommendations [2]. Topolor is a social personalized adaptive e-learning system that builds social features upon an adaptive educational hypermedia system [3, 4]. It is under iterative development, and its first version has been used as an online learning system for the MSc level students in the Department of Computer Science at the University of Warwick [5]. We have evaluated its first version from various perspectives and the results showed high satisfaction from the students [6, 7, 8, 9]. At this stage, we aim to focus on the relatively lower rated features, and improve them to make Topolor more engaging.

Gamification is defined as *the use of game design elements in non-game contexts* [10]. It describes an effective way of engaging users and motivating their behaviors using game mechanisms and/or elements. The use of gamification in enhancing engagement in an e-learning environment has gained traction in recent years, and its benefits have been stated by a number of researches [11, 12, 13]. Hence, we expect gamification can also help with the improvement of those identified features as well as enhance engagement in Topolor. This paper presents our plan of introducing gamification into Topolor, focusing on building light gamification mechanisms upon those identified social interaction features with relatively lower rating.

2 Evaluation Summary of Social Interaction Features in Topolor

As a social personalized adaptive e-learning system, Topolor[1] provides not only adaptation services, such as learning topic adaptation and learning path adaptation, but also social interaction features to support social e-learning and collaborative e-learning [14]. For example, learners can post and share their own learning status, as well as favorite, share and common on their peers' learning status. They can send messages to each other; they can also ask questions and answer others' questions.

We have identified 110 types of actions that students could perform in Topolor, 18 of which were annotated as social interaction actions, as shown in Table 1. We have evaluated each action's *usefulness* and *ease of use*, using a Likert Scale questionnaire. For each action, a student had to assign a score (1: very useless - 5: very useful) to its *usefulness*, and a score (1: very hard - 5: very easy) to its *ease of use*.

Table 1. Social Interaction Actions Annotated in Topolor

Learning Status	Create, Edit, Remove, Comment on, Favorite, Share
Message	Send, Reply
Q&A – Ask	Create, Edit, Remove, Share, Favorite, Add Tag, Edit Tag
Q&A – Answer	Create, Edit, Remove

The experiment was conducted in the Department of Computer Science at the University of Warwick, with the help of 21 MSc students. The optional questionnaires were collected after a time-controlled (2 hours) online learning session. Finally, we collected 10 questionnaires to analyze. The results shown that all the evaluated actions were positively rated (score ≥ 3, the natural score), as the minimum score for usefulness was 3.70 (mean = 4.30, standard deviation = 0.25) and the minimum score for ease of use was 3.80 (mean = 4.15, standard deviation = 0.29). We have examined the evaluation results' reliability. The *Cronbach's Alpha* value for *usefulness* was 0.934, and that for *ease of use* was 0.948. Both are larger than 0.8, suggesting a high level of reliability of the results [15]. We focus here only relatively lower, (whilst still above average) rated actions, as shown in Table 2, with minimum values in bold.

Table 2. Ranking list of the Usefulness and Ease of Use of the Evaluated Actions

Action	Mean value of *usefulness*	Mean value of *ease of use*
Favorite a learning status	**3.7**	**3.9**
Share a learning status	**3.9**	4.2
Edit a learning status	4.2	**3.9**
Edit a question	4.4	**3.9**
Edit a tag on a question	**3.8**	4.1

1 http://www.topolor.com

3 Building Light Gamification Mechanisms

Gamification is implemented for creating more interest, attention and interaction to make a system more engaging [16]. Light gamification mechanisms in this paper literally mean that we intend to introduce gamification as a solution to symbiotically make Topolor easier to use and more engaging, rather than replace its social learning community. In this section, we will specifically address the initial reasons for starting this work, propose three light gamification mechanisms, and depict the scenarios how these mechanisms work in the system.

From the evaluation results we can see that the students felt positively about 'editing', but they liked 'editing' slightly less than the other social interaction actions. 'Editing' features we considered minor features in Topolor, because in a modern social system, what's more important is the timeliness of posts, and it does not make sense to dramatically update a post which has already had comments. For instance, Facebook[2] and Twitter[3] do not allow users to edit their posts, but users can delete their posts. However, these 'editing' features somehow influence the overall satisfaction of the students. Therefore it is necessary to provide mechanisms, which can guide users to use the provided features, especially those that are more complicate while still useful. Inspired by this, we intend to introduce a so-called *tip mechanism*, in order to help the students to get used faster to various features, as well as a *badge mechanism* that encourages them to explore more features.

The evaluation results also reveal some issues about favorite-ing and sharing a learning status, which were designed for the engagement of informal communication and collaboration. While the *tip mechanism* and the *badge mechanism* might make such features easier and more pleasant to use, there may be side effects such as the 'noise' produced. For instance, students may abuse such features only to earn more badges, rather than for learning purposes only. Therefore, to reduce the 'noise', we intend to introduce another mechanism, a *peer-review mechanism*.

Details about the 3 proposed gamification mechanisms are as following:

Tip Mechanism. The first time that a student views a part of a user interface where there is/are a potential feature(s) to use, Topolor will pop up a canvas to remind the student to use the feature(s), and then, if the student clicks 'details', the pop-up will expand to show more information about how to use the feature(s), step by step. These tips will be packaged as missions [17] that a student may be willing to achieve, and when a student completes a mission, s/he will be awarded for the engagement. Additionally, these *tips* can also show the explanation of why the student is offered a specific learning topic, path or peer recommendation, which are expected to improve Topolor's usability. In such a way, the *tips mechanism* acts as a navigator guiding students to use various features in Topolor, and make it more engaging.

Badge Mechanism. Badges are among the most visible elements of gamification. The process of earning badges engages participation, and the exhibiting of badges earned can cultivate an environment of collaborative and competitive e-learning [18, 19]. There will be three ways to earn badges in Topolor: one is to earn by completing a

2 http://www.facebook.com
3 http://www.twitter.com

task of (a) learning topic(s); another is to earn by trying Topolor's features (in association with the *Tip Mechanism* mentioned above); the third one is to earn by helping peers, namely, awarding for the social interaction, e.g., answer a question asked by her/his learning peers. To make the badge mechanism work effectively, we will also design 1) user profile pages to support the visualization of badges that a learner has earned, 2) badge info pages to show what badges can be earned and how to earn them, and 3) user level mechanism to offer privileges according to the badges earned, which might be further extended to a *role-play mechanism*, e.g., along with earning more badges, a learner can 'grow-up' from a 'pupil' with limited privileges to a 'senior scientist' with full privileges.

Peer-review Mechanism. This mechanism will actually be implemented for two objectives. One is to prevent learners from abusing features in Topolor, as mentioned above; the other is to improve the quality of posts. For instance, when posting a learning status or when asking or answering a question, peers can cast a vote on the post's quality as well as click 'Like' button to praise its contribution. The results from the *peer-review* will be finally transferred to the credit to earn badges, which can enhance a learner's reputation, so that s/he has greater weights in determining peer posts' quality. It can also inform the display options of posts (e.g., high quality first).

4 Conclusions and Future Works

In this paper, we have revisited the evaluation results of social interaction features in Topolor, focusing especially on the lower rated features. We have also proposed three light gamification mechanisms to improve those identified features. These mechanisms are also expected to influence the entire system, so that the next version of Topolor can support more engaging social personalized adaptive e-learning.

The follow-up work has started already with the implementation of these three proposed light gamification mechanisms. Then we will conduct experimental studies on the evaluation of the new version of Topolor, analysis on learning behaviors, and comparison between the two versions of Topolor, in order to investigate the impact brought by the newly introduced light gamification mechanisms. Care will be employed to differentiate between the effects of these different mechanisms. To compare the two versions of Topolor, we will consider various criteria and evaluate them from more perspectives. For instance, besides the evaluation of usefulness and ease of use, mentioned in this paper, learning success, task performance, improvements in communication, etc. will be considered as well.

