Proceedings of the

Eighth York Doctoral Symposium on Computer Science & Electronics

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

28th October 2015

Editor: Colin Paterson

Sponsoring Institutions

The Organising Committee are extremely grateful to the following institutions for generously sponsoring YDS 2015. Without such sponsorship YDS would not be able to provide doctoral students of Computer Science and Electronics the opportunity to experience such a well rounded, fully fledged academic conference in their field as a training experience.

Academic Sponsorship

THE UNIVERSITY *of York*

The Organising Committee would like to thank both the Departments of Computer Science and Electronics for their sponsorship.

Industrial Sponsorship

The Organising Committee would like to thank the following organisations for their sponsorship.

IBM

RAPITA SYSTEMS LTD

DRIVING EMBEDDED EXCELLENCE ETAS

SimOmics

THALES

Preface

The York Doctoral Symposium on Computer Science and Electronics (YDS) is an international post-graduate student symposium, now in its eighth year. YDS 2015 was again organised jointly between the departments of Computer Science and Electronics. Here we built upon the foundations laid by last year's Organising Committee and I hope that this partnership continues next year. 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 also hopes to promote more interdisciplinary research.

YDS 2015 offered three categories of submissions, full-length papers, extended abstracts and posters. We received 1 extended abstract, 8 full-length papers, and 13 posters. We saw a slight decrease in the number of submissions when compared to last year. Amongst the material received we were pleased to accept 2 papers and 6 posters from external universities. The acceptance rate for full-length papers and extended abstracts combined was 70%. 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 organisers as well as for the contributors and attendees. Therefore, both full-length papers and extended abstract received three anonymous reviews each.

We are very grateful to our sponsors without whom YDS could not take place. YDS 2014 was sponsored by the Departments of Electronics and Computer Science at the University of York; our industrial sponsors: IBM, Rapita Systems, ETAS, SimOmics and Thales. Among other things, their financial help enabled us to offer prizes for the best full-length paper and extended abstract, the best presentation and the best poster voted for by attendees. We were honoured to host invited keynote talks by Prof. Kevin Warwick, Deputy Vice Chancellor of Coventry UNiversity, Dr. Gordon Hollingworth, from Raspberry Pi and Dr. James Ravenscroft, from IBM.

I would like to express my gratitude to the YDS 2015 Organising Committee for their work on the design and logistics of the event, particularly Nakita Johnson, the Organising Committee Chair, and Patricia Ryser-Welsh for their invaluable work; to the members of the Programme Committee for their time and concentration in producing such useful and professional reviews; to the academic staff at the University of York for their promotion and support of YDS, especially Dr Mike Dodds, Prof. John Clark, and Prof. Richard Paige; to the Computer

Science administration staff who were always helpful and cheerful, no matter what we asked for. I would like to extend my personal thanks to Matt Luckcuck, Sam Simpson and Simos Gerasimou who, as previous chairs of the conference, were able to support me through the process..

YDS 2015 has been exciting, exhausting and stressful in equal measure. However, I believe that the chance to chair the Programme Committee of an academic conference is a worthwhile experience, one that I am glad that I did not pass up. YDS has afforded me a valuable insight into the academic conference process and to forge friendships along the way. I would encourage all post-graduate students to take part in YDS; whether that be in organising, reviewing for, contributing to, or simply attending YDS. Finally, I would like to wish the Organising Committee for YDS 2016 all the luck and determination that they will undoubtedly require.

Colin Paterson
YDS 2015 Programme Chair

Organisation

All members of the YDS 2014 Organising and Programme Committees are from the University of York, UK.

Organising Committee

Chairs

Nakita Johnson

Committee

James Baxter
Matthew Bedder
Mark Douthwaite
Nakita Johnson
Stephen Oxnard
Colin Paterson
Patricia Ryser-Welch
Mark Thomas

Programme Committee

Chairs

Colin Paterson
Patricia Ryser-Welch

Committee

Stewart Aitken
Konstantinos Barmpis
James Baxter
Matthew Dale
Alena Denisova
Mark Douthwaite
Tom Fleming
Dr. Alan Frisch
Simos Gerasimou
Dr.David Haliday
Prof. Edwin Hancock
Dr. Atsifumi Harotima
Prof. Tim Kelly
Matt Luckcuck
Prof. John McDermid
Nakita Johnson
Russell Joyce
Colin Paterson
Jo Pugh
Patricia Ryser-Welch
Gabriel Costa Silva
Sam Simpson
James Stovold
Prof. Jon Timmis
Dr. Martin Trefzer
Matt Windsor

Contents

I Keynote Talks

II Full-Length Papers

III Extended Abstracts

IV Poster Presentation Abstracts

Part I

Keynote Talks

Cyborgs, Robots with Brains and the Turing test

Prof. Kevin Warwick

Deputy Vice-Chancellor (Research) at Coventry University

Abstract. In this presentation a practical look is taken at how the use of implant and electrode technology can be employed to create biological brains for robots, to enable human enhancement by creating Cyborgs and to diminish the effects of certain neural illnesses. In all cases the end result is to increase the range of abilities of the recipients. An indication is given of a number of areas in which such technology has already had a profound effect, a key element being the need for a clear interface linking a biological brain directly with computer technology.
The emphasis is placed on experimental scientific studies that have been and are being undertaken and reported on. The area of focus is notably the need for a biological/technological connection, where a link is made directly with the cerebral cortex and/or nervous system. The presentation will consider the future in which robots have biological, or part-biological, brains and in which neural implants link the human nervous system bi-directionally with technology and the internet.
In the presentation we also take a look at the practical instantiation of the Turing test in which machines pretend that they are human. Actual transcripts will be presentation and the audience will have a chance to try for themselves and see if they can tell the difference between humans and machines purely from conversations.

Biography Kevin Warwick is Deputy Vice-Chancellor (Research) at Coventry University. Prior to that he was Professor of Cybernetics at The University of Reading, England. His research areas are artificial intelligence, control, robotics and biomedical engineering. He is a Chartered Engineer (CEng.) and is a Fellow of The Institution of Engineering & Technology (FIET). He is the youngest person ever to become a Fellow of the City & Guilds of London Institute (FCGI). He is the author or co-author of more than 600 research papers and has written or edited 27 books (three for general readership), as well as numerous magazine and newspaper articles on scientific and general subjects. He has broadcast and lectured widely and holds various visiting professorships.

Development of the Raspberry Pi

Dr. Gordon Hollingworth

Raspberry Pi

Biography Gordon Hollingworth is director of software engineering at Raspberry Pi. During his time there he has overseen the production of a number of Raspberry Pi projects such as the recently released display, the official Raspberry Pi case and software projects like the Raspbian image and the NOOBS installation system. Before this he worked in the mobile multimedia division at Broadcom as a manager of the platform software group and before that has worked in IPTV and consultancy companies.

Educated at the University of York, Gordon gained a first class MEng in Electronics before undertaking a PhD concerning the use of evolutionary strategies to develop electronic circuits with self modifying and fault tolerant behaviour.

Natural Language Processing Techniques and Scientific Discovery

Dr. James Ravenscroft

IBM Solutions Architect for Watson in Europe

Biography James currently works for IBM as the Solutions Architect for Watson in Europe; a technical role he carries out part time whilst also studying as a PhD student at Warwick University under the supervision of Dr Maria Liakata. He has been working for IBM for two years after graduating from Aberystwyth University in 2013 with a first class honours degree in AI and Robotics. He also won the BCS prize for best final year undergraduate project for his work on Partridge, a software pipeline for automating information extraction from scientific literature.

In his talk, James will discuss the history of scientific innovation and the growing importance of natural language processing techniques in keeping up with the ever-accelerating progress of scientific discovery. He will introduce Partridge and his paper processing pipeline and talk on how IBM Watson is also being used for scientific discovery.

Part II

Full-Length Papers

Adapting the Cloud-COVER Threat Modelling Tool to Cover Threats for the Internet of Things

Mustafa Aydin and Jeremy Jacob

University of York

Abstract. We propose an adapted design to the threat modelling tool Cloud-COVER, encompassing the threat landscape for the Internet of Things. By changing the model within the tool, in particular the assumptions made about the devices and their security features, users can be presented with a list of the most likely threats they face according to priority. In addition, other changes, such as the way that threats are ranked within the results, are presented. This adaptated design is the first attempt to provide a threat modelling framework for the IoT environment.

1 Introduction

Threats to computer security are a massive problem, with these problems exacerbated by the connections allowing them to be attacked from anywhere in the world. Increasing numbers of ordinary household items are starting to contain connected embedded devices — the so called 'Internet of Things', presenting additional security and privacy worries [1]. Like all computers these embedded devices also have vulnerabilities, and the security dangers they provide need to be paid attention. Cars, fridges, medical devices, and phones may not be the most obvious targets for hackers, but they are likely to contain valuable data or allow the control of dangerous or expensive devices. Research has demonstrated the feasibility of attacks on each type of device, and the required need to protect the security and privacy of each [2] [3] [4] [5].

One of the most important ways of understanding the threats faced by computers is to use threat modelling [6]. Threat modelling provides an understanding of the existing threats according to the way the system under observation is configured and connected, and can help to reduce some of the complexity of threat analysis. Threat modelling for the Internet of Things is one way in which users could better understand and protect themselves against any possible threats against their devices, and take countermeasures to reduce the likelihood of those threats.

Our threat modelling tool Cloud-COVER (Controls and Orderings for Vulnerabilities and ExposuRes) was developed to look at threats for cloud computing. Cloud-COVER works by using information provided by the user about their cloud deployment to model their system and analyse the way in which threats propagate between the cloud instances [7]. By changing some of the assumptions we make about the things represented within the model itself, Cloud-COVER can be adapted to consider threats within the Internet of Things. In this paper, we present the adapted design to Cloud-COVER to provide threat modelling for the Internet of Things.

Our paper is structured in the following way: we describe Cloud-COVER and its model in Section 2. We provide a stage by stage comparison of the most important changes in Section 3. Related work is covered in Section 4, followed by a conclusion in Section 5.

2 Cloud-COVER

The threat modelling tool Cloud-COVER was developed due to the lack of threat modelling tools catering for the specific issues present in cloud computing, one of which is threat propagation. As distributed systems may have any number of weaknesses, attackers may choose to break into less well protected instances which have the privileges necessary for entry to much better protected instances. As Cloud-COVER is the only threat modelling tool which takes threat propagation into account, it is the only tool which can be adapted to apply the same scenario to the Internet of Things.

Cloud-COVER's work can be broken down into four main stages. The first stage involves taking an input from the user about their deployment — the instances, any hosted data, connections, and their properties. Given these details, the presence of specific types of threats can be determined, followed by three stages of analysis in which the threats are then ranked. The second stage asks the user to specify their preferences for which security attributes they consider to be the most important, relative to each other. The third stage is of threat value propagation, in which the user valuations are re-evaluated, based on whether threats could use lower prioritised instances to attack higher priority ones, depending on the connections between them as well as other factors. The resulting preference graph is then used to initially order the threats. In the last stage, the threats are then reordered based on the likelihood of those threats occurring, values which are determined based on statistics derived from the security literature. It is the changes to these stages which will be highlighted in this work, with a summary of the adaptations in Table 1.

2.1 A Problem with The Internet of Things

Like cloud computing, the Internet of Things is also a distributed system of sorts. Although there are differences which need to be taken into account, it makes sense that a threat modelling tool which was developed specifically to cater for the properties of cloud computing could also provide the same service for the Internet of Things. Due to the differences between regular computers and embedded devices, Cloud-COVER needs to be adapted to better represent the threat landscape of the IoT. These differences will include the assumptions being made about the devices and what they can do, and the users of the tool.

Property	Cloud-COVER	IoT Adaptation
Asks for inputs of:	Instances, data and connections	Devices and connections
User valuations by:	Importance of data security attributes	Importance of device or data
Threat propagation:	Depending on relevance to attributes and connection properties	Depending on user valuation of attributes
Statistics for threat ranking:	Compiled from security literature	None available
Listed results:	Specific threats and countermeasures	General threats and countermeasures

Table 1. Main differences between Cloud-COVER and its adaptation for IoT devices.

We propose the kinds of changes needed to be made to our own threat modelling tool, Cloud-COVER, in order to demonstrate how this can be done. We discuss the important stages of analysis in Cloud-COVER, and how these are adapted in order to make this possible.

3 Adaptations

In this section, we explore the major differences between Cloud-COVER and the changes with the adapted design for IoT. Table 1 summarises the main stages in the way Cloud-COVER finds the important threats and ranks them, and lists the main differences in the adaptation in the right hand side column. These adaptations are explained in more detail below.

3.1 First Stage

In Cloud-COVER, users were asked about the different kinds of data hosted on their system. IoT devices are generally not used to host many different kinds of data like regular computers are, so monitoring the various threats to many data items is unnecessary. Instead, the only kinds of things we need to understand are the kind of data stored on the device. This includes personal identification information, health information, or even the credentials for other devices. Even seemingly useless information such as device use statistics is important to know about, as it can provide knowledge of times of home occupation, a clear security threat. Requiring more knowledge than this about data requires more complex considerations more suited to threat modelling tools aimed at regular systems.

One of the most important assumptions we make about most computers is that they offer basic security precautions to users, such as password protections for access. With embedded systems, even this cannot be assumed to be the case. A computer which has been secured with a password is not invulnerable to hacking, but at the very least presents some obstacles to hackers. One important example of Cloud-COVER inputs which demonstrate such an assumption was when asking about the properties of each connection, specifically the rights given to connections to perform operations (such as data creation and deletion) at each end. With embedded devices, mechanisms to control such operations cannot be assumed to exist, so instead only the existence of connections between devices is important to know about in order to analyse the threat propagations (discussed in Section 3.3).

Instead of asking users about the properties of their data and instances as is done in Cloud-COVER, the properties of the devices themselves is a better way to understand the threats against the devices. These properties include what kind of information is recorded on the device (such as whether or not this is personal information), or whether they allow any level of control over the device they are linked to. One important question to be asked is if the device allows any security access controls, as this is one of the primary ways of protecting the user's devices. After getting this information, threats to each of the devices can be ascertained and in subsequent stages, ranked according to priority.

3.2 Second Stage

The rankings in Cloud-COVER are determined by allowing users to value the most important security attributes of each piece of data. The attributes used were of confidentiality, availability, integrity, and compliance. For each data item, users were asked to rate the importance of those four attributes relative to each other. Using these valuations, only those threats relevant to those attributes could be prioritised, providing the user with an understanding of where to concentrate their efforts to secure their system.

Within the adaptation, users will need to specify preferences but between different kinds of attributes. In the cloud, blind spots in security may occur due to the need to balance data security with accessibility. However in the IoT, the lack of security mechanisms means that there is little balancing required, so using multiple security attributes to consider security implications makes little sense. Instead a simpler representation of threat categorisation is needed.

Using more general security attributes, and using these to analyse security threats, makes more sense when considering devices which are likely to have minimal security capabilities. Instead of dividing threats to data into different categories, threats to data are all considered under a single heading. The other major threat to IoT devices is loss of device control, when attackers hack into devices in order to control some of their functions. Even control of everyday objects can lead to serious outcomes, with one obvious example being control of a garage door, a clear security vulnerability for a household. There have even been examples of hackers taking control of cars. Control of devices is therefore also an important attribute to be rated by the user. The changed nature of the analysed attributes also allows a simpler understanding of the threats in the IoT, and especially in a way that non security experts can understand.

An example of user valuations is given in Figure 1, in which the user provides their valuations for a system containing 3 devices (A, B and C). Relative valuations are used, so users must specify whether attributes are greater than, or much greater than, other attributes. Using this input, threat propagation analysis then takes place.

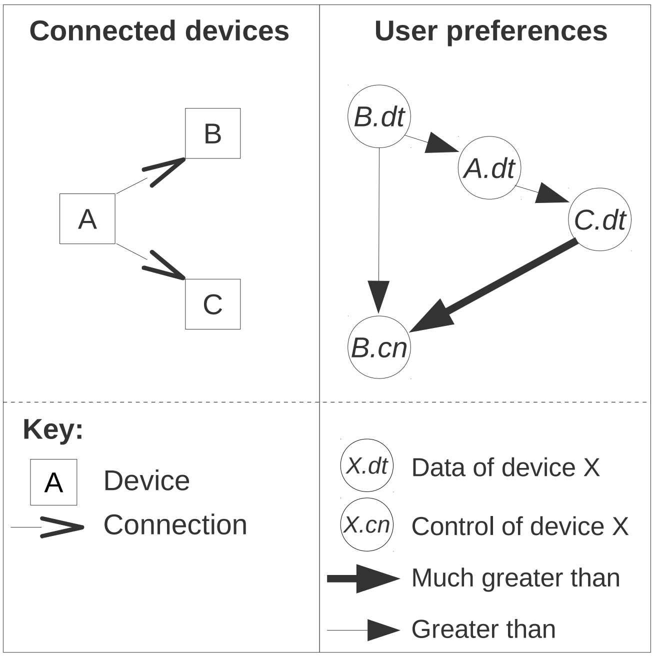


Fig. 1. Examples of connected devices on the left, and user valuations on the right.

3.3 Third Stage

Propagations occur between attributes when they are on connected devices, and when the attribute on the connected to device is higher than that of the attribute on the connecting device. When this occurs, the value of the higher rated attribute is passed over. Propagation examples are shown in Figure 2, in which C's values propagate across to A's attribute values. Although B is also connected to A, B's values could never propagate over to A, as they have been given a lower value than A by the user. C's values can also not propagate over to B, since C and B are not connected (as seen in Figure 1).

When values propagate, they take the incoming and outgoing edges of the attribute that has propagated its value over to them. In Figure 2, *A.cn* takes in the two outgoing edges of *C.cn*, so it becomes greater than *C.dt* and much greater than *B.cn*. *A.dt* also takes the same values from *C.dt*, with no values propagated over to any of B's attributes.

This value propagation helps to ensure that it will be harder for attackers to use less well protected devices to attack better protected devices, since the

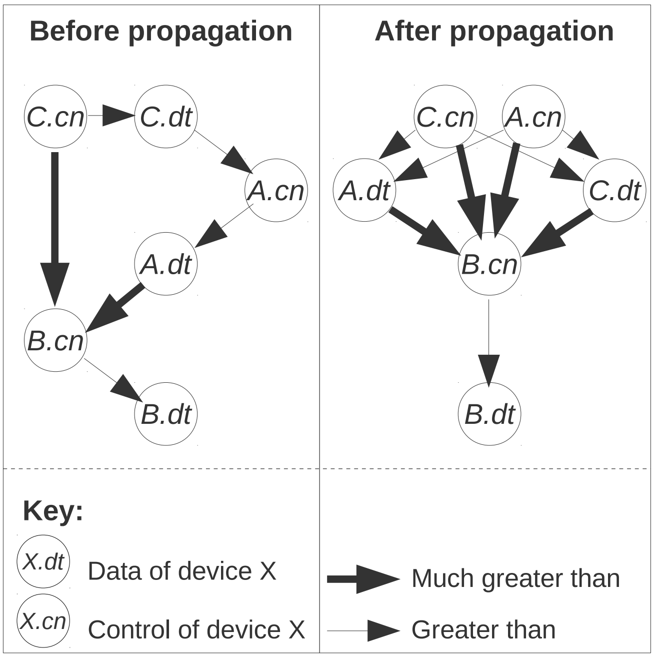


Fig. 2. A before and after example of how values propagate.

priority attached to the threats on each device will have been equalised by the propagation analysis. In Figure 2, before the propagation analysis took place, device C could be more easily attacked through device A since device A's attributes were given lower priorities for securing. After the valuation propagation, A's attributes were given ratings as high as those from C. Although in Cloud-COVER other factors, such as the rights given to the connections, also play a part in the propagation analysis, they are not needed when looking at IoT, because they rely on the assumption that security mechanisms are present in all instances.

3.4 Fourth Stage

After taking account of the user's preferences for security attributes and providing an initial ordering, Cloud-COVER reorders the results according to their probability of occurrence. These probabilities were determined according to information found in literature regarding security breaches. For the Internet of Things, these kinds of statistics do not exist so the rankings will have to be constructed in a different way.

One possibility is to use the growing amount of anecdotal evidence which is starting to appear. By noting the frequency with which different kinds of attacks are being used in the real world (and noting the ways in which researchers demonstrate hacking), it should be possible to at least offer some indications of the likelihood of various attack methods. Although imperfect, the importance of probability of threats is one of main stages of determining the final priority of threat rankings in Cloud-COVER, and should be utilised even if the statistics used may have some issues. Overcoming those issues however is a serious question which will need to be addressed.

3.5 Fifth Stage

In Cloud-COVER, specific threats and countermeasures relevant to the deployment were able to be presented to the user. The fact that different kinds of computers all share broadly similar characteristics means that the differences can be abstracted away in the model. Despite the abstraction, important conclusions about threats can still be determined and presented to the tool user.

The problem with IoT devices is that they cover such a broad range of capabilities. Some may have minimal capabilities, whilst others may even be capable of running some computing applications. This makes the range of possible threats and countermeasures huge, so for the Internet of Things a much more generalised approach to classifying threats and countermeasures is required.