References

1. Brusilovsky, P., and Eva, M.: User models for adaptive hypermedia and adaptive educational systems. J. The adaptive web. 3–53 (2007)
2. Shi, L., Al-Qudah, D. A., Cristea, A. I.: Designing Social Personalized Adaptive E-Learning. In: 18th Annual Conference on Innovation and Technology in Computer Science Education, pp. 341–341. ACM (2013)

3. Shi, L., Al-Qudah, D. A., Qaffas, A., Cristea, A. I.: Topolor: A Social Personalized Adaptive E-Learning System, In: 21st Conference on User Modeling, Adaptation and Personalization. In: Weibelzahl, S., Carberry, S., Micarelli, A., Semeraro, G. (eds.) LNCS, vol. 7899, pp. 338–340. Springer, Heidelberg (2013)
4. Shi, L., Al-Qudah, D. A., Cristea, A. I.: Social E-Learning in Topolor: A Case Study. In: IADIS International Conference e-Learning, pp. 57–64. IADIS Press (2013)
5. Shi, L., Gkotsis, G., Stepanyan, K., Al-Qudah, D. A., Cristea, A. I.: Social Personalized Adaptive E-Learning Environment - Topolor: Implementation and Evaluation. In: 16th International Conference on Artificial Intelligence in Education. LNCS, vol. 7926, pp. 708–711. Springer Berlin Heidelberg (2013)
6. Shi, L., Stepanyan, K., Al-Qudah, D. A., Cristea, A. I.: Evaluation of Social Interaction Features in Topolor - A Social Personalized Adaptive E-Learning System. In: 13th IEEE International Conference on Advanced Learning Technologies, pp. 294–295. IEEE Computer Society (2013)
7. Shi, L., Cristea, A. I., Awan, M., Stewart, C., Hendrix, M.: Towards Understanding Learning Behavior Patterns in Social Adaptive Personalized E-Learning Systems, In: 19th Americas Conference on Information Systems, pp. 1–10. Association for Information Systems (2013)
8. Shi, L., Awan, M., Cristea, A. I.: A Component-Based Evaluation Framework for Social Personalized Adaptive E-Learning Systems. In: 8th European Conference on Technology Enhanced Learning, pp. 633–634. Springer Berlin Heidelberg (2013)
9. Shi, L., Awan, M., Cristea, A. I.: Evaluation of Social Personalized Adaptive E-Learning Environments: From End User Point of View. In: Andrew V. Jones (eds.) 3th Imperial College Computing Student Workshop, pp. 103–110. Schloss Dagstuhl - Leibniz Center for Informatics (2013)
10. Deterding, S., Dixon, D., Khaled, R., Nacke, L.: From game design elements to gamefulness: defining gamification. In: 15th International Academic MindTrek Conference: Envisioning Future Media Environments, pp. 9–15. ACM (2011)
11. Lewis Hobart, M.: Learning from Myself: Avatars and Educational Video Games. J. Current Issues in Education, 15, 3 (2012)
12. Domínguez, A., Saenz-de-Navarrete, J., De-Marcos, L., Fernández-Sanz, L., Pagés, C., Martínez-Herráiz, J. J.: Gamifying learning experiences: Practical implications and outcomes. J. Computers & Education, 63, 380–392 (2013)
13. Kangas, M.: Creative and playful learning: Learning through game co-creation and games in a playful learning environment. J. Thinking Skills and Creativity, 5(1), 1–15 (2010)
14. Sun, P. C., Tsai, R.J., Finger, G., Chen, Y. Y., Yeh, D.: What drives a successful e-Learning? An empirical investigation of the critical factors influencing learner satisfaction. J. Computers & Education, 50(4), 1183–1202 (2008)
15. Carmines, E. G., & Zeller, R. A.: Reliability and Validity Assessment, Sage, vol. 17 (1979)
16. Deterding, S., Dixon, D., Khaled, R., Nacke, L.: From game design elements to gamefulness: defining gamification. In: 15th International Academic MindTrek Conference: Envisioning Future Media Environments, pp. 9–15. ACM (2011)
17. Kim, Bokyeong, Hyungsung Park, and Youngkyun Baek.: Not just fun, but serious strategies: Using meta-cognitive strategies in game-based learning. J. Computers & Education, 52(4), 800–810 (2009)
18. Singer, L., Schneider, K.: It was a bit of a race: Gamification of version control. In: 2nd International Workshop on Games and Software Engineering, pp. 5–8. IEEE (2012)
19. Domínguez, A., Saenz-de-Navarrete, J., De-Marcos, L., Fernández-Sanz, L., Pagés, C., & Martínez-Herráiz, J. J.: Gamifying learning experiences: Practical implications and outcomes. J. Computers & Education. 63, 380–392 (2013)

The Behavioural Uncanny Valley: Social Acceptance and Life in Cognitive Robotics

James Stovold[1], Simon O'Keefe[1], and Jon Timmis[2]

[1] Department of Computer Science, University of York
[2] Department of Electronics, University of York
{jhs503,simon.okeefe,jon.timmis}@york.ac.uk

Abstract. As we move to make robots more human-like, and further integrate them into society, it is becoming more important to consider what the longer-term effects are going to be of such research. By considering life in robots, we present a discussion of social acceptability of cognitive robotics. We also discuss how human-like behaviour could have a significant impact on the acceptability of robots.

1 Introduction

Cognitive robotics aims to produce a robotic system that can reliably make intelligent, informed decisions in the real world, without the need for human intervention, by constructing cognitive abilities in robots. This differs from the traditional approach to robotics in that, instead of specifically programming a robot to perform a task, robots are programmed with cognitive abilities and can decide to perform the appropriate task for a given situation. Because of the lack of an agreed definition of cognition, authors working in the area of cognitive robotics work towards cognition by building systems that exhibit certain properties that are generally accepted as being cognitive. For example, the ability to make decisions or to learn [3]. If it is possible to get to cognition by progressively working towards it, what might happen if we start to consider life and living robots? Many robots are increasingly behaving in a more life-like way, exhibiting more and more properties that could be considered 'living', in particular increased intelligence and emotional responses.

As we move to make robots more intelligent, and more human-like, and as we further integrate them into society (i.e. [8]), it will become important to consider what the long-term effects are going to be of such research. By considering what might happen if it was possible to build a living robot, this paper presents a discussion of social acceptability in terms of the behavioural version of the uncanny valley [6], and discusses the implications of robot life that stems from cognitive robotics research. The paper is organized as follows: section 2 discusses life, section 3 introduces the uncanny valley models, and the paper concludes with some discussion in section 4.

2 Life

> "Just as we know of sight only in connection with a particular kind of material system called the eye, so we know only of life in connection with certain arrangements of matter, of which the biochemist can give a good, but far from complete, account." [5] (p. 10)

This quote by J. B. S. Haldane, along with the short story 'They're made out of meat' by Terry Bisson [1] encapsulates the difficulty we have, as humans, to visualize how a living robot could be possible. On top of this, we tend to define how alive something is by judging how human-like it is (i.e. it has cells, can metabolize, can reproduce etc.). This anthropocentric view of life stems from our frame of reference and it is incredibly difficult to break free from this. The quote by Haldane is discussing a concept known as 'carbon chauvinism' [9], where many humans have no way of being able to comprehend any other basis for life than carbon molecules. While we have various definitions of life, ranging from a series of properties something must exhibit [7] to a simple, physical definition based on entropy [10], these ideas emphasize that we still don't really know what life is.

Considering those in cognitive robotics are working towards cognition by building systems that exhibit similar properties, it seems likely that—whether intentionally or not—we are also working towards living robots by building systems that are increasingly behaving like they are living. There are a number of potential issues associated with this trend, including how well these systems will be received by the general public as they become more life-like.

3 The Uncanny Valley

The 'uncanny valley' is a theory proposed by Mori [6] based on the experience of humans who are unsettled by artificial beings that are very similar to humans, yet are not quite realistic enough to convince us they are human (see Fig. 1). This can result from something as simple as the skin on a model being too waxy, making the model look less familiar. Mori's theory is that this unfamiliarity is what unsettles us when we (for example) shake a prosthetic hand. While Mori describes how movement can exacerbate the effects of this valley, he offers no description of how behaviour can also cause such reactions. Wilson [12] on the other hand, describes how recent work in artificial intelligence is slowly moving towards an uncanny valley of its own, based on the behaviour of robots and virtual models. This 'behavioural uncanny valley' could pose significant problems for cognitive robotics research.