Threat modelling can help to reduce some of the dangers to IoT devices by providing advice to users about where threats lie, but simply providing general advice is not enough. Although the advice cannot be too specific, by determining as much as possible by asking users about the properties of their devices, it is still be possible to determine important threats to their devices. The countermeasures to these threats will include advice such as using ensuring the use of possible security access controls, reducing the number of devices connecting to each other, or to reduce the impact of hacked devices by changing the nature of information stored on them.

The way in which the ordered threats are presented to the user is presented in Table 2. Although *A.cn* and *C.cn* are rated as the most important attributes by the user, threats to *A.dt* and *C.dt* are presented as being of equal value. But as all of B's attributes are rated as much less important (as seen in Figure 2), none of the threats to B mix with those of A and B.

4 Related Work

There are several threat modelling tools and risk management processes used to help with the management of threats.

The STRIDE and DREAD processes are manual threat modelling approaches through which threats are determined by considering the possibility of threats using various techniques to attack systems. STRIDE considers threats taking various approaches (Spoofing, Tampering, Repudiation, Information disclosure, Denial of service, and Elevation of privilege) which form the initials of its name [8][9]. The problem with manual approaches are the time taken to work with them, especially when performing multiple iterations after changes to any work

Order	Data	Attribute	Threat	Probability
1	A	Control	Control of security device	High
=	C	Control	Control of dangerous device	High
3	A	Data	Personal information theft	High
=	A	Control	Resource consumption	Low
=	C	Data	Health information theft	High
=	C	Control	Resource consumption	Low
...	...	...	...	...
11	B	Data	Control of security device	High
12	B	Data	Device use statistics data theft	Low

Table 2. The presentation of threats to the user. Although their top attributes have the same value, the threats present on both devices are different. The properties specified by the user determine the relevant threats to those devices, which are then ordered according to the attributes preferences.

require reevaluation. The most famous example of an automated tool using these techniques is Microsoft's SDL Threat Modelling Tool, which uses the STRIDE methodology to suggest threats to the system under consideration [10].

Attack trees work by considering attacks as tree structures, with goals as the root and the different ways of reaching the root as branches. By attaching different values to the branches reaching the goals, each path can be evaluated to consider how to prioritise defence within systems [11]. They can help to consider any situation when given values for each branch for the situation under consideration. Sea Monster is an automated example of a tool used for attack trees [12].

5 Conclusion

We have discussed the design of an adaptation to our threat modelling tool Cloud-COVER, with a view to creating a threat modelling tool for the Internet of Things. The adaptations keep many of the stages used in the original tool, but changed in order to better relate to threats in the IoT environment.

The next important step will be to make the required changes to Cloud-COVER, with a working tool to provide rankings of threats for users. This could then be tested with everyday users of connected devices, to see how they react to the suggestions made by the tool.

Acknowledgment

Mustafa Aydin is an EngD candidate in the LSCITS project, and is funded by an EPSRC stipend.

References

1. S. R. Peppet, "Regulating the internet of things: First steps toward managing discrimination, privacy, security and consent," *Tex. L. Rev.*, vol. 93, p. 85, 2014.
2. A. S. Elmaghraby and M. M. Losavio, "Cyber security challenges in smart cities: Safety, security and privacy," *Journal of advanced research*, vol. 5, no. 4, pp. 491–497, 2014.
3. A. Grau, "Can you trust your fridge?" *Spectrum, IEEE*, vol. 52, no. 3, pp. 50–56, March 2015.
4. S. Gollakota, H. Hassanieh, B. Ransford, D. Katabi, and K. Fu, "They can hear your heartbeats: non-invasive security for implantable medical devices," *ACM SIGCOMM Computer Communication Review*, vol. 41, no. 4, pp. 2–13, 2011.
5. A. Loo, "Technical opinion security threats of smart phones and bluetooth," *Communications of the ACM*, vol. 52, no. 3, pp. 150–152, 2009.
6. S. Myagmar, A. J. Lee, and W. Yurcik, "Threat modeling as a basis for security requirements," in *Symposium on requirements engineering for information security (SREIS)*, vol. 2005, 2005, pp. 1–8.
7. M. Aydin and J. Jacob, "Cloud-COVER: Using User Supplied Valuations and Propagation Analysis to Prioritise Threats to Systems," To appear in European Intelligence and Security Informatics Conference (EISIC), 2015.
8. Microsoft, "The STRIDE Threat Model," [Online] Available: https://msdn.microsoft.com/en-us/library/ee823878(v=cs.20).aspx , Accessed June 2015, 2002.
9. J. Meier, A. Mackman, M. Dunner, S. Vasireddy, R. Escamilla, and A. Murukan, "Improving Web Application Security: Threats and Countermeasures," Microsoft [Online] Available: https://msdn.microsoft.com/en-us/library/ff648644.aspx, Accessed June 2015, 2003.
10. Microsoft, "SDL Threat Modeling Tool," [Online] Available: https://www.microsoft.com/en-us/sdl/adopt/threatmodeling.aspx, Accessed June 2015, 2015.
11. B. Schneier, "Attack trees," *Dr. Dobbs journal*, vol. 24, no. 12, pp. 21–29, 1999.
12. P. Meland, D. Spampinato, E. Hagen, E. Baadshaug, K. Krister, and K. Velle, "SeaMonster: Providing Tool Support for Security Modeling," 2008, nISK-2008 conference.

The Effect of Vocals and Music Volume on Transcription Typing Speed

Anna Bramwell-Dicks, Helen Petrie, and Alistair Edwards

Department of Computer Science,
University of York, UK
{anna.bramwell-dicks,helen.petrie,alistair.edwards}@york.ac.uk

Abstract. Music psychologists have demonstrated that music is an affective medium that can be exploited to positively impact on a user's task performance across numerous contexts. This concept has the potential to be applied to work-related computing tasks, ideally resulting in improvements to an individual's productivity, efficiency and effectiveness. This paper presents the results from an experiment which investigated how transcription typing performance can be affected, in terms of speed, with a music accompaniment. The experiment focused on manipulating the presence or absence of vocals and the volume of the music, to identify if either or both of these parameters had an affect on typing speed. For the faster typists in the sample, when typing in English the presence of vocals in the music reduced typing speed. Whilst, across all of the typing abilities, louder music led to faster typing speeds when typing in English. With the more difficult Dutch transcription task, the music had no noticeable effect on typing speed, for either the slow or fast typists.

Keywords: Music psychology, human computer interaction, affective computing, transcription typing

1 Motivation and Background

In 2007, an objective study by Wellnomics Ltd [1] collected data regarding the working activity of employees from 95 organisations across different industry sectors from Europe, North America and Australasia. The employees had their computer use monitored over a 4 week period. Overall, the average amount of time spent working at a computer was 12.4 hours per week. But, in the UK and US these averages were higher at 16.8 and 14.5 hours per week, respectively. As such, computer use in these two countries, for these organisations, accounts for approximately 40% of their weekly working activities. Typing on a keyboard is a fundamental interaction method for using a computer. In Taylor's study, the UK users averaged 2,650 keystrokes per hour with 1,730 keystrokes per hour average for the users from the US. This data demonstrates that typing is an activity that forms a significant part of many peoples' working week.

In our research, we are trying to improve typing performance, as a significant work-related computing task, through the accompaniment of music. In particular, we are hoping to establish if and how typing performance is affected by certain dimensions of music. The aim is that if we can understand how particular dimensions affect typing performance, we will later be able to exploit these dimensions to improve performance in work-related computing tasks.

There is a wealth of evidence from the field of music psychology that listening to, or hearing[1], music while undertaking a variety of tasks can affect the users' behaviour in accomplishing that task. This change in behaviour can lead to both increases and decreases in task performance. From exercising [2, 3], to supermarket shopping [4], to drinking soda [5] there are many examples where background music affects the hearers' behaviour[2].

Kallinen [7] demonstrated that the speed of reading news stories on mobile devices while in a noisy café was affected by the tempo of the background music. Significantly faster reading speeds were achieved when listening to fast tempo music over slow tempo music. Though a café is not a typical work setting and reading the news on a mobile device is not a typical work task, this result shows there is potential to exploit the impact of particular dimensions of music to improve performance in work-related tasks, including reading.

The only experiment investigating how music affects typing was performed in 1937 [8]. Jensen performed an experiment with a large sample of skilled typists. The aim of the experiment was to identify what affect Jazz or Dirge[3] music had on typing performance (both speed and accuracy) in comparison with a 'without music' condition. Jensen found that significantly more errors were made when listening to Jazz music than in either the without music or Dirge music conditions, although there was no corresponding affect on typing speed. Contrastingly, the Dirge music resulted in significantly slower typing, though there was no difference in the error rates between Dirge and the without music conditions.

In Jensen's paper, there is no attempt to explain why Jazz music is detrimental to typing accuracy, or why Dirge music made the participants type slower. Neither is there consideration of why an increase in errors did not correspond with a decrease in speed. As such, it is interesting to further Jensen's research looking at the impact of accompanying music on typing transcription within a modern day context, to see if we can understand which, if any, parameters of the music may have caused the effects identified in 1937.

2 Method

2.1 Aims

This experiment is one of many by the first author that is attempting to build a body of work demonstrating that listening to music can have a positive impact

[1] *Listening* and *hearing* are not synonymous terms. The former implies active attention being consciously applied to the audio source, while the latter is a passive activity.
[2] See Bramwell-Dicks [6] for a review of previous literature in this area.
[3] Dirge can be described as slow, melancholy music and is often associated with funerals.

on work performance. Although transcription typing is not a typical task for many computer users, controlling the task in this manner allows us to focus on the mechanics of typing alone and not on typing while thinking, helping to strengthen the internal and construct validity of the experiment.

The primary aim of this particular experiment is to establish whether the presence or absence of vocals in a piece of music, together with the volume of that music, has a negative effect on transcription typing performance. This paper focuses on an analysis of the data collected in relation to typing speed, though accuracy measures were also collected[4].

2.2 Hypotheses

We expect that when transcription typing while listening to music that contains vocals, performance will be impaired as indicated through a reduction in typing speed. On a conceptual level our argument is that as transcription typing involves reading a stream of words using your eyes, hearing a different stream of words through your ears will make it harder to copy the visually presented material. Two different verbal stimuli being presented across two different senses is expected to make the task of transcription typing harder and therefore performance, in terms of typing speed, should be negatively affected.

We also expect that the louder the music is, the more distracting it will be, again having a negative impact on typing performance. As such, it is anticipated that the worst typing performance will be achieved when typing accompanied by loud music condition that contains vocals while the best performance should be achieved in the low volume instrumental music condition.

2.3 Design

Independent Variables (IVs) This experiment used a mixed between and within repeated measures design. Volume is the between groups variable (2 levels - low or high volume). The within participants IVs were vocals (2 levels - with vocals and without vocals), and text (2 levels - English and Dutch text).

Dependent Variables (DVs) Objective measures of typing performance were recorded for typing speed. The number of characters in the transcribed text was established, and then translated into Characters per Minute (CPM)[5].

2.4 Participants

Fifty-five participants (47 male, 8 female) took part in this experiment, all of whom were first year Undergraduate students in the Department of Computer

[4] See Bramwell-Dicks Thesis [9] for accuracy analysis

[5] Although Words per Minute is the typical measure used in typing research, CPM is a fairer comparison measure when the language differs between transcription tasks.

Science at the University of York, enrolled on the Human Aspects of Computer Science (HACS) module in 2013/14. Of these participants, 9 were non-native speakers of English. All of the participants were in the 18 - 24 age group. None of the participants spoke Dutch.

Five participants were randomly selected to win a £10 Amazon gift voucher on completion of the experiment. The fastest typist received a £30 Amazon gift voucher, the 2nd fastest £20 and the 3rd fastest £10.

2.5 Materials and Environment

Setting The experiment took place during two practical classes for the HACS module in the Department of Computer Science's software labs at the University of York. Half of the participants took part in 1 session, with the others doing the experiment the following day. The participants worked with another student who acted as the experimenter, while the first author oversaw the process and was there to ensure the experiment ran smoothly.

Music The piece of music used in this experiment is a piece of "Epic Indie Rock" music called "Atrophy" by The Doppler Shift [10]. To conduct an experiment with strong internal validity, the two variations of the music stimulus must be the same with the exception of the presence of vocals. To vary the vocals parameter while keeping the rest of the music the same requires the creation of new music mixes. It is not possible to simply subtract the vocal track from a pre-mixed recording of music. A multi-track recording[6] of the song was taken and mixed down[7] into two separate pieces of music, one of which will include all of the vocals tracks (typically the lead singer and backing vocals) while the other is mixed down without the vocals tracks included. This music was selected as the raw tracks were readily available and, given the source and obscurity of the band, it was unlikely to be a track that was familiar to the participants. Further, given the age range of the participants, it was hoped that they would largely enjoy the music.

Software Custom built webpages were developed for use in the experiments using HTML5, CSS3, JavaScript and PHP to control the interface and automatically collect and save the data on a server. The webpages were optimised for use in FireFox.

Each typing task had its own webpage which, when opened loaded the source, or presented, text into a non-editable text area on the left hand half of the screen. The webpage also contained a second, empty and editable text area within a form element, in a panel on the right hand side of the screen where the participant

[6] A method of recording where separate recording tracks (typically 16 or 24) are created for each instrument/combination of instruments.

[7] "Mix down" is the term given to the process of taking a multitrack recording and combining the tracks into a stereo (i.e. 2 track) output.

inputs their text transcription. As each participant made their first keypress in the transcribed text input field, the music automatically began playing. After 4.5 minutes an alert message appeared which ended the typing task and stopped the music playing. The experimenter then scrolled down the page to submit the form, which saves the transcribed text as a .txt file on the server, with a unique filename using the participants ID number.

The text passages were counterbalanced so that there was no pairing between music and text, and the order of conditions were varied to remove fatigue effects. Control logic was included in the webpages to ensure that each participant was doing the correct typing task for their identification number, with the right accompanying music for that task, as per the counterbalancing requirements. The webpages also controlled the volume of playback of the accompanying music. For participants with an even ID, the music was played at 100% volume, but participants with an odd ID heard music at 50% volume.

Text The English text came from The Outlaw of Torn by Edgar Rice Burroughs while the Dutch text was from Op Eigen Wieken, a Dutch translation of Louisa May Alcott's Good Wives. Dutch language was chosen as the source of the foreign language text because the alphabet is similar to English. The text was modified to remove any characters with accents or unusual punctuation e.g. instances of 'ő' were changed to 'o'.

Hardware All participants completed the typing tasks on the standard laboratory computers running Windows 7. Each participant was provided with a pair of identical, inexpensive headphones.

2.6 Procedure

Students in the class worked in pairs during the session, with one acting as the experimenter and the other as the participant. Each participant was given an identification number. The experimenter was provided with a set of briefing notes to read, which included information and instructions about the experiment they would be running.

After reading the briefing notes, the experimenter explained to the participant that they would be taking part in an experiment to investigate the impact of different types of background music on performance when copy typing. Informed consent was taken with an emphasis that all participants were able to withdraw at any time, without prejudice.

The participants were asked to make themselves comfortable by adjusting the chair, position of the keyboard, height of the screen etc. They were told that they should type as naturally as possible without prioritising either speed or accuracy and that whatever approach they took to the typing tasks, they should try to be consistent across all of the tasks. The experimenter then set the volume of playback on the computer to its maximum level.

The first typing task in each experiment was a practice task where the participants copied from a passage of the English text for 20 seconds, with the music beginning with the participant's first keypress. The participants were asked to verify that the volume of the music was not too loud and then the first of the "real" typing tasks was loaded into the web browser[8].

Participants then began typing in the real task, again triggering the start of music playback through the headphones. After 4.5 minutes the webpage generated an alert message to end the typing task and stop the music playing. When all 4 typing tasks were completed, the participants completed demographic questionnaires. When all of the participants had finished, the first author debriefed the whole class about the precise experimental hypotheses for the experiment.

3 Results

3.1 Descriptive Statistics

As with any quantitative experiment, it is important to visualise the data prior to analysis. In this instance, visual inspection of Fig. 1 implies a possibly bimodal distribution of typing speed in the English text, without vocals condition, which according to the hypothesis should be the easiest combination of task and music condition. There is a clear gap between 330CPM and 340CPM between the upper and lower tails of two distinct normal distributions. Therefore, the threshold level for designation as either a slow or fast typist was set at 330CPM[9].

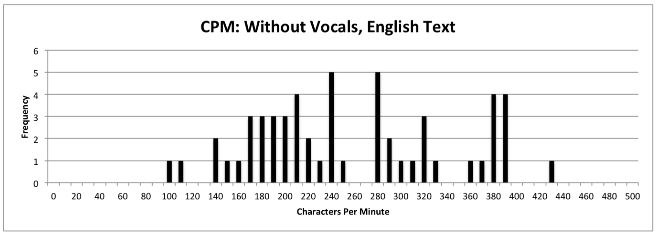


Fig. 1. Histogram of CPM for English Text, Without Vocals condition only

Fig. 2 shows a scatterplot of the CPM values for participants typing in the English text conditions, both with and without vocals in the accompanying music. This scatterplot validates the decision to specify the threshold value at 330 CPM. Participants whose native language is not English are highlighted

[8] If the participant felt the music was too loud they would have been swapped to the quiet group. However, this situation did not arise

[9] There is another gap at 260CPM, which could imply a complex series of separate skewed distributions. However, previous experiments by the first author (see [9] for details) have strongly supported the bimodal analysis method.

in Fig. 2. Participants P14 and P28 achieved comparatively slow typing speeds <120 CPM, inspection of their Dutch typing speeds showed they were achieving similar speeds in Dutch as English suggesting that both were languages that were unfamiliar to the participants. On the other hand, P54 was another non-native English speaker but achieves the fastest typing speeds in both conditions. As such, both P14 and P18 have been removed from the analysis, but data from P54 and the other 6 non-native English speakers was retained.

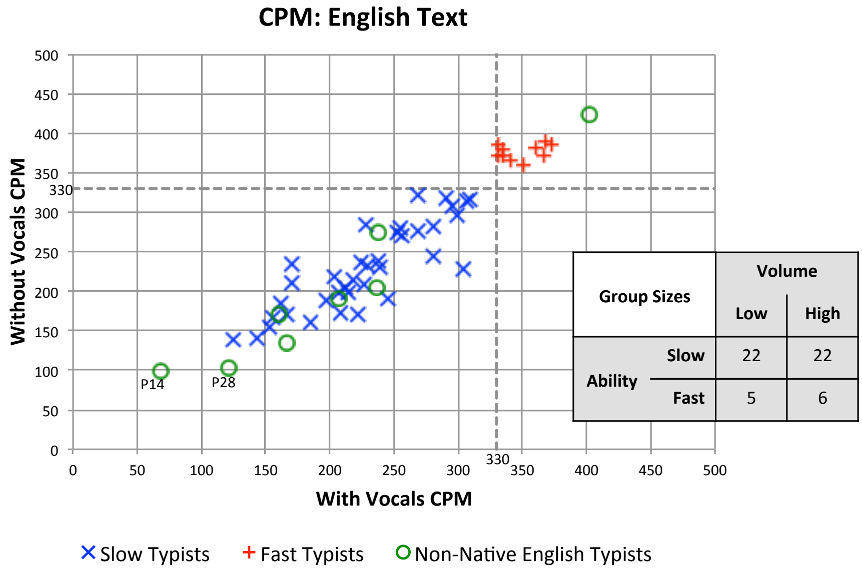


Fig. 2. Scatterplot of CPM for English Text Conditions, including Group Sizes (N)

In this experiment, analysis for normality of distribution using the Shapiro-Wilk's test showed that, when all participants were considered at once, three of the eight tested groups (37.5%) resulted in significantly non-normal distributions. But, when the participants were separated into their slow and fast classification structure the number of significant tests for non-normality reduced to just 2 significant outputs from the 16 tests (12.5%). This provides further evidence that the typing speeds are in fact bimodal distributions, and that the classification structure is appropriate.

3.2 Inferential Statistics

A 2 by 2 mixed within and between participants repeated measures ANOVA was performed on the CPM data, with volume level (low or high) and typist ability (slow or fast) as between-participants IVs. Before presenting the results, it should be noted that Levene's Equality of Variance is violated in both of the English text conditions. In the with vocals condition, the violation was marginal, F (3, 49) = 3.435, p = 0.024, but in the without vocals condition the heterogeneity of variance violation was strong, F (3, 49) = 4.928, p = 0.005.

If the group sizes were equal, this assumption violation could be ignored, however there are vastly differing group sizes in this experiment, as shown in Figure 2. As 2 of the 4 within-group conditions resulted in significant heterogeneity of error variance, and the group sizes are unequal, there needs to be further analysis[10] of the descriptive statistics before an ANOVA can be performed.

When sample sizes are different, heterogeneity of error variance is problematic if the larger variance is associated with the smaller group as the resulting F statistic is inflated leading to a higher chance of a falsely rejecting the null hypothesis (i.e. a type I error or "false positive" outcome) [11]. However, if the larger variance occurs in the larger group, the opposite effect is observed. In this situation, the F statistic is conservative, resulting in potentially not rejecting the null hypothesis in cases where the alternative hypothesis should have been accepted (i.e. a type II error, or "false negative" outcome).

Inspection of box plots of each condition showed that in this experiment the smaller variances were associated with the fast typist ability groups i.e. the groups with a smaller number of participants. As such, the F statistic in this analysis is at risk of being conservative rather than inflated so the analysis can proceed as it is more likely to result in a Type II error than a Type I error.

Significant omnibus effects were identified for text, F (1,49) = 306.715, p <0.000, ${\eta_p}^2 = 0.862$, and for the presence or absence of vocals, F (1,49) = 7.746, p = 0.008, ${\eta_p}^2 = 0.137$. Higher typing speeds were achieved when typing in English (M = 254.16 CPM, SD = 74.95 CPM) rather than Dutch (M = 162.60 CPM, SD = 47.42 CPM). Also, faster typing speeds were achieved when typing accompanied by music without vocals (M = 209.94 CPM, SD = 63.50 CPM) than with vocals (M = 206.82 CPM, SD = 58.87 CPM).

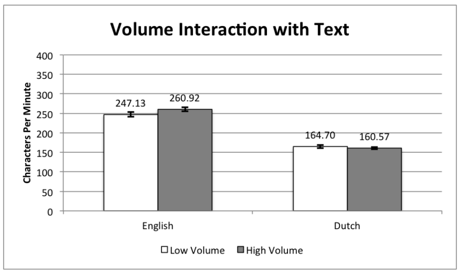
Significant two-way interactions were identified between text and volume, F (1,49) = 8.558, p = 0.005, ${\eta_p}^2 = 0.149$. Fig. 3(a)[11] shows similar means across the two Dutch conditions, but when typing in English the participants were faster in the high volume condition.

Another significant interaction was identified between vocals and volume, F (1,49) = 4.296, p = 0.043, ${\eta_p}^2 = 0.081$. Fig. 3(b) shows that the presence or absence of vocals had no obvious effect in the low volume condition, but in the high volume condition, the participants were faster when typing accompanied by music that did not contain vocals.

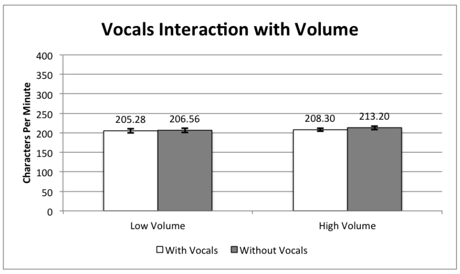
There was a further 2-way interaction between vocals and typing ability, F (1,49) = 8.467, p = 0.005, ${\eta_p}^2 = 0.147$. Fig. 3(c) shows that, the fast group of typists achieved faster speeds when typing accompanied by music without vocals, than music that did contain vocals.

[10] Data transformations (e.g. square root, inverse etc.) did not resolve the violations.

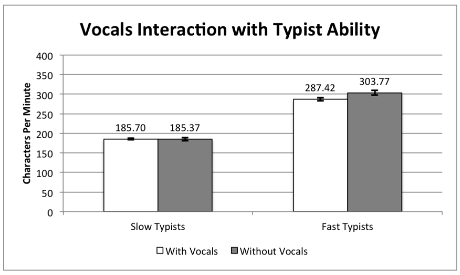
[11] The error bars in Fig. 3 show the standard error of the means.



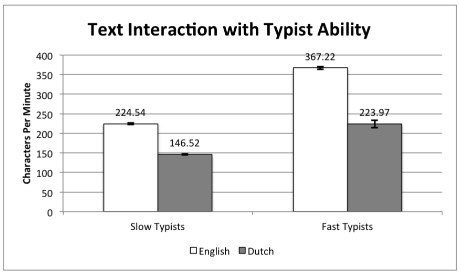
(a) Volume vs Text **(p = 0.005)**



(b) Vocals vs Volume **(p = 0.043)**



(c) Vocals vs Ability **(p = 0.005)**



(d) Text vs Ability **(p <0.000)**

Fig. 3. Significant Two-Way Interactions

The interaction between text and typing ability was significant, F (1,49) = 24.797, p <0.000, $\eta_p{}^2 = 0.336$, as shown in Fig. 3(d).

There was a trend towards a significant interaction between text and vocals, F (1,49) = 3.630, p = 0.063, $\eta_p{}^2 = 0.069$, as shown in Fig. 4[12]. In the Dutch typing tasks, the presence or absence of vocals appears to have had minimal impact on typing speed. However, in the English tasks, there is some evidence that the participants were faster in the without vocals condition. Although the outcome of this test does not quite reach significance, it is approaching significance and should the heterogeneity of error variances have been overcome, the test may have reached the required significant level.

Typing ability was a significant between groups factor, F (1,49) = 90.805, p <0.000, $\eta_p{}^2 = 0.650$, with fast typists achieving higher typing speeds (M = 295.60 CPM, SD = 36.41 CPM) than the slow typists (M = 185.53 CPM, SD = 40.75 CPM). While the volume group itself was not identified as a significant between groups factor, F (1,49) = 0.905, n.s.

4 Discussion and Conclusions

When transcription typing with music that contained vocals the participants were significantly slower. This verifies the initial hypothesis and indicates that

[12] The error bars in Fig. 4 show the standard error of the means.

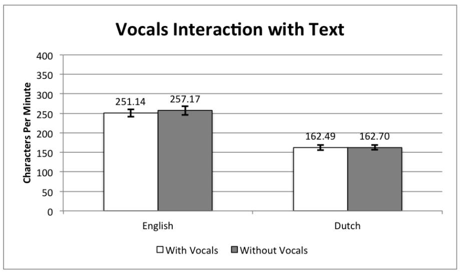


Fig. 4. Trend Towards a Significant Interaction Between Vocals and Text ($p = 0.063$)

when transcription typing the task is more difficult when the music contains vocals, than when the music is instrumental only.

The results have also demonstrated that the effect of background music on typing speed is strongly altered depending on the difficulty of the task. When typing in Dutch both the presence or absence of vocals, and the volume of the background music, had no obvious impact on typing speed. However, when typing in English, the skilled, fast typists found the typing task to be trivial and as a result their speed was affected by the presence of vocals in the music. When vocals were in the background music, these fast typists, were slower.

Although there was an omnibus effect for vocals, the interaction between vocals and typist ability is particularly interesting as for the slow typists the omnibus effect is not evident. However, even with the comparatively small numbers of fast typists in the experiment where the effect is clear, the omnibus effect is achieved. This demonstrates the strength of the effect for the fast typists group.

Although volume in itself was not a significant between groups factor, there was an interaction between the volume group and task language. When typing in English, the task that is easier, the participants were faster with louder music than with quieter music. This outcome goes against our initial hypothesis, where we expected the louder music to be more distracting and have a negative effect on typing performance. There is also an interesting interaction between volume and vocals, where in the low volume condition the presence of vocals in the music had little effect. But, in the high volume condition the participants were faster without vocals than with, indicating that for louder music vocals can have a negative impact on typing speed.

For many computer users there will be very little need to perform transcription typing within their computing activities. As such, the ecological validity of this experiment is somewhat weak. However, in this case, constraining the experimental task allows us to operationalise the parameters of the experiment, and what is lost in ecological validity is gained in internal and construct validity. If the users had been allowed to type freely, without transcribing a prescribed text, it would have been impossible to identify whether it was the speed of typing, or the time spent thinking, which was being affected by the background music.

In this experiment, the participants were listening to music in all of the experimental conditions. There was no experimental condition where the participants had no music playing and as such, their base typing performance without music was not established or included within the analysis. This limitation was caused primarily by the setting of the experiment, i.e. using a classroom-based methodological approach rather than conducting the experiment in a more typical laboratory setting. The classroom is typically a busy environment which is not silent, and although it was quiet when the experiment was taking place. It would have been impossible to control the room sufficiently to guarantee the construct validity for a without music condition and as such, one was not included in this particular experiment.

This experiment has shown that vocals can have a negative affect on transcription typing performance, but high volume music can have a positive influence on typing speed. The effect of accompanying music also seems to differ according to whether the participant is a skilled typist or less experienced. The results from this experiment clearly shows that different parameters of music have different effects, and suggests that we might be able to exploit loud, instrumental music to improve performance when working at a computer.

References

1. K. Taylor, "An analysis of computer use across 95 organisations in europe, north america and australasia," tech. rep., Wellnomics Ltd, 2007.
2. J. Waterhouse, P. Hudson, and B. Edwards, "Effects of music tempo upon submaximal cycling performance," *Scandinavian Journal of Medicine and Science in Sports*, vol. 20, no. 4, pp. 662–669, 2010.
3. J. Edworthy and H. Waring, "The effects of music tempo and loudness level on treadmill exercise," *Ergonomics*, vol. 49, no. 15, pp. 1597–1610, 2006. PMID: 17090506.
4. R. E. Milliman, "Using background music to affect the behavior of supermarket shoppers," *Journal of Marketing*, vol. 46, no. 3, pp. pp. 86–91, 1982.
5. H. McElrea and L. Standing, "Fast music causes fast drinking," *Perceptual and Motor Skills*, vol. 75, pp. 362–362, 1992.
6. A. Bramwell-Dicks, H. Petrie, A. D. Edwards, and C. Power, "Affective musical interaction: Influencing users behaviour and experiences with music," in *Music and Human-Computer Interaction* (S. Holland, K. Wilkie, P. Mulholland, and A. Seago, eds.), Springer Series on Cultural Computing, pp. 67–83, Springer London, 2013.
7. K. Kallinen, "Reading news from a pocket computer in a distracting environment: effects of the tempo of background music.," *Computers in Human Behavior*, vol. 18, pp. 537–551, 2002.
8. M. B. Jensen, "The influence of jazz and dirge music upon speed and accuracy of typing.," *Journal of Educational Psychology.*, vol. 22, pp. 458–462, 1937.
9. A. Bramwell-Dicks, *Title: To be confirmed*. PhD thesis, Department of Computer Science, University of York, UK, 2016.
10. T. D. Shift, "Atrophy." `http://www.cambridge-mt.com/ms-mtk.htm#AltRock`. Accessed: 2015-09-18.
11. B. G. Tabachnick and L. S. Fidell, *Experimental Designs Using ANOVA*. Thomson/Brooks/Cole, 2007.

Verifying a Very Racy Stack

Matt Windsor and Mike Dodds

University of York
{mbw500, mike.dodds}@york.ac.uk

Abstract. We discuss a work-in-progress on verifying the *SP pool*, a concurrent stack, using separation logic. We report on a first draft of a proof of several properties we then use to prove memory safety, moving towards functional correctness. We comment on the advantages and disadvantages of our chosen logic, *iCAP*, and lessons we can take away for applications of separation logic to low-level concurrent software.

1 Introduction

Concurrent programming is more necessary now than ever before. Modern CPU architectures, instead of building up with higher clock speeds and longer pipelines, are growing out with multi-core designs. To use every core effectively, we must spread our computations across them, moving from one sequential process to multiple concurrent threads that may operate in parallel.

For concurrent programs, our choice of data structures is key. Concurrent structures must tolerate interference: as multiple threads can access the structure at one time, we must ensure that no action is interrupted by other actions in a way that disrupts its own work. We want them to be fast, but optimisations tend to raise the possibility of unwanted interference. Data-races, the simultaneous writing and reading (or multiple-writing) of a shared variable, are a concern.

Proving these structures correct is hard—correctness for concurrent programs is more complex than in the sequential world. There are many more issues to worry about: race conditions, liveness, and the existence of more than one valid scheduling in most cases. Various correctness conditions define concurrent correctness: a common one is linearisability [4], the idea that, given a sequential specification of an algorithm, every operation takes effect atomically and in a schedule consistent with the program orders of each sequential method.

We examine the *SP pool*, a single-producer multi-consumer stack. Part of the multi-producer *time-stamped stack* of Dodds *et al.* [3], it uses low-level optimisations which make reasoning challenging. Building up to proving it correct, we discuss our approach to proving well-formedness conditions on the pool, using the separation logic *iCAP* [7]. Such logics extend Hoare logic to deal with disjoint resources, such as areas of memory, that can be separated and combined. With these, we can work on part of a large resource while preserving its context.

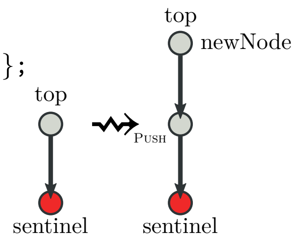
Paper structure. In Section 2, we outline the SP pool. We then discuss what it means for the pool to be well-formed (but not necessarily correct). We outline our proof method in Section 4, and evaluate it in Section 5. We end with an overview of what we have learned, and proposed future work.

2 Outline of the SP Pool

The pool consists of several methods, each acting on some top node `top`. We outline each in pseudo-code, with a diagram of its key atomic action; circles represent nodes of interest (shaded red if taken, or 'dead' to the algorithm). Here, and later on, we use **boldface** for method names, SMALL CAPS for atomic action names, `typewriter` for code, and *`italic typewriter`* for atomic code.

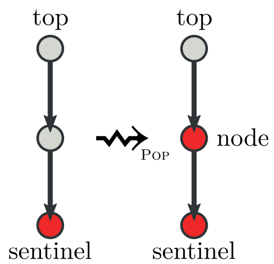
We create pools with **init**. New pools have one node, the sentinel, which marks the end of the pool. The sentinel is taken, has no element, and points to itself. We do not discuss **init** further: we assume it occurs outside shared state, and yields well-formed pools. The **push**[1] method creates a new node, points it to the current `top`, then atomically makes said new node `top`.

```
Node push(Element el) {
   Node newNode = {element: el, taken: false};
   newNode.next = top;
   top = newNode; // atomic action 'PUSH'
   fwdComp(newNode); // defined later
   return newNode;
}
```



The **pop**[2] method finds the newest non-taken node, if one exists. It does so via a simple non-atomic while loop from `top`. It then tries to mark that node as taken, through atomic compare-and-swap[3]. If it fails, another consumer took the node first, and the caller should **pop** again. Else, it returns the taken node's element.

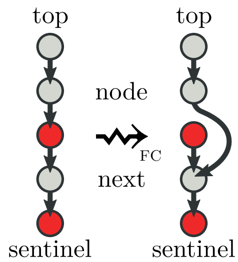
```
<Element, Status> pop() {
   Node oldTop = top;
   Node node = oldTop;
   while (node.taken) {
     if (node.next == node)
       return <NULL, EMPTY>; // Sentinel
     node = node.next;
   }
   if (CAS(node.taken, false, true)) { // atomic action 'POP'
     backComp(oldTop, node); // defined later
     fwdComp(node);
     return <node.element, SUCCESS>;
   }
   return <NULL, CONTENDED>;
}
```



1 **insert** in [3].
2 **getYoungest** and **remove** in [3].
3 For our needs, this is *`CAS(var, old, new)`*, which sets `var = new` if `var == old`.

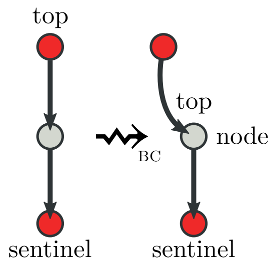
If we ignore **fwdComp** and **backComp**, our code implements all of the functionality of the pool. In practice, this is not sufficient: **push** adds new nodes to the pool, yet **pop** does not dispose of nodes: it marks them taken. This causes time and space leaks: taken nodes remain in memory, and must be traversed, yet are 'dead' to the algorithm. We must remove taken nodes from the chain of nodes between top and sentinel—the spine—, so that garbage collection (or similar) can dispose of them. To do so, we first define **fwdComp**:

```
void fwdComp(Node node) {
   next = node.next;
   while (next.next != next && next.taken)
     next = next.next;
   node.next = next; // atomic action 'FC'
}
```



We call this forwards compression: we sweep from a target node downwards, non-atomically, to find the next non-taken node. Then, we atomically re-point the target's `next` to that non-taken node. We now define backwards compression:

```
void backComp(Node oldTop, Node node) {
   CAS(top, oldTop, node); // atomic 'BC'
   if (oldTop != node)
     oldTop.next = node; // atomic 'FC'
}
```



From earlier, we know that everything between `oldTop` and `node` is taken. We can use this in two ways. Firstly, if `oldTop` is still `top`, we can move the latter to point to `node`: this is the `CAS` shown above. Secondly, if doing so does not create a cycle, we can perform atomic action FC between `oldTop` and `top`.

Unusual observations on the SP pool. The pool, though simple in implementation and intuitively a perfectly reasonable stack, violates many common-sense assumptions about correct stacks. For example, the top node is not always the most recently inserted node (and nodes are not totally ordered by insertion); the pool is not a linked list; and nodes compressed out of the spine can still be traversed and even re-appear later. All of these surprising facts arise from compression: it races with itself and other methods, and alters the pool shape in ways unusual for a stack. However, in terms of adhering to the specification of a concurrent stack, compression has no visible effect on the pool.

As both forwards compression and popping use non-atomic traversals through next-pointers, they can observe a node at the moment a forwards compression moves it out of the spine. This is why we cannot assume the pool is a linked list; while the spine forms one, and all non-taken nodes are within it, there may exist several active branches where a traversal is still sweeping a series of already-compressed nodes. Our SP pool, then, is a reversed tree, with all branches eventually converging onto the spine.

Non-atomic traversals also pose problems. In figure 1, we have a spine containing some taken nodes, and non-taken nodes `A` and `B` in that order.

Step	Thread 1	Thread 2	Thread 3
1	**push** inserts `node`		
2	**push** begins FC on `node`		
3	**fc** loop stops with `next=A`		
4		**pop** marks `A` taken	
5			**pop** marks `node` taken
6			**pop** performs BC on `node`
7			**pop** begins FC on `node`
8			**fc** loop stops with `next=B`
9			**fc** changes `node.next=B`
10	**fc** changes `node.next=A`		

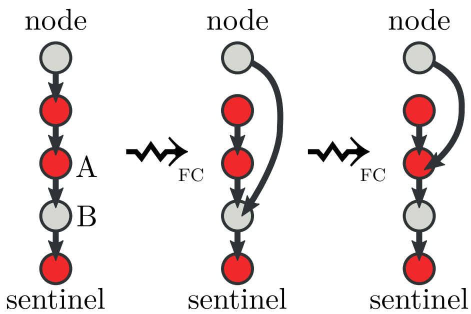


Fig. 1. Sample schedule for data-race between two forwards compressions, and diagram of steps 9 and 10.

Thread 1 replaces Thread 3's compression `node`→`B` (step 9) with the smaller `node`→`A` (step 10), undoing part of 3's work. This is harmless: the only difference is the number of taken nodes on the spine, and this cannot affect behaviour.

A similar effect occurs between backwards compression and insertion. When the new node is created in local state, it takes the current top pointer as `next`. While no other insert can change the top pointer in the meantime, backwards compression can. This is also harmless, as it can only moves a top pointer down past taken nodes, which will be re-introduced by the insertion.

3 Well-Formedness of the SP Pool

We intend to prove well-formedness properties for the pool. To do so, we now examine what we mean by 'well-formedness', first in informal overview and later in a more precise mathematical sense.

Orders on the SP pool. Our concept of well-formedness is heavily based on the observation and constraint of various orders between pool nodes. Before we define well-formedness, we first give each an informal definition:

Insertion predecessor. The relation of node pairs (a, b) where a was inserted directly after b;
Insertion order. The reflexive transitive closure of the above ($a > b$ if a was inserted any time after b), and is total due to the single-producer nature of the pool;
Physical predecessor. The relation of node pairs (a, b) where a`.next` is b;
Physical order. The reflexive transitive closure of the above ($a > b$ if a reaches b by one or more next pointers); this is the reverse of the intuitive order, but is defined as such to make relating it with insertion order easier.

We now define well-formedness as a collection of invariants, mostly on the above orders, each preventing various possible pool failures. This resembles the invariants given by O'Hearn *et al.* [6] for a similar data structure, and several of these invariants are necessary for proving that the methods behave correctly.

Order-subset. Physical order subsets insertion order; actions cannot cause cycles or non-'last-in-first-out' ordering, but physical order may be partial;
Single-predecessor. Each node has exactly one physical predecessor;
Sentinel-predecessor. Only the sentinel can point to itself, preventing cycles;
Sentinel-reach. Each node reaches the sentinel through a path of `next`-pointers; all nodes eventually have a common successor and form the same pool;
Spine. The spine is the only path on which non-taken nodes can lie; compression cannot sweep non-taken nodes off it;
Taken–non-taken disjoint. No node can be both taken and non-taken;
Element-injection. Every non-taken node has exactly one element, and every element is mapped to at most one node.

Stability. As the pool is concurrent, consumers can be scheduled in parallel with each other, and with the single producer. Thus, it is constantly changing under our atomic actions. This poses a threat to well-formedness: we have to show that once we've established it, no atomic action at any point can violate it. This idea of stability is a key component of our proof.