While the technology for socially-able robots is still years away—and so no experiments have been run to confirm the theory presented—there are other clues that the social behaviour of robots could cause issues. The work presented by Dautenhahn [4] shows that being in the same vicinity as a robot can be enough to trigger discomfort in humans. Walters *et al.* [11] report experiments

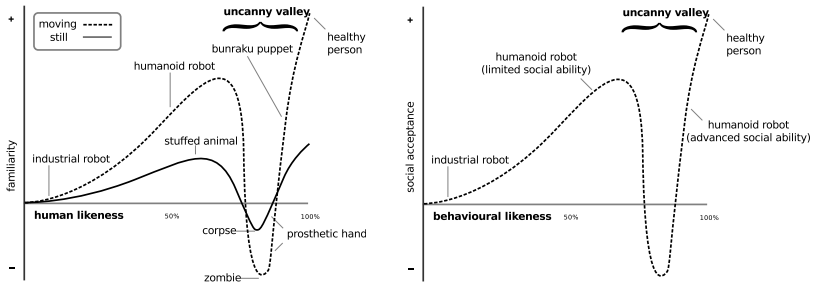

Fig. 1. Graph showing the 'Uncanny Valley' (left). As a model approaches full human likeness, human acceptance of it rapidly drops away until it is almost indistinguishable from humans, at which point its acceptance shoots back up again. (Right) The behavioural uncanny valley. The two proposed approaches to avoiding the robotic *faux pas* are marked on either side of the valley. Edited from [6].

looking at how comfortable humans are when approached by a robot, with some responses from subjects suggesting that there were behaviours that made them uncomfortable. These studies suggest that the behaviour of robots when interacting with humans is not only important, but essential to the acceptance of robots in society.

3.1 Society

Human society maintains a (relatively strict) set of unwritten rules, everything from how to appropriately greet somebody in different situations to which urinal to use in a public toilet. The breaking of said rules, even if the intention was clearly appropriate, is considered a *faux pas*, and can very quickly alienate somebody from society. A key part of judging the appropriate response to a given situation is the emotional interplay amongst members of society.[3]

With robots behaving increasingly like humans, they will need to be able to behave appropriately according to the social situation they are in. If they get this wrong, they risk alienating themselves from the society they are trying to integrate into. There are two potential paths to avoiding this problem: either (a) the robots need to be advanced enough that they can determine the appropriate social etiquette, or (b) they can be simple enough that they aren't judged by the same standards as other members of society. These two points sit on either side of the behavioural uncanny valley (see Fig. 1), and while approach (a) is desirable for full integration of robots, the approach (b) is far simpler, although would arguably not integrate robots into society at a similar level to humans. While we are only discussing social acceptance from behaviour in this paper, it is envisaged that other aspects of human life would also contribute to their

[3] Most of us have got this wrong at some point, whether it is laughing at a funeral or getting upset at a film that is not intended to be emotional. It feels unsettling to us because we know it is not the emotion we should be feeling at this point.

social acceptance, and so would also be susceptible to the behavioural uncanny valley.[4]

Although we are moving towards cognition, emotions and other aspects of human life in robots, they do not necessarily need to be real. Breazeal and Brooks [2] have developed an architecture for a robot to display emotion, yet the robot itself is not actually experiencing the emotion. This 'fake it until you make it' approach to developing emotions is key because, from the external perspective, whether a robot is actually experiencing an emotion or just mimicking is not always relevant. If a robot is mimicking anger and breaks a vase, or if the robot is actually angry and breaks a vase, the outcome is still the same: the vase gets broken. However, this differentiation could ease the impact of the uncanny valley, as it could mean the robot will be treated differently by members of society.

4 Discussion and Conclusion

Social acceptance of robots is just one part of the future of robotics. As robots become more life-like, we are likely to encounter similar issues across the board, and the concept of a robot having a life becomes closer to reality. For example, consider what happens with a learning robot: the robot will, over the course of its service, learn how to do a number of different tasks. If we were to take that robot and copy its memory into the shell of another robot, the chances of it working as before are almost zero. The robot will not only have learnt how to perform tasks, but it will have learnt how to perform those tasks based on data coming from those particular sensors, and with its specific actuators, and will be used to handling a certain level of noise, which will be completely different in another shell. From an external perspective, it could be argued that the robot has been 'living' in its hardware shell too long, and any attempt to move it is likely to 'kill' it. While this is described using terms that are usually associated with living beings, the concepts are remarkably similar.

We have discussed social acceptance of cognitive robotics from the perspective of perceived behavioural appropriateness. The behavioural uncanny valley is a helpful tool for visualizing how different social abilities might affect social acceptance. Two approaches are offered for helping to avoid the behavioural uncanny valley: either limiting or significantly advancing the social abilities of the robot. They both offer advantages, yet limiting these abilities is a far simpler approach, and given that some humans still struggle with social acceptance it seems that this is a far more difficult problem than has been given justice here.

Finally, although it is unlikely that robots will ever be classified as alive, it seems that we are still moving towards this, as it is possible for them to exhibit many of the same properties, including 'death' as they are removed from the shell that they have learned to interact with the environment through. Whether robots need their own definition of 'life', or if they can be integrated into existing definitions of life remains an open question.

[4] Robots may be expected to make mistakes, or even to lie in certain situations in order to better fit into society in the future.

References

1. Bisson, T.: They're made out of meat. OMNI Magazine (1990)
2. Breazeal, C., Brooks, R.: Robot emotion: A functional perspective. In: Fellous, J.M., Arbib, M.A. (eds.) Who Needs Emotions? The Brain Meets the Robot, pp. 271–310. Oxford University Press (2005)
3. Cohen, I.R.: Tending Adam's Garden: evolving the cognitive immune self. Academic Press (2000)
4. Dautenhahn, K.: Socially intelligent robots: dimensions of human–robot interaction. Philosophical Transactions of the Royal Society B: Biological Sciences 362(1480), 679–704 (2007)
5. Haldane, J.B.S.: Science and life: essays of a rationalist. Pemberton Publishing in association with Barrie & Rockliff (1968)
6. Mori, M.: The uncanny valley. Energy 7(4), 33–35 (1970)
7. NASA: Phoenix mars mission. Online: `http://phoenix.lpl.arizona.edu/mars141.php`
8. Prassler, E., Ritter, A., Schaeffer, C., Fiorini, P.: A short history of cleaning robots. Autonomous Robots 9(3), 211–226 (2000)
9. Sagan, C.: Carl Sagan's cosmic connection: an extraterrestrial perspective. Cambridge University Press (2000)
10. Schrödinger, E.: What is Life? The physical aspect of the living cell. Cambridge University Press (1944)
11. Walters, M.L., Dautenhahn, K., Woods, S.N., Koay, K.L., Te Boekhorst, R., Lee, D.: Exploratory studies on social spaces between humans and a mechanical-looking robot. Connection Science 18(4), 429–439 (2006)
12. Wilson, M.: Artificial intelligence is the new uncanny valley. Online: `http://www.fastcodesign.com/1672270/artificial-intelligence-is-the-new-uncanny-valley` (April 2013)

Quantifying the Risk of Multi-tenancy in the Cloud

Abdulaziz Alabdulhafez and Paul Ezhilchelvan

School of Computing Science, Newcastle University, UK

{Abdulaziz.Alabdulhafez,paul.ezhilchelvan}@ncl.ac.uk

Abstract. This paper describes a current piece of research that aims to develop a simulation-based framework to quantify the risk of virtual machines (VMs) co-residency, particularly within the area of multi-tenancy infrastructure-as-a-service (IaaS) cloud computing platforms. Multi-tenancy IaaS clouds allow the sharing of physical resources between multiple users. However, this can lead to a serious and threatening situation, where a malicious user can allocate his VM to the same machine on which his target's VM is running (i.e. VM co-residency). As a direct result of multi-tenancy, the attacker can launch harmful side channel attacks to steal valuable information such as encryption keys. This risk increases whenever the probability of achieving successful VM co-residency increases and vice versa. This framework will also help our understanding of the influence of different cloud elements on the possibility of achieving successful VMs co-residency. In this paper, a brief introduction to the on-going research is provided, along with a background in terms of similar works. In addition, an explanation of the methodology followed in this research to develop the framework is provided, concluding by showing what has been achieved and the possibilities of future work.

Keywords: Cloud computing, security, multi-tenancy, virtualization

1 Introduction

Cloud computing has come to the attention of many researchers recently as an increasing number of users and organizations have started hosting their services and systems in third-party infrastructure-as-a-service (IaaS) clouds. Public IaaS cloud providers such as Microsoft's Azure and Amazon's EC2, allow their users to rent and control virtualized computing infrastructure on demand, by running virtual machines (VMs) as their own servers on a pay-per-use basis according to the providers' Service Level Agreement (SLA) which defines the cloud services' quality level that must be delivered to the user [1]. Providers use virtualization technology to utilize the hardware infrastructure by enabling multi-tenancy. This means allowing multiple users (tenants) the opportunity to share the same machine in which they run their VMs. However, [2] states that multi-tenancy in public IaaS clouds has been proven to introduce a dangerous new threat in the form of side channel attack. A side channel can be defined as an information leakage that occurs as a result of the sharing of physical

infrastructure with other users, such as the sharing of L2 caches of the machine's memory [3]. Side channel attacks require the attacker to perform the following steps [4]: locate where a specific targeted VM is residing; allocate the attacker's malicious VM in the same physical machine as the target VM and detect that both VMs are co-resident; launch several possible Cross-VM attacks on the victim's VM using side channels. The simplicity of this particular threat has motivated this research.