State tuples. In our proof, a pool state is captured by the 8-tuple

$$
\begin{array}{lrr}
\langle N & \subseteq \mathit{Node}, & \text{(non-taken nodes)} \\
T & \subseteq \mathit{Node}, & \text{(taken nodes)} \\
t & \in \mathit{Node}, & \text{(top node)} \\
s & \in \mathit{Node}, & \text{(sentinel node)} \\
\mathcal{P} & \subseteq \mathit{Node} \times \mathit{Node}, & \text{(physical predecessors)} \\
\mathcal{T} & \subseteq \mathit{Node} \times \mathit{Node}, & \text{(insertion predecessors)} \\
E\rangle & \subseteq \mathit{Node} \times \mathit{Element} & \text{(element map)}
\end{array}
$$

We achieve the physical order, which we call $\geq_{\mathcal{P}}$, by taking the reflexive transitive closure of $\mathcal{P}$; insertion order $\geq_{\mathcal{T}}$ is achieved similarly for $\mathcal{T}$.

Abstract atomic actions. Each state-modifying action in the pool is atomic: any other thread either sees a state before the action, or a state afterwards. We can characterise all changes to the pool state in terms of these actions.

The pool has four atomic actions—Push, Pop, FC, and BC—, which correspond to the italicised statements in our pseudo-code. We begin by giving sketches of these in the form of functions on state tuples. These sketches are initially too weak, as they permit ill-formed pools, but we later constrain them. First, we define the atomic actions of pushing and popping some node n:

$$\begin{aligned}\textsc{Push}\ n &\approx \langle N, T, t, s, \mathcal{P}, \mathcal{T}, E, i\rangle \\ &\quad \rightsquigarrow \langle N \uplus \{n\}, T, n, s, \mathcal{P} \uplus \{(n, m)\}, \mathcal{T} \uplus \{(n, i)\}, E \uplus \{(n, e)\}, n\rangle \\ \textsc{Pop}\ n &\approx \langle N \uplus \{n\}, T, t, s, \mathcal{P}, \mathcal{T}, E \uplus \{(n, x)\}, i\rangle \rightsquigarrow \langle N, T \uplus \{n\}, t, s, \mathcal{P}, \mathcal{T}, E, i\rangle\end{aligned}$$

Forwards-compressing a node a to replace its next-pointer b with c, if all nodes moved between branches as a result are taken, corresponds to the atomic action

$$\text{FC}\ a\ b\ c \approx \langle N, T, t, s, \mathcal{P} \uplus \{(a, b)\}, \mathcal{T}, E, i\rangle \rightsquigarrow \langle N, T, t, s, \mathcal{P} \uplus \{(a, c)\}, \mathcal{T}, E, i\rangle$$

It should now be clear why $\mathcal{P}$ and $\mathcal{T}$ are distinct: FC 'abbreviates' the former to remove unnecessary paths through taken nodes. We must show that all edits to $\mathcal{P}$ are consistent with $\mathcal{T}$, only differing by unlinking taken nodes.

We saw in section 2 that **backComp** has two atomic actions. The second is an application of FC, as seen from its signature above. The first, if the new top value is c and everything inserted between t and c is taken, is

$$\text{BC}\ c \approx \langle N, T, t, s, \mathcal{P}, \mathcal{T}, E, i\rangle \rightsquigarrow \langle N, T, c, s, \mathcal{P}, \mathcal{T}, E, i\rangle$$

This is why t and i are distinct: BC can change the top to something other than the most recently inserted node. (However, i and t must still be connected by a run of taken nodes). We keep a record of i so that, when we insert and update $\mathcal{T}$ with the new node, we can attach it to i and preserve the totality of $\geq_{\mathcal{P}}$.

Well-formedness predicate. We now know enough to give our informal well-formedness predicates a mathematical definition. We do this in figure 2.

4 Outline of the Proof

Our chosen proof technique, *iCAP*, is a separation logic. This means that it has a concept of separating conjunction (the $*$ operator) which allows two disjoint resources to be combined into one (and a resource to be split into two disjoint resources). This gives us modular reasoning, in that a proof over one resource is still valid when the resource is combined with any sufficiently disjoint context.

At the concrete level, as per the original definition of separation logic, these resources are areas of memory, which may be split and combined during reasoning. However, *iCAP* only permits this form of reasoning during atomic actions, as these are the only times where we are allowed to modify state. Otherwise, we interact with the state via a region predicate, which gives us a view of the shared memory region as a labelled transition system. In our proof, this predicate is

$$\mathit{region}(\mathcal{S}, T_x, I_x(\texttt{top}), r)$$

If S_x is the x member of the tuple S,

$$\begin{aligned}
wf_{os}(S) &\triangleq \geq_{S_\mathcal{P}} \subseteq \geq_{S_\mathcal{T}} && \text{(Order-subset)}\\
wf_{1p}(S) &\triangleq \forall x \in S_N \cup S_T \cdot |S_\mathcal{P}[\{x\}]| = 1 && \text{(Single-predecessor)}\\
wf_{sp}(S) &\triangleq \forall x \in S_N \cup S_T \cdot (x, S_s) \in S_\mathcal{P} \iff x = S_s && \text{(Sentinel-predecessor)}\\
wf_{sr}(S) &\triangleq \forall a \in S_N \cup S_T \cdot a \geq_{S_\mathcal{P}} S_s && \text{(Sentinel-reach)}\\
wf_s(S) &\triangleq \forall a \in S_N \cup S_T \cdot a \in S_N \implies S_t \geq_{S_\mathcal{P}} a && \text{(Spine)}\\
wf_{TN}(S) &\triangleq S_T \cap S_N = \emptyset && \text{(Taken–non-taken disjoint)}\\
wf_E(S) &\triangleq (\forall n \in S_N \cdot \ |S_E(n)| = 1) \wedge (\forall e \in Elem \cdot |\{n | (n, e) \in S_E\}| \leq 1) && \\
& && \text{(Element-injection)}\\
wf(S) &\triangleq wf_{os}(S) \wedge wf_{1p}(S) \wedge wf_{sp}(S) \wedge wf_{sr}(S) \wedge wf_s(S) \wedge wf_{TN}(S) \wedge wf_E(S)
\end{aligned}$$

Fig. 2. Formal definitions of the well-formedness predicates.

where $\mathcal{S}$ is the set of abstract states the pool can currently take, T_x a transition relation characterising the pool's atomic actions, I_x a interpretation function mapping members of $\mathcal{S}$ to concrete heaps (here parameterised by the address of the pool top), and r a label identifying the specific shared region in question.

Filling the region predicate. Our proof method, then, first involves defining the possible values of $\mathcal{S}$, T_x, and I_x. For $\mathcal{S}$, we lift our 8-tuple representation to sets; then, in region predicates, we can encode facts about the pool. For example,

$$pooled(r, \texttt{top}, \texttt{n}) \triangleq region(\{S \mid \texttt{n} \in S_N \cup S_T\}, T_x, I_x(\texttt{top}), r)$$

has that `n` is in pool r. Predicates on the same region label are $*$-joinable resources, with the resulting $\mathcal{S}$ being the intersection of the individual sets: $pooled(r, \texttt{top}, \texttt{n1}) * pooled(r, \texttt{top}, \texttt{n2})$ tells us that pool r contains `n1` and `n2`.

Lifting our well-formedness predicate $wf(s)$ to a region predicate, we capture what it means to know that a pool r is well-formed:

$$wfpool(r, \texttt{top}) \triangleq region(\{S \mid wf(S)\}, T_x, I_x(\texttt{top}), r)$$

Now, we consider T_x. We define it in figure 3, as a map from the atomic actions to their transitions. As in the weak transitions from section 3, each transition is given as a 'before and after' relationship between two schematics of state tuples. Now, however, we add additional constraints on the pre–and post-states.

Many of the constraints we add here concern insertion orderings. For example, $allT_{\geq}(a, b, T, \mathcal{T})$ means that a was inserted after b, and everything inserted in the range $[a, a \wedge_\mathcal{P} b)$ (where $a \wedge_\mathcal{P} b$ is the meet point, or greatest common physical descendent, between a and b) must be taken:

$$allT_{\geq}(a, b, T, \mathcal{P}, \mathcal{T}) \triangleq a >_\mathcal{T} b \wedge \forall c \cdot (a \geq_\mathcal{T} c >_\mathcal{T} a \wedge_\mathcal{P} b) \implies c \in T$$

The idea behind these constraints is that, if compression works correctly, then the methods can imagine that insertion and physical order are equal. Much of

$$
\begin{aligned}
T_x(\textsc{Push}) &\triangleq \langle N, T, t, s, \mathcal{P}, \mathcal{T}, E, i\rangle \\
&\rightsquigarrow \begin{pmatrix} \langle N \uplus \{n\}, T, n, s, \mathcal{P} \uplus \{(n, m)\}, \mathcal{T} \uplus \{(n, i)\}, E \uplus \{(n, e)\}, n\rangle \\ \wedge\, i \geq_{\mathcal{T}} m \geq_{\mathcal{T}} t \wedge \mathit{allT}_{\geq}(i, m, T, \mathcal{P}, \mathcal{T}) \wedge \mathit{allT}_{\geq}(m, t, T, \mathcal{P}, \mathcal{T}) \end{pmatrix} \\
T_x(\textsc{Pop}) &\triangleq \langle N \uplus \{n\}, T, t, s, \mathcal{P}, \mathcal{T}, E \uplus \{(n, \text{—})\}, i\rangle \\
&\rightsquigarrow \langle N, T \uplus \{n\}, t, s, \mathcal{P}, \mathcal{T}, E, i\rangle \\
T_x(\mathrm{FC}) &\triangleq \langle N, T, t, s, \mathcal{P} \uplus \{(a, b)\}, \mathcal{T}, E, i\rangle \wedge a >_{\mathcal{T}} c \wedge\ \mathit{allT}_{\geq}(b, c, T, \mathcal{P}, \mathcal{T}) \\
&\rightsquigarrow \langle N, T, t, s, \mathcal{P} \uplus \{(a, c)\}, \mathcal{T}, E, i\rangle \\
T_x(\mathrm{BC}) &\triangleq \langle N, T, t, s, \mathcal{P}, \mathcal{T}, E, i\rangle \wedge\ \mathit{allT}_{\geq}(t, c, T, \mathcal{P}, \mathcal{T}) \\
&\rightsquigarrow \langle N, T, c, s, \mathcal{P}, \mathcal{T}, E, i\rangle \wedge\ \mathit{allT}_{\geq}(t, c, T, \mathcal{P}, \mathcal{T})
\end{aligned}
$$

Fig. 3. The transition relation T_x, refining the informal transitions from section 3.

our proof thus involves constraining the relationship between the two orders. Finally, if `top` contains the address of our pool's top node, our I_x is[4]

$$I_x(\texttt{top})(\langle N, T, t, s, \mathcal{P}, \mathcal{T}, E, i\rangle) \triangleq I_{nodes}(\mathcal{P}, N, T, E) * \texttt{top} \mapsto t \wedge t \in N \cup T$$

$$I_{nodes}(\mathcal{P}, N, T, E) \triangleq$$

$$\circledast\, a \in N \cup T \cdot \exists n \in N \cup T,\ b \in \mathit{Bool},\ e \in \mathit{Elem} \cdot$$

$$\mathit{node}(a, e, b, n) \wedge (a, n) \in \mathcal{P} \wedge b = \mathit{false} \iff (a \in N \wedge e = E(a))$$

$$\mathit{node}(x, e, s, n) \triangleq x\texttt{.element} \mapsto e * x\texttt{.taken} \mapsto s * x\texttt{.next} \mapsto n$$

Method proofs. To reduce our stability obligation from considering every possible interleaving of statements in every method to considering only the atomic actions in T_x, we must now show that each pool method only modifies the region through correct use of those atomic actions. We must also show that *wfpool* holds throughout every method, assuming the stability result we prove later.

To do this, we use Hoare-style reasoning: we produce a proof outline, or a sequence of judgements, over every command C_n in the method, $\{P_n\}\ C_n\ \{P_{n+1}\}$, where P_n is the pre-condition of C, and P_{n+1} its post-condition (and the pre-condition for C_{n+1}). We aim to show here that, given some sufficiently strong first and last P that are consistent with our intuition of the pre-conditions and post-conditions of each method, each C changing the shared state represents one of the atomic actions from T_x. This ensures that the proof can rest on the stability of the predicates used in the P conditions over T_x.

We reason about the method entirely using region predicates until we reach an atomic action. At this point we open the region, using I_x to translate our

[4] $\circledast\, a \in A \cdot P(a)$ denotes a separate heap for each a such that $P(a)$ holds in that heap, ie $P(a_1) * P(a_2) * \ldots * P(a_n)$

predicates to a concrete heap onto which we apply our atomic action. After this, we close the region: we must show that the region predicate after close can be mapped by I_x into our modified concrete heap, and that the change from original to modified region predicates is consistent with an action in T_x. This reduces the burden of stability proof from checking each predicate against each method step to that of checking against each action in T_x.

Now, we show that, for every atomic action a and predicate p, p is stable with regards to a. To demonstrate, suppose we want to prove wf_{os} stable. First, only $\mathcal{P}$ and $\mathcal{T}$ are mentioned in wf_{os}; this means that only atomic actions affecting those can invalidate the predicate: Push and FC. For Push, we add $\{n, m\}$ to $\mathcal{P}$ and $\{n, i\}$ to $\mathcal{T}$. From the transition, we know that $i \geq_{\mathcal{T}} m \geq_{\mathcal{T}} t$, and as wf_{os} originally held then we can deduce that $n \geq_{\mathcal{T}} m$, $n \geq_{\mathcal{P}} m$, and the transitive orderings added to both sides must preserve the subset relation. For FC, we note that we substitute some $\{a, c\}$ for $\{a, b\}$ in $\mathcal{P}$, but $b >_{\mathcal{T}} c$, thus the resulting physical order is still a subset of insertion order.

5 Evaluation of the Proof Method

In this section we evaluate whether *iCAP* was a suitable tool for verifying the SP pool. Largely, we feel our use of *iCAP* has been a success. It provides a well-defined structure for organising proofs, reducing mental burden and improving confidence. This has helped us debug our proofs and discover previously unknown properties of the pool. For example, stability of a predicate $S_t = x$ for some x was required by an early insertion proof, and falsified when attempting to prove it under BC; we thus discovered the race between backwards compression and insertion. However, we note room for potential improvement on our method:

Leaky abstraction. The usual stack abstraction is a sequence, with popping and pushing occurring on one end. Our abstract view of the pool is the 8-tuple, which provides only a thin layer of abstraction, and leaks pool implementation details, such as: taken-ness, partial physical ordering, and sentinels. Abstract predicates help by hiding these under a concept of accumulating facts about the pool. However, being able to relate the concrete details of the pool to a more abstract notion of a stack in the logic, perhaps through some form of abstraction refinement, may assist with answering questions about functional correctness.

Unused complexity. The *iCAP* logic targets re-entrant, higher-order concurrent algorithms. Therefore, *iCAP* is large, with many concepts not used in this work. This is not a major problem, as the complexity rarely affects our particular proof, but our case may be better served with a smaller logic. For example, a smaller logic may be easier to automate—the lack of automation in *iCAP* makes our proof difficult to check. Frameworks such as *Views* [2] or *IRIS* [5] may make defining a small, focused logic on top of an existing, sound foundation easier.

Abstract state should be separable. If our abstraction was a sequence, we could use the machinery of separation logic to pull out parts of the sequence and work on them while holding the rest of the sequence invariant. Working with list segments and related shapes is well-understood in separation logic.

However, because our abstract state is effectively a collection of assorted set-theory facts about the concrete state, none of the tools of separation logic are available in *iCAP* until we open the region. This is a missed opportunity, because our transition relations must account for every part of our tuple, even if only one item is changed; this can be seen in the FC transition. If we could separate the parts of state we know will not change, and work only with the changing elements, this would eliminate much redundancy in our proof.

Again, abstract predicates partially help by masking parts of tuples for which we hold no knowledge. However, eventually, our proof must deal with the fact that $\mathit{pooled}(r, a)$ expands towards the set $\{\langle N, T, t, s, \mathcal{P}, \mathcal{T}, E, i\rangle \mid a \in N \cup T\}$.

We could abandon abstraction and work with the separable concrete state, but our tree-shaped state is difficult to work with using standard separation logic machinery. Work on logics for tree structures, such as that by Calcagno *et al.* [1], may fit better because the pool tree would be a first-class structure in the logic. This would give us logical tools to separate and re-combine pool fragments.

6 Conclusion and Future Work

Our aim was to prove the stability of various well-formedness properties on the SP pool; we now have a first draft of that proof. The ideal next step is to finalise and release it. One could then adapt the proof to a more suitable concurrent logic, to address the impedance mismatch described in section 5.

To prove the SP pool correct, we would look to a formal correctness condition, such as linearisability. Initial attempts using the Stack Theorem of Dodds et al. [3] show promise. These attempts depend on our *wf* invariant, but do not interact with *iCAP*. The Stack Theorem is based on the demonstration that linearisability-violating chains of events across the pool's lifetime cannot happen; iCAP is more suited to per-method and per-action specifications. A technique for order-based linearisability in the logic would complement the existing work.

References

1. Calcagno, C., Dinsdale-Young, T., Gardner, P.: Adjunct elimination in context logic for trees. Information and Computation 208(5) (2010)
2. Dinsdale-Young, T., Birkedal, L., Gardner, P., Parkinson, M., Yang, H.: Views: Compositional reasoning for concurrent programs. POPL (2013)
3. Dodds, M., Haas, A., Kirsch, C.M.: A scalable, correct time-stamped stack. POPL (2015)
4. Herlihy, M.P., Wing, J.M.: Linearizability: A correctness condition for concurrent objects. TOPLAS 12(3), 463–492 (Jul 1990)
5. Jung, R., Swasey, D., Sieczkowski, F., Svendsen, K., Turon, A., Birkedal, L., Dreyer, D.: Iris: Monoids and invariants as an orthogonal basis for concurrent reasoning. pp. 637–650. POPL (2015)
6. O'Hearn, P.W., Rinetzky, N., Vechev, M.T., Yahav, E., Yorsh, G.: Verifying linearizability with hindsight. PODC '10 (2010)
7. Svendsen, K., Birkedal, L.: Impredicative concurrent abstract predicates. In: Shao, Z. (ed.) Programming Languages and Systems (2014)

A Pilot Study Investigating The Process of Risk Assessment and Re-Accreditation in UK Public Sector Systems

Michael Davies, Jassim Happa, and Ioannis Agrafiotis

University of Oxford
{michael.davies,jassim.happa,ioannis.agrafiotis}@cs.ox.ac.uk

Abstract. The provision of security of information and its supporting ICT infrastructures within the UK Public Sector is long-standing and well established, underpinned by a wide range of standards and literature. For the many security practitioners that are employed in the sector, a number of important concerns have experientially emerged over several iterations of policy and standards that have been developed over time to govern this activity. The aim of this qualitative pilot study was to explore these concerns from the perspective of security practitioners employed in the sector. Data was collected from six security practitioners via semi-structured interviews. Subsequent transcripts were analysed using a Thematic Analysis approach. This identified four significant themes that suggest that re-accreditation rarely occurs outside of the formal accreditation cycle, and point to the underlying reasons why this is the case.
Given that the National Technical Authority (NTA) is undertaking a comprehensive review of Information Assurance in the Public Sector, this pilot study is well-timed. This qualitative investigation of the issue is novel, and while aspects of these themes may be recognised anecdotally by consultants in this field, this pilot study provides an inductive, data-driven account of the issues with re-accreditation that transpired both within and across participants transcripts. Finally, this study gives some indication of potential further research that could be undertaken in the area.

Keywords: Public Sector, System Accreditation, Thematic Analysis

1 Introduction

For the large number of information security practitioners that are employed in the UK Public Sector, a number of important concerns have experientially emerged in the areas of residual risk assessment and accreditation, over the several iterations of policy and standards that have been developed to govern this activity. The main aim of this qualitative pilot study was to investigate these concerns from the perspective of the security practitioners employed in this area, and in the process attempt to answer the research question

Is there evidence that Public Sector organisations undertake the analysis of Residual Risk and system re-accreditation outside of the mandated accreditation cycle?

Accreditation is an assessment of an ICT system or service against its information assurance requirements, resulting in acceptance of residual risk, in the context of the business requirements and information risk appetite, and is in essence a precursor for the approval of a system to operate [1]. The clear implication being that without a current and valid system accreditation in place, a system should not be operating.

The pilot study collected qualitative data of the experiences of consultants during semi-structured interviews [2]. The subsequent transcripts were interpreted using Thematic Analysis and four superordinate themes were identified:

- System re-accreditation rarely occurs outside of the mandated accreditation cycle
- Key concepts and processes not understood at Senior Information Risk Owner (SIRO) level
- Under-resourcing in terms of sufficiently trained and experienced personnel in key areas
- Disproportionate nature of accreditation/re-accreditation in todays Information Assurance environment

Thematic analysis is a method for identifying, analyzing and reporting patterns within data sets [3]. This has enabled the issues considered by the participants as most relevant to emerge as themes in the data, with participants also providing insight as to what improvements they believe should be undertaken to the current processes. Finally, this study gives some direction on potential further research in the area.

2 Methodology

A qualitative approach to answering this research question was chosen based on the ability of such an approach to encourage participants to provide open and expansive responses to the range of question posed, and not be limited by trying to artificially impose quantitative values on their uniquely qualitative experiences. Information Risk consultants will have formed their own impressions and opinions of the way information risk is managed and communicated in organisations. In expressing these qualitative, personal impressions, the researcher anticipated, that the experiences of failures to fully utilise residual risk analysis in the decision to re-accredit, within these organisations, would be recalled, identified, and articulated by the participants to the researcher. The interview methodology selected for the study was a one-on-one, semi-structured interview based approach, where the participants were asked the same set of questions. The questions were phrased in an open-ended way and participants were encouraged to elaborate on their responses to each question while minimising the influence of the researcher. A complete script has been employed, but the questions posed to the participants have been designed to minimize the follow on inputs of the researcher, allowing the participants to freely explore and consider their responses to each question. Given that the participants all have considerable experience in this area of information assurance consultancy the researcher was confident in the ability of the participants to respond positively to this approach, the researcher may have prepared some questions beforehand, but there is a need for improvisation [2]. This semi-structured approach provides the opportunity to introduce a carefully planned and consistent interview schedule and range of questions to the participants, with the flexibility to probe deeper into participant responses of particular interest. McNamara [4] has stated that this style of interview approach will enable the researcher to en-sure that the same general areas of information are collected from each participant; this provides more focus than the conversational approach, but still allows a degree of freedom and adaptability in getting information from the participant. This was the approach employed here.

This pilot study employed a small number of qualitative interviews undertaken with 6 individuals who are practitioners in the area of UK Public Sector system accreditation or who have been closely involved professionally in public sector IT. Five of the six participants are certified as Security Information Risk Advisors (SIRAs) under the CESG CCP[1] scheme. One of the participants was a Lead Practitioner with extensive experience of Public Sector accreditation. The sixth participant was originally recruit-ed to undertake a test interview, but with wide experience of information and system management in the Public Sector, the contribution made during the interview was of considerable relevance to the pilot study and the researcher was of the view that data of such relevance should not be excluded. The interviews witnessed these participants discuss their broad range of experiences of system accreditation in the UK Public Sector, before being requested to consider in detail, the sub-processes of residual risk analysis and the decision making processes around re-accreditation.

The questions used in the study are provided below:

1. What are the main issues you have experienced when consulting to the UK Public Sector?
2. What specific issues have you encountered when undertaking residual risk assessments?
3. When consulting to the Private Sector have you experienced similar or different issues?
4. How was the risk appetite of the organisation communicated to you during the consultative process?
5. Have you ever known full re-accreditation of a system to be undertaken out-side the accreditation cycle?
6. What were the key decision-making processes involved?
7. In your opinion what influence did the risk appetite of the organisation have on the decision-making process?

[1] CESG Certified Professional (CCP)

8. Can you identify ways in which the residual risk analysis process can be improved?
9. The National Technical Authority is currently reviewing the full Information Assurance process, what do you expect to see coming out of this review?

A formal test of the interview was undertaken using an individual who was originally unconnected with the project but who did have wide experience of the area of investigation. This test-run followed the structure of the planned interview sessions in detail, enabling the researcher to identify any weaknesses in the planned interview format in terms of timings and question content [5]. Some issues with the structure of the inter-view were discovered. Feedback from the test interview was employed to refine the interview question to ensure that they supported the aims of the project. The interviews were audio recorded with participants granting consent for this on the Project Ethical Consent Form. Full written transcripts were produced with anonymisation provided for all participants. Following transcription the audio recordings were destroyed, in accordance with Ethical Consent. Subsequent analysis of the transcripts was then undertaken to draw out the key themes from the participant responses.

Thematic analysis, given its flexibility and utility, was used to identify patterns across the range of data provided by the participant consultants, using coding and theme development to draw out the recurring and relevant themes in participant responses. A theme is able to capture something important about the data in relation to the research question [3].

3 Preliminary Results

The decision to adopt a Thematic Analysis of the transcribed data provided the op-portunity to explore in detail individual and personal lived experience and to exam-ine how participants are making sense of their personal and social world [6]. The Information Assurance community in the UK is small with many consultants, particu-larly those employed in the public sector, having many shared experiences. This facil-itated the exploration of individuals experiences of residual risk and re-accreditation across a wide area of the public sector, and lead to the identification of criteria for the selection of illustrative quotations that identified key superordinate themes and sub-themes. A tabular summary of the key themes identified, and a selection of supporting statements provided by the participating consultants is provided in Table 3.

Preliminary analysis of the important themes identified by the participants pointed to an emerging relationship between these themes, which is hierarchical in nature. This relationship is shown in Figure 1, and considered in the Discussion Section below.

4 Discussion

The first emerging theme identified the common practice of failure to re-accredit systems outside the formal accreditation cycle. This process is supposed to be

Key Theme 1. Lack of re-accreditation outside the cycle	Key Theme 2. Key concepts not understood by SIRO	Key Theme 3. Under-resourcing of key personnel	Key Theme 4. Accreditation Process inappropriate
I struggle to think where I have seen anything re-accredited on a 12-monthly cycle . . . there is no point in going back and re-accrediting the whole system if there hasnt been any major changes Never in my experience . . . I think resources are the main issue	SIROs arent actually sufficiently educated to understand (residual) risk . . . the SIRO is often too remote doesnt understand it doesnt under-stand what he is being asked to sign up for There is a definite lack of under-standing especially by the SIROs and even senior management what is the definition of residual risk	People with appropriate skills not being in positions . . . the government witnesses too many changes that the pace of government accreditors not keeping up that could be down to resources that could be down to people with appropriate skills not being in positions . . . Danger is we may be starting to rely on the service integrator doing risk management due to lack of roles within the department	. . . the RMADS which ends up as a door stopper its a 300-page document that no one really uses it and in my experience I have spoken to lots of accreditors and pan-government accreditors and they have never found any practical use of that It has just ended up as a tick box exercise now basically you have got this whole stack of documents which are RMADS which no one looks at them You have a heavy duty work-force and policies to follow that makes it (re-accreditation) not impossible but its almost it can take as long as the whole implementation

Table 1. Key Themes and supporting statements (data) from participant interviews.

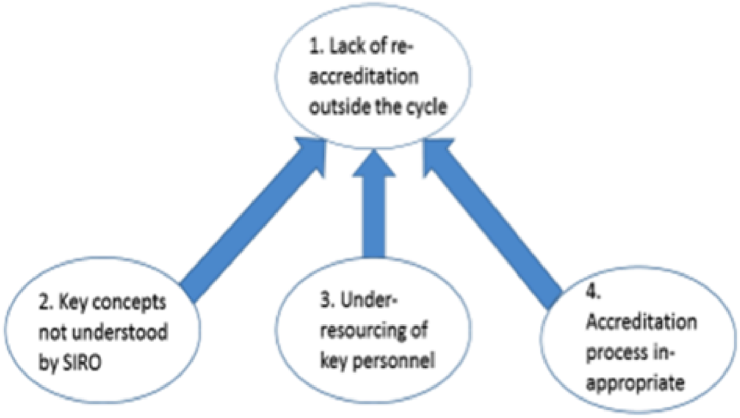


Fig. 1. Inter-relationship between the four superordinate themes identified in the Pilot Study.

under-taken in response to any one of a series of mandated conditions for re-accreditation being met, and is compulsory, and not optional. A plausible relationship of this theme to the subsequent themes identified is shown graphically above, and represents the idea that these are in fact the underlying causes for re-accreditation not being under-taken. For example, in the second theme there was a perceived causal relationship between the lack of re-accreditation and the fact that key concepts and processes are not well understood at the important senior stakeholder level of SIRO. If this is the case, there is the further related consequence that there will be a subsequent detrimental effect on the communication of key risk information in organisations and therefore on the effectiveness of decision making. Theme 3, which identified issues around the under resourcing of sufficiently skilled and experienced personnel is not contentious, reflecting the frequently observed shortage of skilled information assurance practitioners faced by all sectors. Finally, Theme 4 considers the inappropriate nature of the full accreditation / re-accreditation process in todays information security environment. The focus of all of the participants here was on the process of Risk Management Accreditation Document Set (RMADS) generation, in support of the accreditation decision. In the fast moving rapidly evolving cyber security environment that challenges government departments and agencies today, the consultants all agreed that a process that can take months to undertake, and that can be undone by some form of offensive cyber activity in seconds, has possibly had its day. It is interesting to consider the possible relationship between themes 2, 3 and 4. For example, is it possible that the issues identified at SIRO level are associated with the under re-sourcing of key personnel? This could be usefully explored in further work.

Conversely, as shown by some of the consultants interviewed, there are occasions when re-accreditation is undertaken within the mandated accreditation cycle.

" ... Yes, when the system being accredited has been moved to a different location and used in a different context ... "

"... The whole process doesnt necessarily need to happen unless there is a major change ... "

This may therefore occur when some significant change to a system has occurred, and where not to re-accredit would represent an obvious organisational failure to follow mandated standards and guidelines. At what level within the organization the decision to re-accredit is actually taken would provide an interesting area for further research.

5 Conclusions

This pilot study examined some of the issues associated with current processes around re-accreditation in the UK Public Sector. There is clear scope for further investigation of the issues identified. This could be achieved with an extension of this study using an increased sample size and with interviews undertaken with the full range of partic-ipants involved in this process, particularly at accreditor and SIRO level. Based on this pilot study there are clear indications that Public Sector organisations do under-take analysis of residual risk and risk appetite to support the decision making process for ICT system re-accreditation, but usually only for situations where some significant change has occurred. Where system re-accreditation should have, and has not been undertaken, Public Sector systems may be operating in an unsafe state. In the second theme there is a perceived causal relationship between this lack of re-accreditation and key concepts and processes not being understood at senior stakeholder level (SIRO), in Public Sector organisations. There is also the suggestion that this has a subsequent detrimental effect on communication of important elements of risk infor-mation and on decision making. These would appear to be suitable candidates for more extensive research. Moreover, the results identify measures that could be undertaken, such as enhanced training for SIROs, and more effective recruitment and reten-tion of key personnel, which could be adopted in the short-term to address some of the issues identified.

References

1. CPNI: Risk Management and Accreditation also released as HMG INFOSEC STANDARD NO. 2 Risk Management and Accreditaion of Informaiton Systems. (2005).
2. Myers, M.D., Newman, M.: The qualitative interview in IS research: Examining the craft. Inf. Organ. 17, 226 (2007).
3. Braun, V., Clarke, V.: Using thematic analysis in psychology. (2006).
4. McNamara, C.: General Guidelines for Conducting Research Interviews, http://managementhelp.org/businessresearch/interviews.htm. (2009).

5. Kvale, S.: Doing Interviews. (2007).
6. Lyons, E., Coyle, A.: Analysing Qualitative Data in Psychology. Sage Publications Ltd (2007).

A Study of Students' Motivations to Participate in a Crowdsourcing Project to Support their Disabled Peers

Fatma Layas, Helen Petrie, Christopher Power

University of York, York YO10 5GH, United Kingdom

{fal503, helen.petrie, christopher.power}@york.ac.uk

Abstract. Crowdsourcing is currently popular as a way to draw on the resources from a large number of people to complete tasks. One area in which there has some exploration of crowdsourcing is to support people with disabilities. However, there is little exploration of what motivates people to participate in such socially responsible crowdsourcing projects. In this study we sought to explore what students think would motivate them to participate in a crowdsourcing project to support their blind and partially sighted peers by describing images in electronic learning resources. Using an online questionnaire, we found that students thought they are particularly motivated by their sense of altruism and earning money. Other motivators included knowing that they are contributing to a large project, improving their academic skills and enhancing their job opportunities in the future. Whereas, students were neutral in what they thought about: getting academic credits, the fun and entertainment of the activity, being connected with other students on their course. Students thought they would not be motivated by being in a competition with other students, drawing attention to their skills, the social recognition they would receive, and to pass the time. Implications for how to motivate socially responsible crowdsourcing projects are discussed.

1 INTRODUCTION

With the rapid development of the use of technology in society, higher educational institutions now often use virtual learning environments (VLEs) to facilitate teaching and learning. Lecturers typically upload their digital learning resources to the VLE, which can include digital slide packs, used for lectures and seminars. Students can study these before the class, look at them during the class and review them afterwards. Such digital learning resources now include many images: photographs, graphs, diagrams and drawings. While some of these images are for decorative reasons, many of them are vital to understand the materials and being able to learn from them. Blind and partially sighted students can be hindered in their learning if these images are not accessible to them.

Screenreaders and screen enlargement programs used by visually impaired students can now deal well with the text in digital learning resources, and with specialist notations such as mathematics and music. However, these assistive technologies cannot extract any information from images, these need descriptions provided by a sighted person, and a sighted person who understands the context and educational authorial intent of the image. Unfortunately lecturers do not generally provide explicit descriptions themselves of every image in their learning resources and lack the time and resources to do so. This is a situation very parallel to the image description problem on websites [11]. However, this is a situation in which crowdsourcing could be applied successfully to support blind and partially sighted students.

1.1 Crowdsourcing to support people with disabilities

Although crowdsourcing is still at an early stage of development, it has been applied in many contexts. A number of crowdsourcing projects have supported people with disabilities. Early initiatives in this area were surveyed by Bigham, Ladner and Borodin [2]. One of the interesting early initiatives was the ESP games project [14] which involved the players tagging images on the Web as part of the game play. In addition to being fun and entertaining for players, this contributed to describing images on the web, which is particularly helpful for screenreader users. While this may not be considered a classical crowdsourcing project, as it is a game rather than merely asking people to undertake tasks, it has the characteristic of using the power of the crowd to solve a problem for particular groups of web users. In another initiative Takagi et al. [13] developed a true crowdsourcing project to address web accessibility problems encountered by disabled web users. The Social Accessibility Project allowed disabled web users to report website problems which would then be posted for the crowd of supporters to solve.

These initiatives have shown that crowdsourcing has the potential to be a useful means of providing support for people with disabilities in the problems they face in daily life and in particular in access to information. However, many technical and practical challenges need to be addressed to achieve a successful crowdsourcing project in this area. One of these challenges is the need to understand what motivates people to participate in such crowdsourcing projects.

1.2 Motivation of the crowd

Although crowdsourcing projects to support people with disabilities have not investigated the motivation of the crowd, research has investigated motivational factors that lead people to participate in crowdsourcing projects in general. These factors can be divided into intrinsic and extrinsic motivations. In the context of crowdsourcing, extrinsic motivation means that crowd members are motivated by factors external to the crowdsourcing task; these might be financial rewards, social recognition, or job opportunities. Whereas intrinsic motivation means that crowd members are motivated by factors related to the task itself, such as the satisfaction they get from undertaking the task.

Lakhani et al. [9] explored the motivation of the crowd to participate in a scientific problem solving project, InnoCentive. In an online survey active crowd members rated the importance of 16 motivational factors. While crowd members were motivated by monetary rewards, they were significantly more motivated by intrinsic motivation factors, such as the intellectual challenge of problem solving. Similar research by Oreg and Nov [10] explored the motivation of the crowd in the domain of open source software (Sourceforge) and content (Wikipedia) developments. An online survey with 185 crowd members found that crowd members rated self-development and reputation building higher as motivations than members of the open source content group. However, members of the open source content group rated altruism higher than the open source software group.

Brabham [3,4] examined the motivation of the crowd in a number of artistic design projects. He investigated the motivation of crowd members on iStockPhoto [3], an online royalty-free photography agency for amateur photographers. An online survey with 651 "iStockers" showed that the desire to make money, improve one's creative skills, and to have fun were the strongest motivators, whereas passing the time and building a network of friends were the least popular motivators. Brabham [4] conducted instant messaging interviews with 17 members of Threadless, an online t-shirt company that crowdsources the design process of its product through an ongoing online competition. This study found that crowd members had a mix of intrinsic and extrinsic motivations: money, develop creative skills, the potential to leverage freelance design work, and the love of the Threadless community. Most interestingly, one motivation was addiction to the Threadless community, whereby the crowd members see themselves as part of Threadless rather than just external contributors.

While there is an overlap between what would motivate people to participate in crowdsourcing projects in general, the motivators for crowd members in different projects such as, InnoCentive, Threadless and iStockphoto are different. As a result each project has a group of motivations that are not necessarily helpful to apply in other projects.

Thus, what would motivate people to participate in one crowdsourcing project, would not necessarily motivate them to participate in another one. We are interested in investigating how best to motivate people to participate in crowdsourcing projects to support people with disabilities, and more specifically how to motivate sighted students to participate in a crowdsourcing project to support their blind and partially sighted peers. This is because we are developing a crowdsourcing project, Describe-IT, to create descriptions of images in digital learning resources. We are investigating what factors to incorporate into the project to motivate student participation. In this initial study, we studied students' perceptions of their motivations, and the motivations of other students, to participate in such a crowdsourcing project. In subsequent research, we will compare their perceptions with their actual behavior using the project when it goes live.

1.3 DescribeIT: A crowdsourcing project to describe images to help visually disabled students

The particular crowdsourcing project we are developing is to support students with visual disabilities in higher education. While there are many types of materials in higher education settings that require adaptation for people with disabilities, this project is targeting the provision of descriptions of images in digital learning resources for students with visual disabilities as an example of a crowdsourcing task that students can undertake in an educational setting. Therefore, we are developing a crowdsourcing project called DescribeIT, which allows sighted students to provide descriptions of images in PowerPoint presentations for their blind and partially sighted peers. The application will allow students to describe as many or as few images as they wish, at any time they wish.

The aim of this paper is to investigate what students think would motivate them to participate in such a crowdsourcing project and their perceptions of other students' motivations.

2 METHOD

2.1 Design

The study investigated how students perceive different motivational factors that would affect their participation and the participation of other students in a crowdsourcing project to describe images for their blind and partially sighted peers.

2.2 Participants

144 students responded, 87 women and 57 men. Their age range was from 18 to 51 years, with a mean age of 23.7 years (SD=6.5). The University of York is a very international institution, and while majority of participants were from the United Kingdom (52%, 75), 10% (14) were from China, 3% (4) were from Malaysia, 3% (4) were from the USA, 2% (3) were from each of India, Saudi Arabia and Spain. The remaining participants (26%) were spread relatively evenly across 20 other countries.

All participants reported using social media, such as Facebook. However, only 15% of the participants reported participating in crowdsourcing projects, including Wikipedia, Mechanical Turk, Kickstarter, and many other projects.

As an incentive to complete the questionnaire, participants were entered into a prize draw for one of 10 £10 Amazon vouchers.

2.3 Materials

The online questionnaire consisted of three sections:

Likelihood and motivation to participate in the project: consisted of four questions, which investigated students' perceptions of their likelihood of participating in the image description project and the likelihood of other students on their course, and their own statements of their motivations to do so and their peers.

Motivational factors: In this section, a set of 12 motivational factors was presented, taken from the research on the motivation of the crowd in crowdsourcing projects. These factors are listed in Table 2. Students were asked to indicate on a 7 point Likert scale how much each factor would motivate them to participate in the image description project. For each factor, students were also asked to explain why they had given that rating.

Demographic and online activities questions: this section collected demographic data and collected data about participants' online activities.

2.4 Procedure

A recruitment email was sent out, with one reminder email sent five days later. A further round of recruitment was conducted several months later, to increase the sample size. On each occasion, the online questionnaire was available for one week.

3 RESULTS

Students gave a rating mean of 4.6 (SD=1.7) that they would participate in the project. However, students thought the likelihood of their peers participating in the project less, with rating mean of 4.0 (SD=1.4). To investigate students' ratings of the likelihood of participation was significantly above the midpoint of the rating scale, a one-sample t-test was conducted comparing the ratings with the midpoint rating of 4. This showed that students' ratings of their own likelihood to participate were significantly higher than the neutral midpoint ($t = 4.27$, $df = 143$, $p < .001$). Obviously, their ratings for their estimation of the likelihood that other students would participate was not significantly different from the neutral midpoint of the rating scale, as the mean rating was exactly 4.0.

Students were also asked to state what they thought would motivate them and what they thought would motivate other students in their course to participate. Students noted a number of intrinsic and extrinsic factors, including: rewards and prizes, wanting to help others, knowing that the project is useful to others, knowing a friend or a family member who is blind and any form of monetary reward. Table 1 shows all the motivational factors mentioned by students which they thought would motivate themselves or motivate other students.

Students were then asked to rate to what extent each of 12 motivational factors would motivate them to participate. The results are summarized in Table 2 which shows that students rated their "sense of altruism, wanting to help other students" as the highest rated motivational factor to participate in the image description project. Whereas, "Being in a competition with other students" was the lowest rated motivational factor.