Considering the aforementioned dangerous threats imposed by co-residency and side channel attacks, and assuming that different elements in cloud environments (e.g. number of hosts, clusters, number of users etc) might have an impact on the co-residency possibility, the main contribution of this research is to introduce a novel simulation-based framework for cloud services providers and users which can be used to analyze and estimate the VM co-residency probability in certain cloud computing settings, and to allow them to understand how the VM co-residency probability can be affected by each cloud element. This framework can help in conducting various quantitative and qualitative simulation-based analyses using the VM co-residency simulator (VCS) which has been implemented for the purpose of this research.

2 Background and related work

The work of Ristenpart et al. [4] is one of the main contributions in the field of VMs co-residency. The results show that there is a noticeable risk when shifting data and service hosting and processing to a third-party in the form of the cloud provider. Using Amazon's EC2 service as a case study, they demonstrated that it is possible to map the internal cloud infrastructure in order to locate the targeted VM's location then launching a number of VM probes until one of these probes becomes co-resident with the targeted VMs. They have also outlined a number of possible side channel attacks which enable the attacker to gather sensitive information from VMs that share the same machine. They have achieved a 40% successful co-residency with the targeted VMs. However the strategy which they have applied is cloud provider dependent, and cannot be used to pre-estimate the probability of achieving co-residency in other cloud platforms. Other works on this field have tried to determine co-residency and develop some countermeasures [5]. Home-Alone has been introduced by [6] to detect and warn cloud users when a malicious co-residency occurs. Another work by [7] has tried to eliminate side channels completely. The rest of the research focuses on exploiting side channels in multi-tenancy clouds such as in the case of theft-of-service attacks [8] and running DoS on the targeted VM [9].

Although a number of pieces of research into VM co-residency have been carried out, the cloud elements that can influence the probability of co-residency as well as estimating this probability, have not been explored yet. This research tries to examine the influence of these cloud elements in order to quantify the co-residency possibility.

3 Framework Specifications

Considering the scope of this research (i.e. studying virtual machine co-residency), a higher risk of co-residency is usually associated with IaaS platforms where a mali-

cious user is able to fully control the attacking virtual machines [4]. Consequently, the research focuses on VM co-residency in public IaaS platforms, with the possibility of studying other cloud platform models in the future. The proposed framework applies an efficient sensitivity analysis of the results of the simulation to assess the impact of each cloud element on the probability of achieving successful co-residency. This framework is expected to help cloud providers and users to assess the potential risk of co-residency, depending on the cloud platform's settings. This research assumes the following hypotheses: there is always a possibility that an attacker will be able to co-reside with targeted VMs in order to launch harmful VMs side channel attack against them. This probability is affected by a number of cloud elements such as the number of hosts, clusters, users and the VM allocation placement being used. Since assessing the probability of co-residency is of an experimental nature [10], the proposed framework applies a simulation-based analysis using a new VM co-residency simulator (VCS) that can model and simulate a detailed VMs co-residency behaviour in IaaS platforms (Figure 1). This simulator has been successfully developed using Java, and it is ready to be used in the proposed framework. Simulation is used as a test bed because it provides a fully controllable and repeatable environment where various settings can be assessed and compared easily.

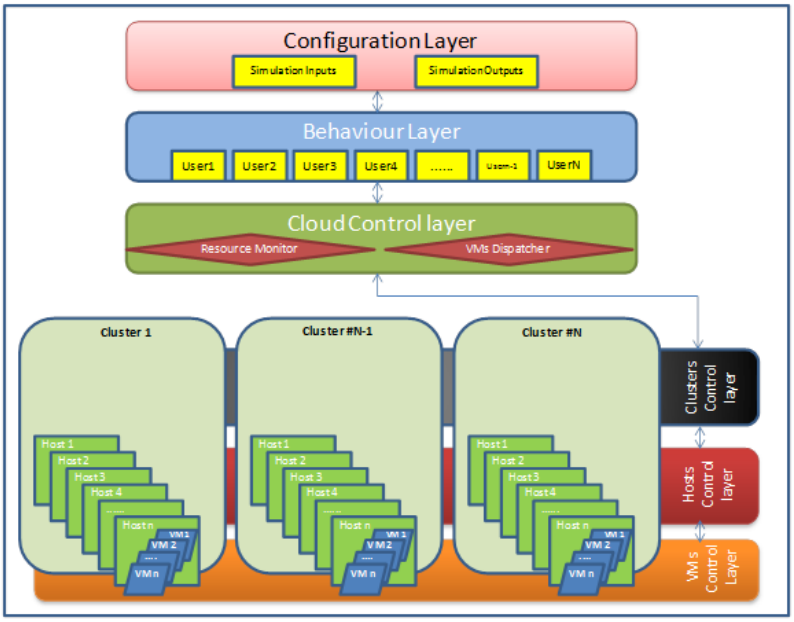

Fig. 1. The VM co-residency simulator's (VCS) internal structure

The framework consists of both qualitative and quantitative analysis of the simulation results in order to quantify the probability of VM co-residency using the Design of Experiment's principles (DoE) [10] as follows:

1. Design a full factorial experiment including as many simulation input variables (i.e. cloud elements) as necessary as the experiment's factors. Assign to each factor

both minimum and maximum values "levels" (i.e. the simulator variables' settings). The use of this type of factorial experiment instead of the one-factor-at-a-time method is an efficient way to evaluate the impact of several factors and their interactions using a smaller number of simulation runs.
2. For each level combination of the experiment design, run the simulator and analyze the VM co-residency results and obtain the Main Effects [10] for each factor. This analysis compares and identifies the level of influence for each of the factors on the co-residency probability under each factors setting combination. It shows which cloud elements (factors) impact the co-residency probability the most and, most importantly, reveals what is the best setting for these factors at which the co-residency probability can be minimized or maximized.

In addition, the simulator can estimate the likelihood of an adversary VM(s) co-residing with a particular VM within a defined time period in a particular cloud environment's setting. This can be estimated using the following steps:

3. For a particular cloud environment's setting, and using the previous simulation data, develop a number of advanced VM placement strategies to exploit the simulated cloud VM allocation polices in order to maximize the likelihood of achieving targeted VM co-residency.
4. After feeding the simulator with the attacking strategies, suggest a time period in which a particular VM (the victim VM) will be running. Then suggest the percentage of the adversary VM that will be created per time unit then run the simulator. This simulation run estimates the likelihood of an adversary VM(s) co-residing with that particular VM within a defined period of time in a particular cloud environment's setting following the suggested attacking strategies.

4 Conclusion and Future Work

This paper outlines the work in progress to develop a simulation-based framework that can be used to quantify the risk in a given IaaS cloud associated with VM co-residency. This framework consists of four steps that involve (1) designing the simulation experiment following the design of experiment principles, (2) simulating the required IaaS cloud, (3) deriving clever methods for achieving VM co-residency by analyzing the simulation results then (4) running the simulation again after applying the derived co-residing methods in order to find out how likely it is that a successful co-residency will occur for a given VM. Currently, the first two steps have been accomplished and a VM co-residency simulator has already been implemented and tested successfully. Furthermore, this framework can help cloud providers to develop more sophisticated countermeasures based on the simulation results, in order to minimize the possibility of malicious users achieving co-residency. It can also help cloud users to pre-examine the probability of VM co-residency in any cloud environment. An interesting application of the proposed framework can include analyzing the impact of a number of cloud elements such the VM allocation policy, the number of cloud users, the number of hosts, and the number of clusters on VM co-residency.

Moreover, the simulation results can reveal the best VM placement algorithm. This can then be used to minimize the possibility of achieving targeted VMs co-residency. Also the simulation results can be analyzed to develop a number of VM placement strategies which can be used by attackers to improve the possibility of VM co-residency and determining if it is worth deploying costly anti co-residency mechanisms such as Home-Alone.

Future work will include completing the work on the framework and conducting an experiment using the implemented simulator by applying different cloud settings to find out what are the most influential cloud elements (e.g. VMs placement policies, Number of Cloud Users, etc.) that influence the VMs co-residency the most, and to what extent they impact on the results. We will also assess the probability of achieving VM co-residency in different IaaS cloud platforms and settings.