Table 1. Factors which participants thought would motivate them or other students

Motivational Factors	Self N (%)	Others N (%)
Rewards, such as vouchers, gift cards, cash	43 (29.8%)	65 (45.0%)
Knowing that the project is useful to others	26 (18.0%)	4 (2.7%)
Wants to help others	19 (13.0%)	11(7.8%)
Knowing a friend, a colleague, or a family member who is blind	10 (6.9%)	4 (2.7%)
Reviewing the slides' materials	8 (5.5%)	3 (2.0%)
If the task is easy to do	7 (4.8%)	3 (2.0%)
Getting more information about the project	6 (4.0%)	9 (6.0%)
Personal interest to accessibility area	5 (3.0%)	5 (3.0%)
Recognition from University	4 (2.7%)	0 (0.0%)
The ability to stop whenever they want	4 (2.7%)	1 (0.7%)
Raise awareness about partially sighted and blind students' needs	3 (2.0%)	6 (4.0%)
If they were asked from lecturer	3 (2.0%)	0 (0.0%)
Improving one's skills	2 (1.0%)	2 (1.0%)
For Fun	2 (1.0%)	1(0.7%)
Competition with others	1 (0.7%)	0 (0.0%)
Course credits	1 (0.7%)	2 (1.0%)
Enhance their CV	1 (0.7%)	3 (2.0%)
Feedback from blind students	1 (0.7%)	0 (0.0%)

A series of one-sample t-tests were carried out to see if students' motivations are statistically different from the neutral midpoint rating of 4 (See Table 2). The results showed that students rated their sense of altruism, wanting to help other students, improving their academic skills, being paid for their efforts, enhancing their job opportunities in the future, and knowing that they are contributing to a large project significantly above the midpoint. On the other hand, drawing attention to their skills, the social recognition they would receive, to pass the time and being in a competition with other students were rated significantly below the midpoint. Table 3 presents some typical comments from students about each of the motivational factors.

4 DISCUSSION AND CONCLUSIONS

The present study explored what students perceive would motivate them to participate in a crowdsourcing project to support their blind and partially sighted peers. Students were asked the likelihood that they would participate in such a crowdsourcing project and their perception of the likelihood of other students on their course participating. They were also asked what would motivate them. Then they were asked to rate 12 motivational factors.

Table 2. Mean, standard deviation and one-sample t-tests for the top 12 motivational factors

Motivational Factor	Mean Rating	SD	t value	p (df=143)
Your sense of altruism, wanting to help other students	5.35	1.45	11.2	0.001
Being paid for your efforts	4.91	2.05	5.36	0 .001
Knowing that you are contributing to a large project	4.66	1.66	4.76	0 .001
Enhancing your job opportunities in the future	4.59	2.03	3.51	0 .001
Improving your academic skills	4.38	2.01	2.27	0 .005
The fun and entertainment of the activity	4.14	1.89	0.93	n.s.
Getting academic credits	4.06	2.27	0.22	n.s.
Being connected with other students on your course	3.92	1.96	-0.47	n.s.
Drawing attention to your skills	3.22	2	-4.65	0 .001
The social recognition you would receive	2.84	1.81	-7.63	0 .001
To pass the time	2.61	1.82	-9.11	0 .001
Being in a competition with other students	2.50	1.84	-9.76	0 .001

Overall, students showed an interest in participating in the socially responsible crowdsourcing project to describe images in digital learning resources for their blind and partially sighted peers. Students rated their likelihood of participating positively although they rated the likelihood of other students participating as neutral. This difference could be due to what social psychologists call the "fundamental attribution error" [12], which proposes that people interpret their own behavior very much in terms of the specific situation, but interpret the behavior of others in terms of persistent personality traits. Thus, when asked to predict their own participation, students think of the specific situation of helping other students who are at a disadvantage, but when asked to predict the behavior of other students, they think of the general helpfulness of other students, not in the context of helping disabled students.

The study found that altruism was rated highly as a motivation. This finding is in agreement with previous studies [8], [10] that found people's sense of altruism is one of the leading self-reported motivations for participating in socially responsible crowdsourcing projects such as Wikipedia. The present study suggests that students were not particularly motivated by the fun and entertainment of the activity, which is contrary to some other studies [3], [5,6,7]. This could be due to the nature of the proposed task in our project, as students did not think of describing an image as fun task, whereas Brabham's studies examined artistic design tasks. Probably for this reason some students suggested adding aspects of gamification to make the task more fun to do.

Table 3. Typical comments from students about the top 12 motivational factors

Motivational Factor	Typical Comments from students
Your sense of altruism, wanting to help other students	I feel that all students deserve an equal chance to learn effectively. I have been helped many times … helping others is a way to return the favour.
Being paid for your efforts	Money is a pretty good incentive for students. I volunteer a lot. Getting paid is an even bigger incentive.
Knowing that you are contributing to a large project	I like to do things that do good to the world. I don't think the size of the project is relevant aside from the effect of helping more people.
Enhancing your job opportunities in the future	Probably a nice thing to have in your CV. I prefer to enhance my opportunities building networks, improving my knowledge and skills.
Improving your academic skills	The more you describe the visual materials to someone else the better you understand it.
The fun and entertainment of the activity	Seems like fun. The idea of wracking my brains to describe something is intriguing. It would be a new experience for me, and probably fairly enjoyable.
Getting academic credits	Academic credits would help me towards my job goal, hence it is partially important.
Being connected with other students on your course	Colleagues of mine will be interested, then we have a common ground to build a friendship. It would probably bring me closer to my classmates especially if there were any visual impaired.
Drawing attention to your skills	So any excuse for us to show any skills is worthwhile. I don't think that the activity involves any really impressive skills.
The social recognition you would receive	I believe people should do things for self-satisfaction and to help - not for other people to think highly of them. Maybe a weekly leaderboard would make me want to do more.
To pass the time	There are numerous other things I could be doing to pass time. It's not hard work, so I can see this being viable procrastination.
Being in a competition with other students	I'm not very competitive. I feel this would detract from the positive aspects of the project.

Monetary reward was the highest rated self-reported factor in the extrinsic motivation group, in agreement with the findings of Brabham [3] and Lakhani et al. [9]. The fact that improving one's skills was a significant motivator was similar from those from Brabham's studies [3, 4, 5]. Previous studies [4], [8] emphasized the importance of the community within the crowdsourcing projects and being connected to other participants in crowdsourcing projects. However, the present study suggests that being connected to other students was not thought to be a primary motivation to participate in the project. This finding is similar to Brabham's results from his study of iStockphoto [3], where he found a lack of connection between iStockphoto members due to trust issues. The motivation of getting academic credits has not been examined in any previous studies that we are aware of. However, the present study showed a neutral self-report rating for this factor. Students did not think they would be motivated by "drawing attention to their skills" or "the social recognition they would receive". While competing to solve difficult scientific problems might sound appealing to students, being in competition to help their blind and partially sighted peers was not. As with iStockphoto members [3], students thought they would not be motivated by passing the time. As our participants are full-time students they often have much work to do and according to them if they have spare time they would spend it doing more enjoyable things.

It is worth noting that participants repeatedly mentioned that knowing someone blind or partially sighted would motivate them to participate in the project. Obviously that is a particular motivation when the crowdsourcing project is to help these groups, but it has wider implications for socially responsible crowdsourcing.

These results present a first insight into what students say would motivate them to participate in a crowdsourcing project to support their blind and partially sighted peers. Participants said they would be particularly motivated to participate in the crowdsourcing project by their sense of altruism and monetary rewards. In addition, participants said that they would be motivated by knowing that they are contributing to a large project and that this contribution would enhance their job opportunities in the future. New motivational factors emerged from the study including: knowing a friend or colleague who would benefit from the project, and knowing that the project is useful to other colleagues. Whereas participants were neutral in what they thought about getting academic credits, improving their academic skills, the fun and entertainment of the activity, being connected with other students on their course. Participants thought they would not be motivated by being in a competition with other students, drawing attention to their skills, the social recognition they would receive, or to pass the time.

An important point is that the present study and most studies that have investigated crowd members' motivations have relied on self-report of motivations. As Antin and Shaw [1] note, this methodology is very vulnerable to the "social desirability bias", meaning participants may respond to questions in ways that they believe that they should, that are socially acceptable. Therefore, we are currently conducting studies with the same group of students to investigate their actual behavior in participating in the crowdsourcing project to provide image descriptions for electronic learning resources.

The findings of the current study will be compared with students' actual behavior, to examine the differences between their self-reported predictions of their behavior and motivations and their actual behavior.

5 REFERENCES

1. Antin, J. & Shaw, A. 2012. Social Desirability Bias and Self-Reports Of Motivation: A Study of Amazon Mechanical Turk in the US and India. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2012) (May 5-10, Austin, TX, USA)*. 2925-2934. New York: ACM Press.
2. Bigham, J.P., Ladner, R. E. & Borodin, Y. 2011. The design of human-powered access technology. *In Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '11) (October 24 – 26, Dundee Scotland).* 3-10. New York: ACM Press.
3. Brabham, D. C. 2008. Moving the crowd at iStockphoto: The composition of the crowd and motivations for participation in a crowdsourcing application. *First Monday* 13(6).
4. Brabham, D. C. 2010. Moving the crowd at Threadless: Motivations for participation in a crowdsourcing application. Information, *Communication and Society* 13(8): 1122-1145.
5. Brabham, D. C. 2012. Motivations For participation in a crowdsourcing application to improve public engagement in transit planning. *Journal of Applied Communication Research* 40(3): 307-328.
6. Hossain, M. 2012. Crowdsourcing: activities, incentives and users' motivations to participate. *International Conference on Innovation, Management and Technology Research (ICIMTR2012), (May 21-22, Malacca, Malaysia)* 501-506. IEEE.
7. Kaufmann, N., Schulze, T. & Veit, D. 2011. More than fun and money. Worker motivation in crowdsourcing: A study on Mechanical Turk. *In Proceedings of the Seventeenth Americas Conference on Information Systems (August 4-7 2011, Detroit, Michigan)*. 1-11. Association for Information Systems (AMCIS).
8. Kuznetsov, S. 2006. Motivations of contributors to Wikipedia. *ACM SIGCAS Computers and Society* 36(2): 1.
9. Lakhani, K. R., Jeppesen, L. B., Lohse, P. A. & Panetta, J. A (2007). The value of openess in scientific problem solving. Division of Research, Harvard Business School.
10. Oreg, S. & Nov, O. 2008. Exploring motivations for contributing to open source initiatives: The roles of contribution context and personal values. *Computers in Human Behavior* 24(5): 2055-2073.
11. Petrie, H., Harrison, C. & Dev, S. 2005. Describing images on the Web: a survey of current practice and prospects for the future. *Proceedings of 3rd International Conference on Universal Access in Human-Computer Interaction. (July 22 – 27, Las Vegas)*.
12. Ross, L. & Nisbett, R.E. 2011. *The person and the situation: perspectives on social psychology*. London: Pinter and Martin.
13. Takagi, H., Kawanaka, S., Kobayashi, M., Itoh, T. & Asakawa, C. 2008. Social accessibility: Achieving accessibility through collaborative metadata authoring. *In Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility (ASSETS '08) (October 13–15, 2008, Halifax, Nova Scotia, Canada). 193-200*. New York: ACM Press.
14. Von Ahn, L. & L. Dabbish 2004. Labeling images with a computer game. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '04) (April 24–29, 2004, Vienna, Austria).* 193-200. New York: ACM Press.

Towards a Fitness Function for Musicality using LPM

Sahar Arshi, Darryl N. Davis

Department of Computer Science, University of Hull, UK.
{ S.Arshi@2014., D.N.Davis@ }hull.ac.uk

Abstract. In this article the features of Liquid Persian Music (LPM) software in employing cellular automata (CA) evolution is briefly described. Controlling synthesizer parameters by means of pattern matching over the emergent nature of CA is an important characteristic of LPM. The hypothesis is that LPM will enable us to explore new dimensions of music composition. Here, the main focus is to analyze the LPM output voices in the search for finding proper tools for enhancing them in a musical way. This paper focuses on applying Zipf's law on various configurations of produced audio to fulfill aesthetic goals. The nature of a fitness function for measuring the pleasantness of music has been targeted as a concluding remark for future research.

Key words: Cellular Automata, Liquid Persian Music, Synthesis Toolkit, Zipf's Law, Computer Music, Artificial Music Composition, Evolutionary Algorithms, Fitness Function.

1 Introduction

The advent of cellular automata originally dates back to 1940s, when Von Neumann was looking forward to develop a system capable of reproduction, comparable in certain respects with biological breeding [1, 2]. Cellular automata was studied as a dynamical system in 1960s [3] . Cellular Automata are discrete dynamical systems whose global intricate behaviour is determined by reciprocal influence of elementary identical individuals [3–6] .Each cell has *k* finite states at time *t*, and all the cells evolve simultaneously. The state of a cell at time *t* depends on its state and its neighbors' states at time *t-1* [7]. In the one dimensional elementary CA (which is the subject of this study), the permutations of each cell with its two adjacent neighbors specifies eight situations. Once allocated to binary states, the selection of one of the 256 local transition rules specify the CA evolution.

Wolfram studies on CA recognize four classes of behaviour, namely, fixed, cyclic, chaotic, and complex [3]. Li and Packard [8] subdivided the second class to three further subgroups, namely heterogeneous, periodic with intervals greater than one, and locally chaotic.

Cellular automata exhibiting myriad genres of behaviour have been targeted as a creative tool for artists. The generated patterns by CA possess self-similar characteristics known as 1/f noise or pink noise. In music composition pink noise yields a pleasing balance between regularity and abrupt variations [9]. Various cellular automata musical systems have been designed since the advent of this field. Amongst these systems, CAMUS and Chaosynth [10, 11] have gained popularity. CAMUS exploits Game of Life and Demon Cyclic Space, and uses a Cartesian space mapping to MIDI for achieving the musical triplets. Chaosynth is a cellular automata sound producer based on the generation of sound granules.

Liquid Persian Music is a CA based toolkit for exploring various musical possibilities. This paper expands the capabilities of LPM and establishes connections with genetic algorithms. The article consists of five sections. In the second section a brief overview of LPM is given, followed by an investigation throughout the output of the software. In the third part the Zipf's law has been applied on the output distributions of cellular automata as a musical aesthetical measurement. Zipf's law is able to recognize the pink noise distribution in data series. The fourth section focuses on further experiments and clarifies some future research directions for LPM.

2 LPM Project Features Overview

Liquid Persian Music is an auditory software tool developed at the University of Hull. LPM explores the idea of artificial life systems in producing voices. The software takes advantage of Synthesis Toolkit (STK) [12] for implementing the physical model of a stringed musical instrument. A model of which its parameters are controlled by defined pattern matching rules. Pattern matching rules classify output from CA and update the parameters of the synthesiser to yield musical effects [13, 14]. OpenAl library is responsible for propagating the producing voices.

The elementary CA used in LPM consists of an assembly of cells arranged in a one dimensional array. In every time step of CA evolution, the pattern matcher extracts the difference between the consecutive generations. Twenty different pattern matching rules have been defined in this software as well as Dice's coefficient, and Jaccard similarity. The obtained values from pattern matchers are then fed into synthesizer for producing sounds. Some of the synthesizer parameters include ADSR envelope, loop gain, and the musical instrument string length for defining frequency. Further information about the software can be found in [15].

An important point needed to be mentioned is that the aggregation of a CA rule and a pattern matching rule on each of the synthesizer elements does not produce a single note but a collection of notes which are referred to as *voices* throughout this paper.

Studying the musical behaviour derived from one-dimensional (1D) CA does not require the investigation of the 256 rules' behaviours. The rule space can be reduced to 88 fundamental behaviours [16] by applying conjugate, reflection, and both transformations together [3], since they lead to rule sets with inherently equivalent behaviour (The interested reader is referred to [1] for formulation of conjugate and reflection transformations and how they are applied to find equivalent CA rules). The 88 1D CA rule behaviours, 7 defined synthesizer parameters, together with 20 pattern matching rules, expand the number of voices to $88 * 20^7$.

The musical output of LPM has been investigated through two different approaches: the decision of a human subject and the plots of pattern matcher outputs. The values of

20 pattern matchers for 10000 generations of 88 CA rules have been extracted. The initial seeds for CA are selected randomly. These values change the parameters of the synthesizers within the acceptable defined ranges for parameters.

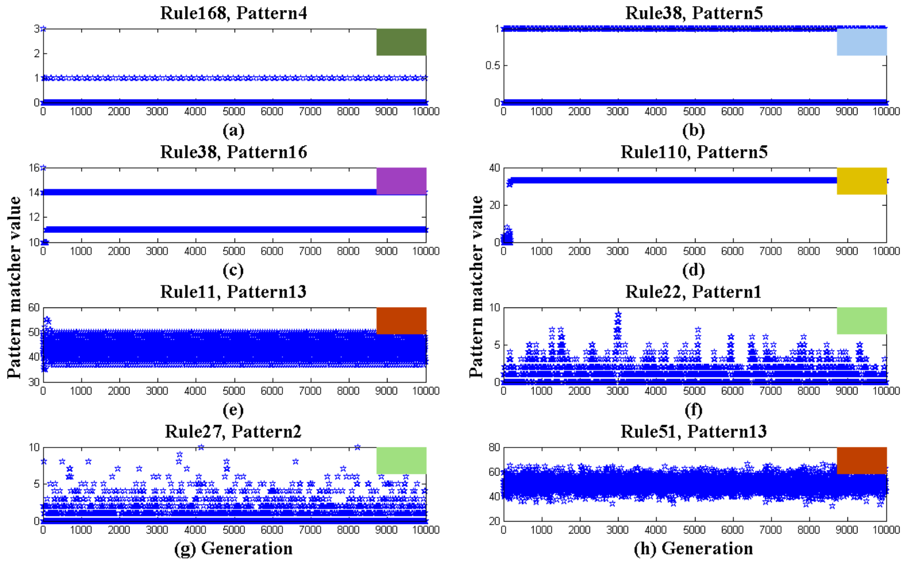


Fig. 1. Pattern matcher values over CA evolution of 10000 generations for (a) rule 168 : pattern 4, (b) rule 38 : pattern 5, (c) rule 38 : pattern 16, (d) rule 110 : pattern 5, (e) rule 11 : pattern 13, (f) rule 22 : pattern 1, (g) rule 27: pattern 2, (h) rule 51: pattern 13.

Three sorts of behaviour have been observed in the experiments with CA and our pattern matchers in LPM (It needs to be stated that this classification into three behaviours is independent from Wolfram classes and are based on the conducted experiment). In the first group, the evolution converges to a homogeneous state, in which no differences can be measured through consecutive generations, resulting in a uniform sound. Amongst this behaviour are the first Wolfram class and first subgroup of the second class. The second behaviour consists of an ordered oscillation between two or more values. Some examples include: rules 6, 178, 26, 73 from the third subgroup of the second Wolfram class; rules 22, 30 from the third Wolfram class; and rule 110 from fourth Wolfram group. The third pattern of behaviour is observed as a disordered fluctuation between a large number of values. Among these, rule numbers 14, 28, 51 from classes 3 and 2 are notable. Figure 1 illustrates some of these behaviours for CA rule numbers 11, 22, 27, 51, 38, 110, 168 (from Wolfram classes 2, 3, 2, 2, 2, 4, 1 respectively). Subfigure (d) shows the homogeneous behaviour, while subfigures (a), (b), (c), (e), (f), (g) suggest almost an ordered oscillatory pattern within different ranges. Subfigure (h) depicts disordered fluctuation behaviour. These three obtained behaviours from our pattern matchers provide insight into the nature of the musical behaviour of our system.

3 Investigation on LPM output and Zipf's law

Zipf's law characterizes the scaling attributes of many natural effects including physics, social sciences, and language processing. Events in a dataset are ranked (descending order) according to their prevalence or importance[17]. The rank and frequency of occurrence of the elements are mapped to a logarithmic scale, where linear regression is applied to the events graph. The slope and R^2 measurements demonstrate to what extent the elements conform to Zipf's law. A linear regression slope of -1 indicates Zipf's ideal. Zipf's law can be formulated as $F \sim r^{-a}$, in which r is the statistical rank of the phenomena, F is the frequency of occurrence of the event, and a is close to one in an ideal Zipfian distribution. The frequency of occurrence of an event is inversely proportional to its rank [17]. $P(f) = \frac{1}{f^n}$ is another way to express the Zipf's law. $P(f)$ is the probability of occurrence of an event with rank f. In case of $n = 1$(Zipf's ideal), the phenomenon is known as pink noise. The cases of $n = 0$ and $n = 2$ are called white and brown noises, respectively [17].

Voss and Clarke [18] have observed that the spectral density of audio is 1/f like and is inversely proportional to its frequency. They devised an algorithm which used white, pink, and brown noise sources for composing music. The results show that pink noise is more musically pleasing due to its self-similarity characteristics, the white noises are too random, and the brown noises are too correlated [17] producing a monotonous sound.