5 References

1. Hashizume, K., N. Yoshioka, and E.B. Fernandez. *Misuse Patterns for Cloud Computing*. in *the 2nd Asian Conference on Pattern Languages of Programs*. 2011. Tokyo, Japan.
2. Trend Micro, *Security: Making Virtual Machines Cloud-Ready*. 2010, Trend Micro, Incorporated.
3. Song, D.X.D., D. Wagner, and X.Q. Tian, *Timing analysis of keystrokes and timing attacks on SSH*. Usenix Association Proceedings of the 10th Usenix Security Symposium, 2001: p. 337-352.
4. Ristenpart, T., et al., *Hey, You, Get Off of My Cloud: Exploring Information Leakage in Third-Party Compute Clouds.* in *Proceedings of the 16th Acm Conference on Computer and Communications Security*, 2009: p. 199-212.
5. Amittai, A., et al., *Determinating timing channels in compute clouds*, in *Proceedings of the 2010 ACM workshop on Cloud computing security workshop*. 2010, ACM: Chicago, Illinois, USA.
6. Yinqian, Z., et al., *HomeAlone: Co-residency Detection in the Cloud via Side-Channel Analysis*, in *Proceedings of the 2011 IEEE Symposium on Security and Privacy*. 2011, IEEE Computer Society.
7. Yinqian, Z., et al., *Cross-VM side channels and their use to extract private keys*, in *Proceedings of the 2012 ACM conference on Computer and communications security*. 2012, ACM: Raleigh, North Carolina, USA.
8. Fangfei, Z., et al. *Scheduler Vulnerabilities and Coordinated Attacks in Cloud Computing*. in *Network Computing and Applications (NCA), 2011 10th IEEE International Symposium on*. 2011.
9. Huan, L., *A new form of DOS attack in a cloud and its avoidance mechanism*, in *Proceedings of the 2010 ACM workshop on Cloud computing security workshop*. 2010, ACM: Chicago, Illinois, USA.
10. Schultz, T.A., *An introduction to the design of experiments: A simplified approach.* Quality Progress, 2000. **33**(9): p. 122-122.

Part IV

Poster Presentation Abstracts

Optimization of Cluster Head Rotation in Self-Organizing Wireless Sensor Networks

Arezoo Vejdanparast[1], Ali Karami[2]

[2] Department of Computer Science, University of York, UK
[1] arezoo.vejdanparast@gmail.com
[2] ak552@york.ac.uk

Abstract. A wireless sensor network (WSN) consists of several sensor nodes deployed over a geographical area for the purpose of target tracking and monitoring physical phenomena like temperature, vibration, humidity and so on.[1] The power source of a sensor node supplies the energy needed by the device for data acquisition and data transmission to the base station. This power source often consists of a battery with limited amount of energy. In addition, it is difficult or sometimes impossible to recharge or replace the battery as nodes may be deployed in an unattended or unpractical environment. Therefore, improving the energy efficiency and maximizing the lifetime of the network are the major challenges in WSNs. Clustering is a key technique used to both extend the lifetime of WSNs and make them scalable by forming clusters. However, clustering in WSNs faces several challenges such as proper cluster formation, optimum selection of cluster heads, maintaining intra- and inter-cluster connectivity and establishing cluster head rotation frequency to prevent rapid discharge of cluster heads. LEACH algorithm is one of the fundamental clustering protocols proposed for WSNs.[2] It uses random cluster head selection and frequent cluster head rotation strategy to distribute the energy load uniformly among all sensor nodes. In this work, we propose an energy efficient continuous working time (C.W.T.) strategy that could apply to the data transmission phase of LEACH algorithm to improve its performance. In our model a selected cluster head keeps working continuously until its residual energy reaches a predefined threshold. With this mechanism, the frequency of cluster updates and the energy dissipation of new cluster head establishments can be reduced. The results of simulations demonstrate that using C.W.T. model can reduce the energy consumption of each cluster effectively and increase the system useful lifetime.

[1] I. F. Akyildiz, W. Su, Y. Sankarasubrammaniam and E. Cayirci, "Wireless Sensor Network: a survey", Computer Network 38 (4), (2002).

[2] W. Heinzelman, A. Chandrakasan and H. Balacrishnan, "Energy-Efficient Communication Protocol for WSN", Hawaii International Conference on system science (1), (2000).

The effect of line spacing and text justification on performance and preferences on web search tasks

Sorachai Kamollimsakul, Helen Petrie, Christopher Power

Department of Computer Science, University of York

{sk750, helen.petrie, christopher.power}@york.ac.uk

Abstract. Most research on the effects of text presentation on the web is conducted with English speaking participants on English language web pages. There is a very small amount of research on web pages in other languages, Chinese and Arabic. Currently, there is no research about text presentation on Thai speakers reading Thai web pages. This study investigated the effects of two variables of text presentation on web pages: line spacing and text justification, with UK English speaking and Thai speaking web users. Measures of both performance and preference were taken.

There were 12 UK participants (mean age: 26.42 years, range: 24 - 31 years, SD: 2.4, 9 men, 3 women) and 12 Thai participants (mean age: 28.83 years, range: 20-38 years, SD: 5.3, 3 men, 9 women). The participants were asked to find particular pieces of information on a specially constructed website. The UK participants used an English language version of the website and the Thai participants used a Thai version. To ensure results from the two groups were comparable, the content of the website and all materials in the study were translated using careful quality control mechanisms. Each participant undertook six tasks, one with each of the combinations of three levels of line spacing (1, 1.5, 2) and two levels of text justification (left only and left-right). The order of website presentation was counterbalanced to avoid practice and fatigue effects. Three aspects of performance were measured (task completion rate, time spent reading per page and number of web pages visited) and a 5 point Likert ratings of preferences for the six combinations of line spacing and text justification was also collected.

The results show that line spacing and text justification had no significant effect on performance. There was also no significant effect of nationality and no interaction between the variables. On the preference measure, there was a significant main effect for line spacing, $F(2,44) = 64.92$, $p < .001$, but no main effect for text justification or nationality. In post-hoc pairwise comparisons, single line spacing was significantly lower in ratings than 1.5 line spacing and double line spacing ($p < .001$) while 1.5 line spacing and double line spacing were not significantly different from each another. There was a significant interaction between line spacing and nationality, $F(2,44) = 5.13$, $p < .01$; English speaking participants preferred 1.5 line spacing (Mean rating = 4.17, SD = 0.17) and Thai speaking participants preferred double line spacing (Mean = 3.96, SD = 0.21).

An improved rate control algorithm for SVC with optimised MAD prediction

Xin Lu

Department of Computer Science, University of Warwick, United Kingdom
xin@dcs.warwick.ac.uk

Abstract. Scalable Video Coding (SVC), the scalable extension of the H.264/AVC standard, provides solutions for video applications with different network bandwidths, device capabilities and user demands[1]. Rate control algorithms aim to ensure successful transmission of an encoded bitstream and to make full use of the limited bandwidth. They affect not only the stability of the bit rate, but also the picture quality of the entire video sequence. Bandwidth is a valuable resource, and often there are situations in which the bandwidth is insufficient or the network is unstable. Sometimes even the bitstream comprising the basic layer cannot be transmitted completely, resulting in frame skipping. In these cases, effective rate control is essential. With a proper control scheme, overflow and underflow of the buffer are prevented, which means that frame skipping and wastage of channel resources can be avoided. Furthermore, rate control appropriately allocates the available bits according to the complexity of the image content, so that the quality of the video is maximised.

The most widely used rate control mechanism is JVT-G012[2] which has become the default rate control scheme for the base layer of SVC. There are two models in the scheme: the quadratic Rate Distortion (RD) model and the linear Mean Absolute Difference (MAD) model. In the JVT-G012 algorithm, the MAD is predicted from the MAD values of previous frames. This is not always accurate as the MAD can change abruptly due to scene changes or fast motion. Furthermore, a number of rate control methods for SVC have been proposed. Most of them focus on the improvement of the RD models for Qp estimation and a few consider the refinement of MAD prediction.

In the SVC encoding process, the base layer is encoded first, before the enhancement layers. This leads to the idea that some encoding results of the base layer can be used to inform the coding of the enhancement layers, thus benefitting from the bottom-up coding structure of the standard.