In the musical domain, Zipf's metrics are obtained by enumerating the different musical events' frequency of occurrence and plotting them in a log-log scale versus their rankings. The slope of Zipf's distribution differs from $-\infty$ to 0. The R-squared value is between 0 and 1. The decreasing of the slope to minus infinity speaks of the level of monotonicity increased. Various publications explore the utilization of Zipf's law in musical data analysis and composition. Previous experiments [17, 19] show its successful application in capturing significant essence from musical contents. In [17] the Zipf's metrics consist of simple and fractal metrics. The simple metrics include seventeen features of the music as well as the ranked frequency distributions of pitch, and chromatic tone. Fractal metrics gives a measurement of the self-similarity of the distribution. These metrics were later used to train neural networks to classify musical styles and composers, with an average success rate of over ninety percent; demonstrating that Zipf's metrics extract useful information from music in addition to determining the aesthetical characteristics of music pieces [17].

Here, the values obtained from the pattern matching rules for the 88 CA are used to study the behaviour of LPM in terms of Zipfian distribution. These measurements are ranked in compliance with their redundancy. After applying linear regression on the rank and frequency of occurrence of the set values, the slope and R-squared measurements are obtained.

The values of the 20 pattern matching rules have been extracted from over 10000 generations of CA evolution. Figure 2 shows linear regression lines fitted to Zipfian data distribution of LPM outputs for specific CA rule numbers. Table 1 depicts the slopes obtained from applying Zipf's law on the dataset after the five hundredth generation to ten thousandth generations for these rules. The reason for this time delay is to let the CA to reach stability in its evolution after the initial state. The coloured cells in the figures are obtained after a scrutinized comparison between the output graphs and Zipfian data. Most of the parameters attained in the table for Zipfian slopes are as expected according to the output graphs from the previous section. The brown cells indicate the conditions where the distribution follows Zipf's law. The yellow cells show the distributions with minus infinity slopes. The dark green cells illustrate monotonous outputs. The light green cells demonstrate the situations in which the author would expect Zipfian ideal parameters, despite the obtained parameters being far from ideal. The light blue cells depict the cases where the parameters have nearly ideal Zipfian distribution; however, the graphs in the previous section (Fig.1) show the contrary. In this case the Zipf's distribution is not sufficient for showing the musicality of the data distribution, due to the limited diversity of events. The remaining cells (dark purple) show slopes near to zero, with the graphs in Fig. 1 suggesting tedious outputs.

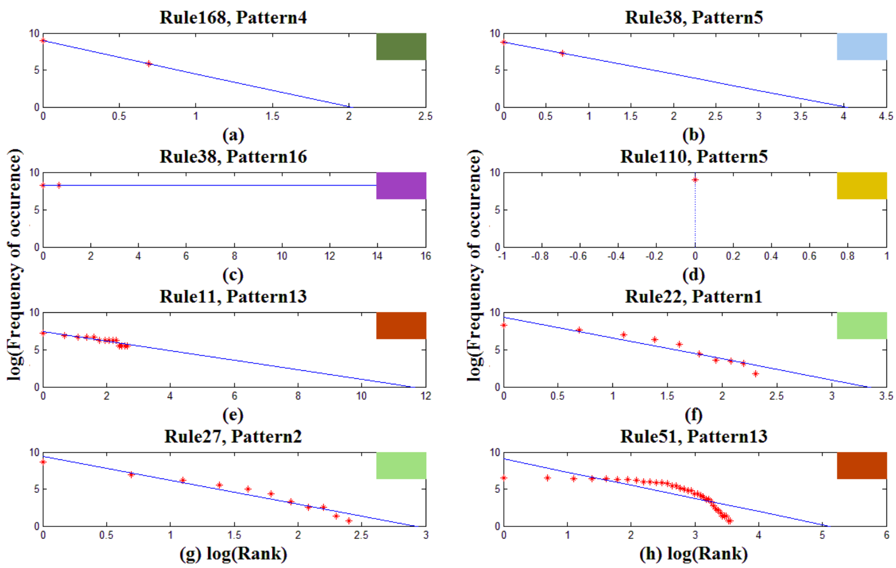


Fig. 2. Zipfian distributions for (a) rule 168: pattern 4, (b) rule 38: pattern 5, (c) rule 38: pattern 16, (d) rule 110: pattern 5, (e) rule 11: pattern 13, (f) rule 22: pattern 1, (g) rule 27: pattern 2, (h) rule 51: pattern 13.

The remainder of this section presents an analysis of the colour coding given for table 1. A confusion matrix allows good visualisations over the performance of a classifying algorithm. It shows how the predicted and actual classes overlap each other due to performance of the algorithm [20].

The target classes are obtained by human decision. The human labelling was performed according to the studies on the graphs as in figure 1 and also by interactive auditory tests from LPM outputs. The predicted classes are gained by labelling them in terms of Zipf's law musicality; the slopes which are mostly near the Zipf's ideal (-1) are categorized as being musical. Here, it has been decided that Zipf slopes between -2.1 and -0.6 are expected to be musical.

The confusion matrix in table 2 is defined as follows, with the related colour coding and observed behaviour in Wolfram classes and Li and Packard subclasses shown in table 3:

- TP (True Positive) items are expected to be musical and their Zipf slopes show they have musical attributes (brown cells).
- FP (False Positive) indicate items not expected to be musical but their Zipf's graphs show they have musical attributes (light blue cells).
- TN (True Negative) are items not labelled as musical and are correctly classified outside of the musical group (yellow, dark green, and purple cells).
- FN (False Negative) are expected to have a musical output, however, Zipf's metric is not showing that (light green cells).

Table 1. Zipfian slopes for some examples, the first column on left depicts CA rules and the first row stand for pattern matching rules (please refer to table 3 for colour coding).

	M1	M2	M3	M4	M5	M6	M7	M8	M9
168	-2.53	-4.43	-3.12	-2.87	-2.53	-2.53	-Inf	-4.43	-0.61
11	-2.33	-1.21	-2.52	-2.78	-1.94	-2.19	-1.96	-2.00	-0.75
27	-1.54	-3.37	-1.57	-3.15	-1.60	-3.30	-1.61	-3.13	-1.99
38	-3e-4	-1.18	-0.11	-1.74	-1.22	-0.93	-1.36	-1.19	-3e-4
51	-3.30	-3.18	-3.19	-2.84	-3.28	-3.28	-3.34	-3.00	-1.95
22	-2.60	-2.86	-2.96	-3.07	-3.26	-2.94	-3.44	-3.29	-1.64
110	-3e-4	-3e-4	-3e-4	-3e-4	-Inf	-Inf	-Inf	-Inf	-3e-4

Table 2. Confusion Matrix.

Confusion Matrix	Musicality = True (Positive)	Musicality = False (Negative)
Zipf = True	TP (254 items)	FP (30 items)
Zipf = False	FN (207 items)	TN (1269 items)

Table 3. Colour coding interpretations in terms of confusion matrix.

Colour Coding	Colour interpretation	Wolfram CA Classes	Number of occurrences
	FN	2, 3	207
	FP	1, 2-2, 3	30
	TN	1,2,3,4	1031
	TN	1, 2-2, 3	71
	TN	1, 2-2, 2-3, 3, 4	167
	TP	2-2, 3	254

The Accuracy ($\frac{TP+TN}{TP+FP+TN+FN}$), sensitivity ($\frac{TP}{TP+FN}$), and specificity ($\frac{TN}{TN+FP}$) [20] of the classifier are calculated as 87, 55, 98% respectively. High accuracy suggests Zipf classifier is likely to predict the musical and non-musical samples correctly. Middle range rate for sensitivity indicates its average ability for identifying musical elements. The specificity shows its success rate in correctly excluding non-musical individuals.

4 Applying Zipf's Law on a Crafted Sequence of Voices

Regarding the outputs of previous stage, it was assumed that juxtaposing a collection of voices yield Zipfian slopes. On this account a 100 sequences, consisting of a random selection of voices, were defined in Matlab. As a first stage, the length of voices were randomly chosen amongst the total number of CA generations. The obtained Zipfian slopes range from -3.06 to -1.44 with 0.81 and 0.66 as their respective r-squared values. Figure 3 illustrate the Zipfian distributions for these cases and for a third case with an ideal fit (graph c).

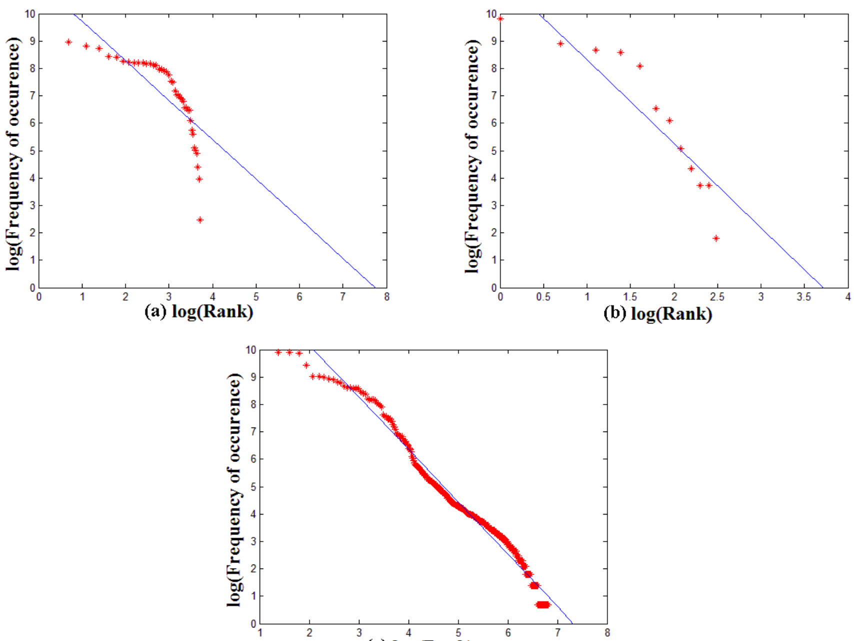


Fig. 3. Examples of Zipfian distributions for the first experiment with voices of random length up to 10000, with obtained results of (a) $s = -1.44$, r^2= 0.66, (b) $s = -3.06$, r^2= 0.81, (c) $s = -1.91$, r^2= 0.96.

In the consequent stages the random selection of whole generations was performed with a smaller sample size (max 20). By decreasing the length of voices, the monotonicity of those characteristic elements will be lowered, which can produce acceptable range of numbers for the Zipfian distribution. In this last experiment the min and max achieved Zipfian slopes are -1.56 and -0.89; and the r-squared values are 0.92 and 0.91. Some examples are depicted in figure 4, with the case of maximum r-squared value in graph c. The results imply that crafting voices with the appropriate characteristic beside each other give musically pleasant outputs in LPM. Tailoring the best musical combinations from the possible space of emerging voices is a task which is appropriate for the use of evolutionary algorithms.

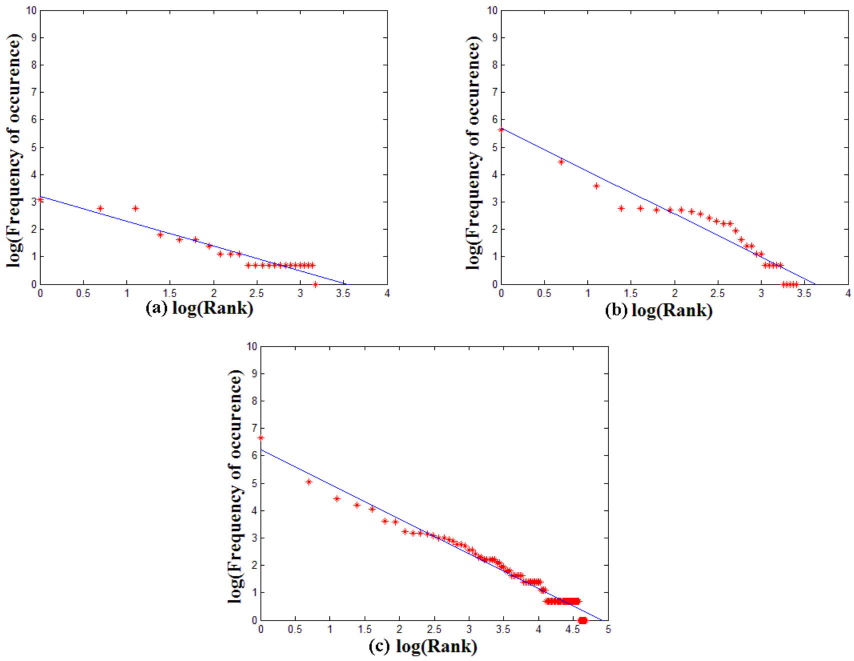


Fig. 4. Examples of Zipfian distributions for the second experiment with voices of random length up to 20, with obtained results of (a) $s = -0.89$, r^2= 0.91, (b) $s = -1.56$, r^2= 0.96, (c) $s = -1.36$, r^2= 0.97.

If Pattern Matching over CA defines musical structures, then the application of Genetic Algorithms (GA) could tailor the musical sequences to make them aesthetically acceptable to audience and conform to musical rules. The search to find optimal solutions is guided by assigning higher fitness to competent individuals. Applying Genetic Algorithms on search and optimization of musical sequences has special requirements. For example, defining the search space; specifying the musical knowledge and rule representation; and the choice of an appropriate fitness function [21]. Since there are infinite possibilities for producing music; it is necessary to define suitable constraints to limit the search space. An idea could be to keep those melodies which conform to ideal Zipfian slopes. The second issue would be to clarify how music composition evolution is effected by GA progression by means of rule representation.

5 Conclusion and Future Directions

In this paper, pattern matching over CA evolution was employed as a controller for the parameters of a synthesizer. The outputs of the software have been explored through graphs and auditory tests. The output distributions have been investigated regarding their compliance with Zipf's law. The results are categorized according to the expectations from studying the behaviours. In the third experiment collections of voices are sequenced; some with pleasing Zipfian slopes. Some of the CA and rules do not contribute to musical outputs by themselves, but, experiments with crafted pieces have shown that the proper combination of those elements can give in acceptable musical results.

The measured Zipfian slopes characterize the global features [17] of the produced music. Attention was kept on one dimension of the synthesizer (frequency) and on global measurement of aesthetics throughout this study, for simplicity. Although, Zipf's law can be considered a good approach for investigating the pleasantness of the output melody, there are other approaches which can be taken into account in our research. Other pattern matching rules are under investigation to assist us in attaining some other aspects of visual aesthetics from cellular automata evolution [22]. By extending the dimension of the produced sound, our future direction will investigate different possibilities of ordering those elements horizontally and vertically for yielding melody and harmony regarding aesthetical measurements.

In the next step the author is to investigate the feasibility of different approaches for crafting an appropriate fitness function and whether it will be automatic or interactive [9]. In interactive mode, users train evaluators which are applied as fitness functions in the evolutionary system. Fitness bottleneck is a challenge in this type of evaluation [9]. There are various automatic fitness functions for the application of GA in music composition given a set of musical training sets. These include a model based on Zipf's law [19], Markov model, or artificial neural networks. Fitness functions based on the distance from an original melody [23] are other approaches; as are using constraint satisfaction with genotypes conforming to music theory [24], and the weighted sum of statistical compositional features [25].

References

1. Burks, A.W.: Von Neumann's Self-Reproducing Automata, (1970).
2. Burks, A.: Essays on Cellular Automata. University of Illinois Press (1970).
3. Wolfram, S.: A New Kind of Science. (2002).
4. Burraston, D., Edmonds, E.: Cellular automata in generative electronic music and sonic art: a historical and technical review. Digit. Creat. 16, 165–185 (2005).
5. Beyls, P.: The Musical Universe of Cellular Automata, (1988).
6. Cellular automaton, from wikipedia, http://en.wikipedia.org/wiki/Cellular_automaton.
7. Burraston, D., Edmonds, E., Livingstone, D., Miranda, E.R.: Cellular Automata in MIDI based Computer Music. Proc. Int. Comput. Music Conf. 4, 71–78 (2004).
8. Li, W., Norman, P.: The Structure of the Elementary Cellular Automata Rule Space. Complex Syst. 4, 281–297 (1990).
9. Fernández, J.D., Vico, F.: AI methods in algorithmic composition: A comprehensive survey. J. Artif. Intell. Res. 48, 513–582 (2013).

10. Miranda, E.R.: On the evolution of music in a society of self-taught digital creatures. Digit. Creat. 14, 29–42 (2003).
11. Miranda, E.R.: Sounds of artificial life. In: Proceedings of the 4th Conference on Creativity & and Cognition, ACM. pp. 173–177 (2002).
12. Cook, P.R., Scavone, G.P.: The Synthesis ToolKit in C++ (STK), https://ccrma.stanford.edu/software/stk/.
13. Woods, N.V.: Liquid Brain Music: Phase II. (2009), Department of Computer Science, University of Hull.
14. Turner, C.: Liquid Brain Music, (2008) , Department of Computer Science, University of Hull.
15. Davis, D.N.: Computer and Artificial Music: Liquid Persian Music, http://www2.dcs.hull.ac.uk/NEAT/dnd/music/lpm.html.
16. Jen, E.: Global properties of cellular automata. J. Stat. Phys. 43, 219–242 (1986).
17. Manaris, B., Romero, J., Machado, P., Krehbiel, D., Hirzel, T., Pharr, W., Davis, R.B.: Zipf's Law, Music Classification, and Aesthetics. Comput. Music J. 29, 55–69 (2005).
18. Voss, R.F., Clarke, J.: "1/f noise" in music: Music from 1/f noise, http://physics.bu.edu/~redner/542/refs/voss-clarke.pdf, (1978).
19. Lo, M.Y.: Evolving Cellular Automata for Music Composition with Trainable Fitness Functions. Electron. Eng. (2012).
20. Witten, I. H., Frank, E. & Hall, M.A. (2011) Data mining : practical machine learning tools and techniques, 3rd ed., Amsterdam; Boston, MA: Morgan Kaufman.
21. Burton, a R., Vladimirova, T.: Generation of musical sequences with genetic techniques. Comput. Music J. 23, 59–73 (1999).
22. Kirke, a., Miranda, E.: Capturing the aesthetic: Radial mappings for cellular automata music. J. ITC Sangeet Res. Acad. 21, 15–23 (2007).
23. Horner, A., Goldberg, D.E.: Genetic algorithms and computer-assisted music composition. In: In Proceedings of the International Conference on Genetic Algorithms (1991).
24. McIntyre, R. a.: Bach in a box: the evolution of four part Baroque harmony using the genetic algorithm. Proc. First IEEE Conf. Evol. Comput. IEEE World Congr. Comput. Intell. 852–857 (1994).
25. Horner, A., Ayers, L.: Harmonization of musical progressions with genetic algorithms. In: In Proceedings of the International Computer Music Conference. pp. 483–484 (1995).

Part III

Extended Abstracts

Unconventional Reservoir Computers: Exploiting Materials to Perform Computation

Matthew Dale

Department of Computer Science,
University of York
md596@york.ac.uk

Abstract. Reservoir Computing (RC) has been identified as a powerful technique for training Recurrent Neural Networks (RNN) and an adaptable computational model for many unconventional systems. The reservoir model can be defined by its two structural partitions; a "reservoir", a dynamical system that creates a complex projection of the input, and a linear readout layer trained upon the output reservoir states. The paradigm is highly competitive to many state-of-art machine learning techniques and provides an excellent theoretical grounding in training and understanding desirable characteristics of self-contained physical systems. In this work, we postulate that *in-materio* systems - comprising of complex nanoscale substrates - could represent ideal, re-configurable physical reservoirs. Moreover, we argue that these substrate-based computers could benefit from optimisation/exploration techniques, configuring materials into highly susceptible states that boost the reservoir training mechanism using techniques found in Evolution-*in-Materio*.

Keywords: Material Computation, Evolution-*in-Materio*, Reservoir Computing, Unconventional Computing, *edge of chaos*

1 Introduction

Unconventional computing is a rising research area in which alternative computing systems, often based on the complex dynamics of non-linear systems, are developed to create new algorithmic paradigms and novel physical computing substrates. In principle, unconventional computers separate themselves by trying to fully exploit the physics and chemistry of the underlying substrate to perform computation. Moreover, such systems are created with the intention of capitilising on inherent physical laws and interactions, requiring little or no intervening layers of abstraction [24]. In comparison, conventional/classic computers are designed to be substrate-independent, where a symbolic "machine" is highly constrained into physical hardware.

Material/substrate-based computers are conceptual machines in which some computational process, or physical mechanism, maybe extracted from a behaviourally "rich" dynamical system. In essence, information processing can be exploited from what the substrate does naturally, for example, how the system reacts and dynamically adapts to some input stimulus [23]. Informally speaking, this can be viewed as “kicking” the dynamical system and observing its behaviour to some given stimulus, where the method of perturbation and observation may vary in type, e.g. electrical, mechanical, optical, etc. From this, we then speculate as to whether some exploitable process, or *computation*, can be extracted and whether the system can be trained to consistently exploit this process.

Exploiting computation directly from materials offers many potential advantages over classical systems. As a paradigm, it potentially offers vast amounts of computational power by capitalising on the massive parallelism and natural constraints of these rich systems where such properties could provide solutions“for free”, or at least computationally cheaper [23].