In the proposed algorithm, a novel MAD prediction model is proposed that considers the MAD from previous frames and the MAD of the reference frame in the base layer, together. Experimental results show that the proposed algorithm achieves good rate control and produces an improvement in rate distortion. In particular, compared with the JVT-G012 algorithm, the proposed method achieves a low bit rate mismatch error, and results in higher coding efficiency. Evaluation results show that the proposed rate control algorithm achieves a 0.27dB gain in PSNR or a 4.81% reduction in bit rate compared to the original JSVM implementation.

[1] H. Schwarz et al. Overview of the scalable video coding extension of the H.264/AVC standard, IEEE Trans.Circuits Syst.Video Technol., 17(9):1103-1120, 2007.

[2] Z. Li et al. Adaptive Basic Unit Layer Rate Control for JVT, 7th JVT Meeting, 2003.

Bringing Sketches to Life with EuGENia Live

Joost van Pinxten[1], Louis M. Rose[2], and Dimitris Lagos[2]

[1] Eindhoven University of Technology
j.h.h.v.pinxten@student.tue.nl
[2] University of York
{louis.rose|dimitris.lagos}@york.ac.uk

Abstract. Scientists and engineers love to sketch. Small diagrams with simple symbols are often used in discussions regarding the interactions between the components of a system. Paper sketches can succinctly describe valuable information and, although flexible, they can be hard to keep consistent, both internally (e.g., notational variations can occur in a single sketch) and externally (e.g., contradictions can arise as the same components are modelled differently across sketches). Moreover, paper sketches cannot fully explore the behaviour of a dynamic system[1].

EuGENia Live[2] is a digital sketching tool that allows users to define their own sketching notations. Like many sketching tools, EuGENia Live cannot be used to represent the dynamic properties of systems. To address this drawback, we have extended EuGENia Live so that users can specify behaviour (dynamics) for sketching notations, in addition to appearance (structure). Our initial work in this direction has involved allowing the appearance of a sketch element to be controllable via its property values. Furthermore, the behaviour of a sketching notation can now be described by specifying how to react to discrete events.

We have applied our extensions to EuGENia Live to construct a language for modeling a biological domain - MicroRNA gene regulatory networks[3] - and to simulate several small networks. These networks can now be expressed as a combination of connected diagram elements and corresponding characteristic simulation properties in EuGENia Live. By using existing mathematical models, we have added simulation behaviour to diagrams that we had originally sketched with pen-and-paper. This simulation model can be used as an explanatory device or as a prototype for a more detailed simulation. Our initial results are promising: we were able to accurately reproduce the dynamics of small gene regulatory networks, using a custom sketching notation that was developed in under a week.

[1] N. Mangano, A. Baker, and A. V. D. Hoek, *Calico: a prototype sketching tool for modeling in early design*, Proceedings MiSE '08, pp. 63

[2] L. M. Rose, D. S. Kolovos, and R. F. Paige, *EuGENia Live: a flexible graphical modelling tool*, Proceedings of the 2012 Extreme Modeling Workshop, pp. 15

[3] M. Osella, C. Bosia, D. Cor, and M. Caselle, *The role of incoherent MicroRNA-mediated feedforward loops in noise buffering*, PLoS Comput Biol, vol. 7, 03 2011.

Domain-Specific Languages for Supporting Web Application Development Processes

Athanasios Zolotas[1,2], Richard F. Paige[1], and Chris Wedgwood[2]

[1] Department of Computer Science, University of York, UK
{amz502, paige}@york.ac.uk
[2] Digital Lightspeed Solutions Ltd.
cwedgwood50@yahoo.com

Abstract. Many Web Applications that are developed do not fulfil the actual needs of the clients. This is a result of poor requirements elicitation, poorly defined requirements and incomplete or inadequate testing. Epner[3] claims that 53% of the web applications built do not include the required functionality while only 16% of them meet the business needs. Many solutions have been proposed to help the better elicitation and management of the requirements, and better testing of Web Applications, but these approaches are not widely used in practice.

We argue for the development of a set of Domain-Specific Languages (DSLs) that will be deployed across the different phases of the web application development lifecycle to facilitate the production of correct Web Applications that deliver on the *contract* defined between clients and developers. MDE can be used in addressing the problem. The proposed solution consists of a set of Domain-Specific Languages that will allow clients to express the requirements in *tendered contracts* and define *test cases* that will either confirm that the produced solution fulfils their needs, or will identify where the solution falls short. This process will be aided by different model management techniques like *Model-to-Model* and *Model-to-Text Transformations*. Some additional applications for the DSLs are the following:

Estimate Approximate Implementation Cost: It would be of interest to have an accurate estimation for the cost of the application and the time it needs to be built. This can be done by having predefined cost and time values for different concepts (i.e. the cost/time for implementing a simple database query, the creation of a static page, etc.). The requirements stored in the models will be analysed to produce an approximation of the cost and time needed to built the web application based on these values.

Code Generation: It would be of interest the generation of parts of the code for the web application, based on the requirements that were declared using the requirements DSL (e.g. the UI elements). However, the primary scenario in this project is that the code will be fully written by the company's developers.

[3] Epner, M.: Poor project management number-one problem of outsourced e-projects. Research Briefs, Cutter Consortium 7 (2000)

Investigation into the Errors in the CISPR 12 Full Vehicle Radiated Emissions Measurements Due to Vehicle Directivity

Max Paterson, John F Dawson

Electronics Department, The University of York, UK

mp638@york.ac.uk

Abstract. This project investigates the errors introduced when recording the radiated emissions of current road vehicles due to the directivity of the vehicle body shell. Within modern cars there is an ever increasing array of electronic devices, all of which pose an electromagnetic interference threat to other electronic devices both within the vehicle and also outside of the vehicles boundary. The threat could range from someone walking past using their I-Pod or mobile phone through to a TV or stereo system in the home. Electronic devices can be considered as unintentional transmitters of radio frequency energy, this energy propagates away from the device with unknown directions and amplitudes. In order to ascertain the exact direction at which the maximum amplitude occurs a full spherical scan of the device with a measurement system is required.

The aim of CISPR 12 is to provide protection against such interference by defining a limit below which all vehicles radiated emissions must fall. The current method utilised during a CISPR 12 measurement program does not maximise the emissions through azimuth rotation and antenna height scanning, as is the case with most other Standards used to test typical domestic and commercial appliances. During a CISPR 12 measurement a fixed receive antenna height of 3m is used, with the DUT being measured at just two azimuth angles (normal to the left hand and right hand side of the vehicle, in line with the centre of the engine block).

In an attempt to investigate the errors introduced by using the current CISPR 12 measurement method, a simple representation of a vehicle body shell has been modelled using a package of commercially available electromagnetic simulation software (CONCEPT II.) The model was excited using a single 2m long harness located in the passenger foot well position. The amplitude of the vertical and horizontal component of the electric field has been simulated for five frequencies between 100 MHz and 500 MHz. From this data, polar plots of the radiation pattern were produced. The maximum amplitude recorded over the full 360 degree rotation of the model was then compared to the amplitude recorded at the two CISPR 12 equivalent azimuth angles. Errors of up to 30 dB have been recorded across the frequency band considered.

Future work will investigate methods of reducing the errors described above, this will include the use of alternate azimuth angles and the use of different heights for the receive antenna above the ground.

Choice Satisfaction in Search Engine Use

Pawitra Chiravirakul

Department of Computer Science, University of Bath, UK

p.chiravirakul @bath.ac.uk

Abstract. In typical search engine use, an extensive set of relevant documents is returned. If only one or a few search results can be read, the size of this set may decrease satisfaction with the chosen document(s).

Recent cognitive-psychological studies and consumer behaviour research have demonstrated that more choice sometimes causes a negative effect on consumers' decision satisfaction. Consumers may be more satisfied with their selected choice when they choose from a small set of options – this is the so-called Choice Overload Effect or The Paradox of Choice. The existence of this effect was found in numerous domains. Nevertheless, only a few researchers attempted to identify the sufficient conditions for the effect.

In this work search engine users' behaviour and their satisfaction were investigated. A series of three related experiments were conducted. Participants were given search tasks and presented with either six options or twenty-four returns to choose from. We examined factors that moderate the choice overload effect by measuring the satisfaction with the final choice. The results showed that decision satisfaction was affected by the ranking of returns, and that choice satisfaction was affected by the number of options and the decision time. Large sets of options yielded positive effect on participants' satisfaction when they made a decision without time limit. When time was more strongly constrained, choices from small sets led to relatively higher satisfaction. This research adds to the existing body of research on the choice overload effect in the search engine use context.

Investigating the Impact of Social Interactions in Adaptive E-Learning by Learning Behaviour Analysis

Lei Shi and Alexandra I. Cristea

Department of Computer Science, University of Warwick
{lei.shi, acristea}@dcs.warwick.ac.uk

Abstract. Adaptive Educational Hypermedia Systems (AEHSs) allow for personalization of e-learning[1]. Social media tools enable learners to create, publish and share their study, and facilitate interaction and collaboration[2]. The integration of social media tools into AEHS offers novel opportunities for learner engagement and extended user modelling, and thereby fosters so-called Social Personalized Adaptive E-learning Environments (SPAEEs)[3]. However, there has been a lack of empirical design and evaluation to elaborate methods for SPAEEs. The goal of research, therefore, is to investigate 1) the learning behaviour patterns within SPAEEs and the use of these patterns for learner engagement, 2) the evaluation methodologies for SPAEEs, and 3) the design principles for SPAEEs. Topolor[4] is a SPAEE that has been under iterative development for achieving our research goals. The first prototype was used as an online learning system for MSc level students in the Department of Computer Science, at the University of Warwick, and usage data was anonymously collected for analysis[5]. This poster focuses on system features and learning behaviour analysis. We firstly present the methodologies applied in the research, followed by the social and adaptive features that Topolor provides[6]. Then we revisit the analysis of learning behaviours[7]. Finally we propose the follow-up work based on the evaluation results.

---

[1] Brusilovsky, P.: Adaptive Educational Hypermedia: From generation to generation (Invited talk). In: Proc. of Hellenic Conference on Information and Communication Technologies in Education, Athens, Greece, pp. 19–33 (2004)

[2] Shi, L., Al-Qudah, D. A., Cristea, A. I.: Social E-Learning in Topolor: A Case Study. In: IADIS International Conference e-Learning, pp. 57–64. IADIS Press (2013)

[3] Shi, L.,et al.: Designing Social Personalized Adaptive E-Learning. In:18th Annual Conference on Innovation and Technology in Computer Science Education, pp. 341–341. ACM (2013)

[4] Shi, L., Al-Qudah, D. A., Qaffas, A., Cristea, A. I.: Topolor: A Social Personalized Adaptive E-Learning System, In: 21st Conference on User Modeling, Adaptation and Personalization. LNCS, vol. 7899, pp. 338–340. Springer, Heidelberg (2013)

[5] Shi, L., et al.: Evaluation of Social Interaction Features in Topolor - A Social Personalized Adaptive E-Learning System. In: 13th IEEE International Conference on Advanced Learning Technologies, pp. 294–295. IEEE Computer Society (2013)

[6] Shi, L., et al.: Social Personalized Adaptive E-Learning Environment - Topolor: Implementation and Evaluation. In: 16th International Conference on Artificial Intelligence in Education. LNCS, vol. 7926, pp. 708–711. Springer Berlin Heidelberg (2013)

[7] Shi, L., Cristea, A. I., Awan, M., Stewart, C., Hendrix, M.: Towards Understanding Learning Behavior Patterns in Social Adaptive Personalized E-Learning Systems, In: 19th Americas Conference on Information Systems, pp. 1–10. Association for Information Systems (2013)

Scalable Parallelism in Game Simulations

Joseph Kehoe and Joseph Morris

Dublin City University

Abstract. Most parallelism in games occurs during the graphics stage of the game simulation pipeline. This takes the form of Single Instruction, Multiple Data (SIMD) concurrency and executes on the graphics card. As yet parallelism in the rest of the game pipeline has proven difficult to implement[1]. We will show that synchronous parallelism can be adapted to game simulations, that it scales well, removes deadlock and is deterministic.

Multicore architectures are becoming the de facto standard for computer game platforms. In computer games the main program loop consists of three parts: physics simulation, graphics display and game entity simulation (AI). The graphics and physics stages are amenable to SIMD parallelism and as a result can be executed on specialized graphics cards designed for SIMD parallelism. The simulation code, on the other hand, is not amenable to SIMD parallelism and so must be run on the main processor. Over the past decade, processor advances have been achieved by switching from single core architectures to multicore architectures. Writing code that can efficiently utilize shared-memory multicore architectures has proven difficult thus game simulation code has been unable to take full advantage of these architectures. We propose a model of concurrency that combines the Bulk Synchronous Processing and Active Object models and demonstrate that it is deadlock free and deterministic[2].

We intend to implement a prototype of this model to demonstrate that this model is suitable for games simulation. We will implement a number of standard algorithms used in games such as: flocking, swarming and path finding to test the model. As well as these we will also develop a game test bed that will be used to simulate various loads and game scenarios. These benchmarks will be used to test the efficiency of our approach. Efficiency will be measured in two parts. First, we will show that any overhead introduced by the model must not cancel out any gains from the introduction of concurrency. We will also determine the maximum frame rate (fps) that is possible before the overhead becomes too burdensome. Secondly, we will measure the scalability of the solution by varying the size of the simulations and recording the maximum possible fps rate. To be successful we require the solution to scale linearly on shared memory multicores (from two to thirty two cores) and to degrade gracefully as the size of the simulation grows.

[1] Tim Sweeney. The Next Mainstream Programming Language: A Game Developer s Perspective. SIG-PLAN Not., 41(1):269269, January 2006.

[2] Morris, J., & Kehoe, J. (2011). A Concurrency Model for Game Scripting. 12th International Conference on Intelligent Games and Simulation (pp. 10-16). EUROSIS.

Towards reflective search of archival interfaces

Jo Pugh

Department of Computer Science, University of York, United Kingdom
jjp513@york.ac.uk

Abstract. Archival collections present particular problems for both users and systems developers. These collections are generally comprised of many different groups of records, each with a distinct form of arrangement and consist of a range of different media (electronic, paper, microfilm). This array of collections has been called a "crazy quilt" and users' interactions with it are characterised by frustration and confusion.[1] For systems developers, an additional problem is the variable quality of archival metadata.

Dozens of models of the information seeking behaviours of users exist, ranging from simple query-response system focused models to complex examinations of environment and behaviour.[2] It has been suggested that we need to move beyond this query response-paradigm.[3] Yet when we consider the "advanced" components for search interfaces suggested by authors such as Shneiderman[4], these still seem primarily occupied with query building or new ways of presenting results.

Beyond the phases of query execution and result examination, users need support in applying the most useful and relevant search strategies and tactics to enable them to develop their information seeking expertise.[5] Developers need to incorporate new components into archival systems which support search regulation, information processing and the organisation and presentation of research outcomes[6]. The result will be sytems which support reflective information seeking, in which searches are well planned and results well considered.[7]

---

[1] Duff, Wendy, Stoyanka, Petra: Transforming the Crazy Quilt: Archival Displays from a Users Point of View. Archivaria 45, p.60 (1998)

[2] Wilson, Tom D. Models in information behaviour research. Journal of documentation 55, no. 3 (1999): 249-270.

[3] White, Ryen W., Roth, Resa. "Exploratory search: Beyond the query-response paradigm." Synthesis Lectures on Information Concepts, Retrieval, and Services 1, no. 1 (2009): 1-98.

[4] Shneiderman, Ben, Plaisant, Catherine. Designing the user interface, 5th edition, p.506 (2009).

[5] Bates, Marcia J. Information search tactics. Journal of the American Society for information Science 30, no. 4 (1979): 205-214.

[6] Brand-Gruwel, Saskia, Iwan Wopereis, Amber Walraven. A descriptive model of information problem solving while using internet. Computers & Education 53, no. 4 (2009): 1207-1217.

[7] Edwards, Syvia, Bruce, Christine: Reflective internet searching: an action research model. The Learning Organisation 9, no.3, p.183 (2002)

Dynamic Pick-up and Delivery Problems and Pricing

Philip Mourdjis[1], Prof. Peter Cowling[2], and Martin Robinson[3]

[1] YCCSA, University of York, York, YO10 5GE, United Kingdom
pjm515@york.ac.uk
[2] peter.cowling@york.ac.uk
[3] Transfaction Ltd. B1 Adastral Park, Martlesham Heath, Suffolk IP5 3RE, United Kingdom
martin.robinson@transfaction.com

Abstract. The Dynamic Pick-up and Delivery Problem (DPDP) is a collection of extensions to the vehicle routing problem. Firstly, multiple depots and vehicles are considered each having its own working hour rules. Secondly, orders have collection and delivery locations such that a vehicle must make a collection before the associated delivery. Thirdly, each location has a time window during which it should be serviced with penalties associated with late service. Finally, the problem is dynamic, meaning that orders cannot be entirely pre-planned and must be updated "on the fly" in real time.

As a set of extensions to an existing NP-Hard problem, the DPDP is inherently complex, when thousands of orders are involved it becomes infeasible to solve using exact methods. Metaheuristics have shown promise on many NP-Hard problems and can be used to produce high quality solutions quickly[4]. To develop a metaheuristic for Transfaction Ltd. we have carried out a detailed analysis of existing order information and created a method to generate randomised orders matching the known distribution of the data.

We present an abstract model of Transfaction's problem, developed to study and optimise vehicle schedules in the DPDP environment, and an overview of existing metaheuristics for similar problems such as Tabu Search[5], Variable Neighbourhood Search[6] and Monte-Carlo Tree Search with discussion as to how these are suitable for our problem.

To investigate pricing of orders in such a system we plan to develop an agent based system where carrier companies can bid on new orders as they arrive. A number of pricing and bidding strategies will be considered and interactions between agents employing different strategies will be noted.

[4] El-Ghazali Talbi, "Metaheuristics: From Design to Implementation", John Wiley & Sons, 2009.
[5] Jean-Franois Cordeau, "A Parallel Iterated Tabu Search Heuristic for Vehicle Routing Problems", Computers & Operations Research, 2012
[6] Pierre Hansen, "Variable Neighbourhood Search: Methods and Applications". 4or, 2008.

Intelligent Scheduling and Routing
gully maintenance problem

Yujie Chen[1], Prof. Peter Cowling[2], and Stephen Remde[3]

[1] YCCSA, University of York, York, YO10 5GE, United Kingdom
yc1005@york.ac.uk
[2] peter.cowling@york.ac.uk
[3] stephen.remde@york.ac.uk

Abstract. The last decade has seen an increasing utilization of optimization techniques for the effective management and provision of services in distribution systems. Large numbers of real-world applications have already shown that the use of computerized planning techniques for the global distribution process can produce substantial savings, generally in the range of 5 to 20%[4]. It is easy to see that the impact of these savings on the global economic system is significant.

Our research looks at a large scale asset maintenance issue using management strategies. In collaboration with our industrial partner Gaist Solution Ltd, we focus on a gully maintenance problem (GMP) consisting of 28,290 gullies distributed over Blackpool. Gullies are an integral component of highways responsible for draining surface water, maintaining these in a timely manner is important for the city. The complexity behind the GMP is that each gully has a number of features affecting our decision making including geographic position, social value and a condition that changes over time. Combining ideas from the period routing problem (PVRP)[5] and the routing problem with profit (VRPP)[6], this paper presents a mathematical model for the GMP and proposes a 2-phase dynamic scheduling strategy to solve it. Our aim is to construct a set of tours for each vehicle for each day over a long period (1 year) that maximizes the average maintenance quality while satisfying operational constraints.

[4] Toth P. , & Vigo D. (Eds.) (2002), "The vehicle Routing Problem", Society for Industrial and Applied Mathematics
[5] Christofides, N., & Beasley, J. E. (1984). The period routing problem. Networks
[6] Archetti, C., Speranza, M., & Vigo, D. (2012). "Vehicle Routing Problems with Profits"

Knowledge based search and decision making for Magic: The Gathering

Nick Sephton

Computer Science,
University of York, York, YO10 5DD

Abstract. *Magic: The Gathering* is a physical trading card game created by Richard Garfield and first published by Wizards of the Coast in 1993. It is a complex strategy game in which players take turns to play cards onto the table and attack each other in an attempt to deplete their opponent's life totals (and thus win). Magic is of interest as it has a very high branching factor and tree depth, caused by a highly technical game rule structure, which is hidden during normal play by human ability to abstract away complex information. In an initial turn of Magic, technically there are 14 "priority passes" (opportunities for a player to act), which are automatically and silently ignored by human players until relevant, but each must be considered by any AI attempting to play optimally. A "Learn to Play" demo is available on the Wizards of the Coast website[1].

Magic: The Gathering offers an interesting research target for a number of reasons. Firstly the game is incredibly complex and has a high branching factor, making it a very challenging target. Secondly, the game is highly successful and has many different levels of play, including professional players.

For the purposes of experimentation, a vastly simplified version of Magic: The Gathering was created, which doesn't consider many of the options and priority passes that are included in a standard game of Magic. The simplified version only included creatures and land (just two of the seven different types of cards available in Magic).

[1] http://www.wizards.com/magic/tcg/newtomagic.aspx?x=mtg/tcg/newtomagic/learntoplay

Reducing Number-Entry Errors in Medical Systems: A Tale of Two Studies

Frank Soboczenski, Aniket Datta, Paul Cairns

Department of Computer Science, The University of York, Deramore Lane YO10 5GH, UK
{fs596,ad835,Paul.Cairns}@york.ac.uk

Abstract. Number entry is a ubiquitous task in healthcare. Despite potentially fatal consequences of errors, very little is known about the errors people make using these devices and how the devices themselves influence the errors made[1].

Focusing only on infusion pumps, some of these devices are fitted with serial interfaces, others with chevron or 5-Key interfaces. Additionally, 5-Key and chevron interfaces are generally very common in healthcare and other safety-critical domains. This makes it important to understand the errors that people make with such devices and how these errors can be mitigated. We present two studies which look into the errors made in number entry using a 5-Key and a chevron interface.

In the first study, we chose to investigate specific design features for 5-Key devices and their effect on error reduction as proposed in earlier work by Cauchi et al[2]. Specifically, [2] found that the cursor start position consistently affected the errors made with the interface. Therefore, we address the design feature in our 5-Key study of where the cursor starts to verify its effect on the errors by an empirical evaluation. In the second experiment we chose to investigate under what conditions people tend to make errors during a number-entry task. For this purpose we studied participants in three situations when entering numbers in a chevron interface, namely *Relaxed*, *Accuracy* and *Pressure*.

What we discovered is a persistent error rate of 1-3% in both studies but we could not detect a difference between a left or right cursor start design in the 5-Key study. However, our results show that there is a promising tendency towards a significant difference in the amount of errors that people make between the three conditions in the chevron study. To our surprise, the latter also revealed that number-entry in the *Pressure* condition was not only the least accurate of the three but it also took the longest time for participants to enter the numbers. Moreover, 69% of the people participating in the 5-Key and 52% participating in the chevron experiment made at least one error while entering numbers.

[1] H. Thimbleby and P. Cairns. Reducing number entry errors solving a widespread, serious problem. Journal of the Royal Society, Interface / the Royal Society, 7(51):142939, Oct. 2010.

[2] A. Cauchi, A. Gimblett, H. Thimbleby, P. Curzon, and P. Masci. Safer 5-key number entry user interfaces using differential formal analysis. Proceedings of the BCS HCI 2012, 12:12, 2012.

Educare: Democratic Framework for Educational Data Analysis

Artur Goulão Ferreira[1], Cláudia Antunes[1]

[1] Instituto Superior Técnico – Universidade de Lisboa, Av. Rovisco Pais, 1 – 1049-001 Lisboa
{artur.ferreira,claudia.antunes}@ist.utl.pt

Abstract. With the democratization of education in societies and the rise of technology there is a need for better tools to understand the behaviours of students and educators. Educational Data Mining (EDM)[1] as an emerging discipline applies data mining techniques to data from educational settings. These processes usually require specific data mining tools in a well-defined context and their capability to adapt to new data or environment changes are, for most of them, nonexistent or very expensive/out of context for governments.

The present work is part of the *educare* project and aims for exploring the possibility of creating an Online Analytical Processing (OLAP)-based framework for data analysis using common technologies and methodologies, and proposes a new architecture for a framework that can answer four questions, can it be developed at a low cost or even for free? Can it be permanently accessible but be fully functional without internet access? How would it interact with future data? Are any of those components reusable in other scenarios, either technological (cheap mobile devices) or social?

Both cost and interoperability were given a significant role in the study. Many OLAP based browsers are either platform specific or don't expose an API and therefore cannot be integrated. The present study represents a significant step forward into Educational Data Mining providing power tools based on open source technologies.

The next step in this work is to explore the impact of external actor's contexts (other than the educational environment) on actor's results. Our first step will be the integration of information from student social networks, in order to analyze and visualize how the student's educational behavior is influenced by their social context. After being able to model this type of information, and successfully integrating it on results prediction, other contexts may be explored.

[1]The research project educare is being developed at Instituto Superior Técnico since 2011, and is supported by Fundação para a Ciência e Tecnologia (PTDC/EIA-EIA/110058/2009).