2 Reservoir Computing

Reservoir Computing (RC) was first introduced as competitive method to training Artificial Recurrent Neural Networks (RNNs). Classical RNNs typically succumb to inefficiencies such as; suffering from bifurcations, computationally expensive update cycles due to small parameter changes and generally difficult to train using backpropagation and gradient-descent based methods. Reservoir Computing overcomes many of these difficulties by reducing the training complexity, training only one set of connection weights instead of all weights recursively. To understand how the system functions it has to be conceptually viewed in a different manner. The “reservoir”, which is here an untrained, randomly created network of recurrently connected neurons, represents a dynamical system possessing *kernel*-like properties (see Kernel Methods [21]). By this, we infer that it produces a high-dimensional expansion/projection of the input into reservoir states. In essence, turning the reservoir into a non-linear filter, acting as a pre-processing mechanism which can be exploited by a subsequent (trained) linear readout layer. This simplifies the training of the readout and helps avoid many of the aforementioned inefficiencies, but, it also implies that the reservoir should be, by itself, sufficiently rich to create an exploitable projection. This is typically referred to as the kernel ”quality” [12].

Using an RNN as a computational model of a complex non-linear system has its advantages as RNNs are assumed to be universal approximators of dynamical systems [14]. Before the unifying name of “Reservoir Computing”, it was divided into many “flavours”. The *Echo State Network* (ESN) [6] approach was conceived from the observation that given an RNN possessed certain key properties one only needed to train a linear readout to produce a high-performance system. The two main properties identified were to possess both a fading memory - referred to as the *echo state property* - and an adequate dynamic memory capacity for the given task. ESNs are somewhat abstract from biological neural networks and generally consist of rather simplistic neuron models such as sigmoidal or leaky-integrator neurons, but this simplicity also makes them easy to create, simulate and manipulate.

The *Liquid State Machine* (LSM) [17] model came forth as a method for defining the computational properties and power of neural microcircuits, described as; “an alternative to Turing and attractor-based models in dynamical systems” [16]. Maass (its creator) argues that such classical models cannot handle real-time data streams (of continuous-time) such as spike trains in spiking neuron-based RNNs. This model is viewed as much more biologically plausible compared to the ESN model.

2.1 Criticality Hypothesis

In the literature, it has been hypothesised and demonstrated that there may be a critical state in which a system can exhibit maximal computational power, a point where maximal complexity can be acquired. This state is referred to as a region near (or at) the “edge of chaos” [2, 10]. The edge of chaos is described as the transitional border between ordered and chaotic behaviour, where perturbations to the system quickly fade and settle, or significantly alter predictability and long-term stability.

This theory is discussed as it has been observed to play some useful role in quantifying reservoir performance and optimisation [28]. Moreover, it could have direct connotations to material computation whereby a material can only exhibit “richness”, and therefore be exploitable, by operating close to or within this region.

2.2 Unconventional Systems

In what is referred to here as the second epoch of Reservoir Computing - the first being confined to simulations of recurrent neural networks - there has been a marked shift towards hardware-based reservoir systems. These new systems exploit the structure and methodology of the RC paradigm, separating the physical system into to two constituent parts; the “reservoir” (the dynamical system) and the interface mechanism (the input and readout).

An early example of this change to hardware was first demonstrated using the LSM model, where the “liquid” (reservoir) was physically represented by a bucket of water [4]. In this experiment, by applying an external stimulus one could identify dynamic states encoded in the ripples on the surface of the water. Then, using these states as the basis of a trained system they could exploit the resulting wave interactions to solve logic gate and classification problems.

Jones *et al.* [7] later observed that LSM properties could also be found in models of gene regulation networks of E.coli. The same LSM mechanisms were also identified much earlier in biological systems, described as a possible process used by mammalian brains in speech recognition [3]. The LSM model has also been further verified as a process by which trained linear readouts can extract information from non-linear projections created both by a cats primary visual cortex [19] and neural tissue from the primary auditory cortex [8].

More recently, a wide variety of new and interesting unconventional systems have been proposed. The first (in no particular order) is an example based on non-linear systems that incorporate delayed feedback, creating pseudo-networks in optoelectronic and photonic hardware [1, 11]. The second, encompasses small photonic networks created on-chip using either active [26] or passive components [27]. The last, demonstrates realised networks of memristive (resistance which depends upon input history) components such as; electro-chemical, self-assembling networks of silver nanowires [22, 25] and randomised memristor networks [9].

3 Evolving Configurations of Matter

A key vision to the proposed work is that one could evolve high-performance reservoirs given the desirable reservoir properties as objectives (e.g. criticality, memory capacity and kernel "quality"). Evolution-*in-Materio* (EIM) [18] has been identified as unique and novel concept to evolving desirable configurations in matter. EIM has shown that hybrid digital/analogue systems can directly manipulate and exploit the physical characteristics of materials to create bespoke computing platforms [5, 15]. An advantage (but should also be viewed with caution) of this methodology is that the training process does not require a full understanding of the computational mechanisms that it exploits, treating the material as a self-encapsulated *black-box.*

It is proposed here that combining both EIM and RC concepts could solve many of the discrepancies within each individual field, such as a lack of optimisation (in physical systems) in RC, training limitations and modeling in EIM, and dealing with both scalability and task complexity in each paradigm.

4 Future Work

This small section highlights some possible avenues and investigations for future work.

At present, searching for, or, designing ideal materials as bespoke computational machines is very problematic and time consuming. We propose that knowing/understanding unique characteristics that make good reservoirs could help guide substrate selection/design.

In Hardware-based Reservoir Computing there has been no discussion of optimisation by evolution in a physical system. Reservoir optimisation (in software) has been widely investigated and has proven to be a significant field of interest in classical reservoirs [14]. We argue that configuration through evolution in intractable complex systems could be, in some cases, one of very few viable methods to both explore the enormous state space and exploit properties hitherto unknown.

The criticality (edge of chaos) proposition has only recently been considered in EIM [20], more substantial investigations are still required. For many materials this may be a significant feature, where to be configurable and effective requires some highly-susceptible and critical state - a feature also said to be directly coupled with a high kernel "quality" [13]. When experimenting with this hypothesis further questions should also be asked such as, what general implications does this have on a physical system? how do we measure and define this region quantitatively in a physical system? can we drive or influence a material to operate, self-organise and self-sustain (with minimal interaction) near/within this critical landscape? and, what are the benefits of such a regime on task-generalisation, i.e. does it make a substrate more task independent?

Real-Time control problems are proposed as computational substrates may be uniquely adapted to such domains, providing low-power, massively parallel, high-speed standalone alternatives. We propose that a hybrid (analogue/digital) controller based on architectural layers that exploit both paradigms could provide an improved robust solution, moving us closer to fully embodied systems.

Acknowledgments. Matthew Dale is funded by a Defence Science and Technology Laboratory (DSTL) PhD studentship.

References

1. L. Appeltant, M. C. Soriano, G. Van der Sande, J. Danckaert, S. Massar, J. Dambre, B. Schrauwen, C. R. Mirasso, and I. Fischer. Information processing using a single dynamical node as complex system. *Nature communications*, 2:468, 2011.
2. N. Bertschinger and T. Natschläger. Real-time computation at the edge of chaos in recurrent neural networks. *Neural computation*, 16(7):1413–1436, 2004.
3. P. F. Dominey. Complex sensory-motor sequence learning based on recurrent state representation and reinforcement learning. *Biological cybernetics*, 73(3):265–274, 1995.
4. C. Fernando and S. Sojakka. Pattern recognition in a bucket. In *Advances in artificial life*, pages 588–597. Springer, 2003.
5. S. Harding and J. F. Miller. Evolution in materio: A tone discriminator in liquid crystal. In *Evolutionary Computation, 2004. CEC2004. Congress on*, volume 2, pages 1800–1807. IEEE, 2004.
6. H. Jaeger. The "echo state" approach to analysing and training recurrent neural networks-with an erratum note. *Bonn, Germany: German National Research Center for Information Technology GMD Technical Report*, 148:34, 2001.
7. B. Jones, D. Stekel, J. Rowe, and C. Fernando. Is there a liquid state machine in the bacterium escherichia coli? In *Artificial Life, 2007. ALIFE'07. IEEE Symposium on*, pages 187–191. IEEE, 2007.
8. S. Klampfl, S.V. David, P. Yin, S.A. Shamma, and W. Maass. Integration of stimulus history in information conveyed by neurons in primary auditory cortex in response to tone sequences. In *39th Annual Conference of the Society for Neuroscience, Program*, volume 163, 2009.
9. Z. Konkoli and G. Wendin. On information processing with networks of nanoscale switching elements. *International Journal of Unconventional Computing*, 10(5-6):405–428, 2014.
10. C. G. Langton. Computation at the edge of chaos: phase transitions and emergent computation. *Physica D: Nonlinear Phenomena*, 42(1):12–37, 1990.
11. L. Larger, M. C. Soriano, D. Brunner, L. Appeltant, J. M. Gutiérrez, L. Pesquera, C. R. Mirasso, and I. Fischer. Photonic information processing beyond turing: an optoelectronic implementation of reservoir computing. *Optics express*, 20(3):3241–3249, 2012.

12. R. Legenstein and W. Maass. Edge of chaos and prediction of computational performance for neural circuit models. *Neural Networks*, 20(3):323–334, 2007.
13. R. Legenstein and W. Maass. What makes a dynamical system computationally powerful. *New directions in statistical signal processing: From systems to brain*, pages 127–154, 2007.
14. M. Lukoševičius and H. Jaeger. Reservoir computing approaches to recurrent neural network training. *Computer Science Review*, 3(3):127–149, 2009.
15. O. R. Lykkebø, S. Harding, G. Tufte, and J. F. Miller. Mecobo: A hardware and software platform for in materio evolution. In *Unconventional Computation and Natural Computation*, pages 267–279. Springer, 2014.
16. W. Maass. Liquid state machines: motivation, theory, and applications. *Computability in context: computation and logic in the real world*, pages 275–296, 2010.
17. W. Maass, T. Natschläger, and H. Markram. Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural computation*, 14(11):2531–2560, 2002.
18. J. F. Miller and K. Downing. Evolution in materio: Looking beyond the silicon box. In *Evolvable Hardware, 2002. Proceedings. NASA/DoD Conference on*, pages 167–176. IEEE, 2002.
19. D. Nikolić, S. Haeusler, W. Singer, and W. Maass. Temporal dynamics of information content carried by neurons in the primary visual cortex. In *Advances in neural information processing systems*, pages 1041–1048, 2006.
20. O. R. Lykkeb S. Nichele, D. Laketic and G. Tufte. Is there chaos in blobs of carbon nanotubes used to perform computation? In *FUTURE COMPUTING 2015, The Seventh International Conference on Future Computational Technologies and Applications*, pages p. 12–17, 2015.
21. J. Shawe-Taylor and N. Cristianini. *Kernel methods for pattern analysis*. Cambridge university press, 2004.
22. H. O. Sillin, R. Aguilera, H. Shieh, A. V. Avizienis, M. Aono, A. Z. Stieg, and J. K. Gimzewski. A theoretical and experimental study of neuromorphic atomic switch networks for reservoir computing. *Nanotechnology*, 24(38):384004, 2013.
23. S. Stepney. The neglected pillar of material computation. *Physica D: Nonlinear Phenomena*, 237(9):1157–1164, 2008.
24. S. Stepney, S. L. Braunstein, J. A. Clark, A. Tyrrell, A. Adamatzky, R. E. Smith, T. Addis, C. Johnson, J. Timmis, and P. Welch. Journeys in non-classical computation i: A grand challenge for computing research. *International Journal of Parallel, Emergent and Distributed Systems*, 20(1):5–19, 2005.
25. A. Z. Stieg, A. V. Avizienis, H. O. Sillin, R. Aguilera, H. Shieh, C. Martin-Olmos, E. J. Sandouk, M. Aono, and J. K. Gimzewski. Self-organization and emergence of dynamical structures in neuromorphic atomic switch networks. In *Memristor Networks*, pages 173–209. Springer, 2014.
26. K. Vandoorne, J. Dambre, D. Verstraeten, B. Schrauwen, and P. Bienstman. Parallel reservoir computing using optical amplifiers. *Neural Networks, IEEE Transactions on*, 22(9):1469–1481, 2011.
27. K. Vandoorne, P. Mechet, T. Van Vaerenbergh, M. Fiers, G. Morthier, D. Verstraeten, B. Schrauwen, J. Dambre, and P. Bienstman. Experimental demonstration of reservoir computing on a silicon photonics chip. *Nature communications*, 5, 2014.
28. D. Verstraeten and B. Schrauwen. On the quantification of dynamics in reservoir computing. In *Artificial Neural Networks–ICANN 2009*, pages 985–994. Springer, 2009.

Part IV

Poster Presentation Abstracts

Sustainable Transportation Modes and Infrastructure

Hairul S. Radzuan

University of York, York, UK.
hsmr500@york.ac.uk

Abstract. With the aim of achieving a lower carbon footprint on the environment, green technology has been introduced in transportation systems in the past decade. Vehicles powered by renewable energy and smart road plans will need to be utilised. In this hypothetical project, glow-in-the-dark roads, road heating elements, intelligent road lighting system, Urban Maglev, bicycle hire system, electric vehicle charging system and Bus Rapid Transit were proposed for implementation in York, to pioneer an eco-transport city. Problems this project sought to tackle included: bad weather conditions such as flood; snowy and icy roads; road congestion; excessive carbon emission; wastage of clean renewable energy. Public transport services provided through this project will reduce car usage in York, encourage cycling and walking, provide easy access for York residents and visitors, and pioneer renewable energy usage for private vehicles. We advocated the aforementioned solutions based on a review of the literature, case studies, archival research and a survey. As York is a city full of historical structures and with a growing population, factors of both technology and local needs have been integrated to produce a pioneering city in the United Kingdom while advancing the technology growth. As a result, renewable energy is hoped to be used in a much wider context as it would help to pave a healthier and more sustainable future.

Keywords. Sustainable, transportation, Maglev, solar energy, car charger, glowing road, future transport, BRT, road heater

Combined electrochemical-photonic biosensor arrays

Jose Juan-Colas, Alison Parkin, Thomas Krauss and Steven Johnson

University of York

Abstract. Label-free photonic biosensors provide rapid, highly sensitive and specific detection of biologically and clinically relevant biomarkers [1]. Here we show an innovative technology that integrates photonic sensing with electrochemistry, which is capable of combining the complementary information revealed from the optical and electro-chemical domains and in a multiplexed format. Our technology consists of high Q-factor optical ring resonators [2] fabricated in silicon-on-insulator substrates in which the silicon device layer was doped n-type to an optimum carrier concentration of 7.51016 cm-3. The combination of the two measurement domains in a single sensor permits optical detection and quantification of electrochemical processes in situ. For example, we studied electrochemical grafting of a range of diazonium salts onto silicon and found that our technology was able to distinguish conclusively between diazonium assembling in monolayer or multilayer structures. We also exploited the high spatial resolution inherent to electrografting to demonstrate an array of electro-photonic ring resonators for multiplexed detection of molecular binding. Electrografting was used to selectively functionalise individual resonators with the array, each with a different singlestranded DNA probe molecule and show that this selectively functionalized photonic array is capable of detecting multiple DNA strands simultaneously. While we focus here on ring resonators and the detection of DNA, our approach is generic and can be further optimised for a range of photonic sensors and profiling other large molecular families, including the proteome.

Keywords: photonics electrochemistry biosensors DNA microarray ring resonator

References

1. X. Fan, I. M. White, S. I. Shopova, H. Zhu, J. D. Suter and Y. sun. Analytica Chmica Acta, 620, 8-16 (2008).
2. K. De Vos, I. Bartolozzi, E. Schacht, P. Bienstman and R. Baets. Opt. Express, 15, 7610-7615 (2007).

A Support for Model Checking Z Specifications

Maria Ulfah Siregar[1], John Derrick[2]

Department of Computer Science, The University of Sheffield, UK
[1]acp12mus@sheffield.ac.uk,[2]j.derrick@sheffield.ac.uk

Abstract. One of the deficiencies in the Z tools is that there is no support for model checking Z specifications. To build a model checker directly for a Z specification would take considerable effort due to the abstraction of the language. Therefore, alternative methods have been explored, such as translating the input of a Z specification into a language that a model checker tool accepts. This approach is quicker rather than writing a model checker for the language from scratch. As part of this work, researchers at the University of Sheffield implemented a translation tool which took a Z specification and translated it into the input for SAL model checker, a framework for combining different tools for abstraction, program analysis, theorem proving and model checking, which they called Z2SAL. We present two pieces of work to support model checking of Z specifications. Firstly, aid Z2SAL in a translation of a generic construct. We could not found any generic parameter in SAL literatures. Unfortunately, a generic construct is common to use in a Z specification. Now, we have just come through one of generic constructs, a generic constant. Our redefinition system pre-processes a Z specification by redefining all generic constants to equivalent axiomatic definitions. Our approach originates from similar behaviours between both definitions; they declare a global variable inside a Z specification. This work also includes a redefinition of an abbreviation definition, a lambda expression, a set comprehension with many declared parameters, and a new translation for function and constant. Secondly, aid Z2SAL in a translation of schema calculus. A new schema could be constructed based on other schemas using schema operators. A constructed schema is used commonly to define a more complex and huge specification of a system. We expand the schema based on other schemas' declaration and predicate part. This expanding needs normalisation and simplification.

Keywords: Axiomatic definition, generic construct, Z specification, Z2SAL, SAL.

An Introduction to Autonomous Agent Modeling and Simulation using MASON Toolkit

Stewart Aitken [A] Susan Stepney [A] Gianluca Tempesti [B]

A. Department of Computer Science B. Department of Electronics
University of York, Heslington, York, YO10 5DD, United Kingdom
Email:- sa120@york.ac.uk

Abstract. This poster describes how simple algorithms applied to two dimensional (2D) grids can model growth. Agent-based modeling and simulation (ABMS) is widely used across multiple disciplines with differing sorts of agents and nomenclature. This poster outlines the differing agents and their usage for investigating complex systems by inspection of the patterns derived from the local rules acting on the agents within the simulation. There are several different framework and libraries available to carry out these simulations. We have chosen to use MASON (Multi-Agent for Simulation of Neighbourhoods and or Networks). The first pair of experiments' zero and one was built for test and validation of the simulation and methodology we intended. To confirm that the simulation platform validated we applied a non-parametric statistical test, Mann Whitney *U* test and the Vargha Delaney ***A*** measure effect size to these two experiments cell occupancy per run. The results for the comparison of experiment zero and one was very close such that it supports the null hypothesis *p* value. Experiment two is currently under development as the next simulation using von Neumann and or other neighbourhoods for the agent to replicate into and has the ability to start with more than a single agent and distance traveled to cover 1, 2 4 and 8 cells away. This is showing some early results for the von Neumann neighbourhood but requires further development particularly to resolve contention issues of certain multiple agents. Thus we have gone through the hierarchy of entities, objects, agents and finally autonomous agents. These are visualised within the simulation and emergent properties are being observed in the multi-agent version of experiment two.

Keywords: Agent-based modeling, individual-based modeling, agents, complex systems, simulation, MASON toolkit, pattern behaviour, emergent, MVC design patterns, autonomous agent.

Coding Style in Source Code Plagiarism Detection

Olfat M. Mirza and Mike S. Joy

Computer Science Department, University of Warwick
Coventry, CV4 7AL
{O.M.Mirza,M.S.Joy}@warwick.ac.uk

Abstract. Plagiarism has become an increasing problem in higher education in recent years. A number of research papers have discussed the problem of plagiarism in terms of text and source code and the techniques to detect it in various contexts. When plagiarism relates to source code (computer programs) there is a variety of easy ways to copy others' work because the source code can be obtained from online code banks and textbooks, which makes plagiarism easy for student. Analysis of coding style is a way to detect source code plagiarism because it relates to programmer personality but does not affect the logic of a program, and can be used to differentiate between different code authors.

There is no agreed definition of coding style used in research to determine the provenance of computer code. Oman and Cook[1] used coding style to identify software theft and to prevent plagiarism of Pascal programs. Krsul and Spaffor[2] improved the accuracy of C program authorship identification by examining over 50 metrics and dividing them to 3 categories: (i) programming layout metrics, (ii) programming style metrics and (iii) programming structure metrics. Ding and Samadezadeh[3] used the techniques reported in previous literature to derive metrics suitably for use when analyzing Java source code style.

The immediate objective of this research is to identify new algorithms for deriving coding style of source code and to apply them to big datasets to evaluate the suitability as tools for plagiarism detection.

Keywords: Coding Style, Source Code Plagiarism Detection, Style Analysis.

---

[1] Oman, P.W. and Cook, C.R., 1991. A Programming Style Taxonomy. *International Journal of Computer Mathematics*, 4, pp.309–325.

[2] Krsul, I. and Spafford, E.H., 1997. Authorship analysis: Identifying the author of a program. *Computers & Security*, 16(3), pp.233-257.

[3] Ding, H. and Samadzadeh, M.H., 2004. Extraction of Java program fingerprints for software authorship identification. *Journal of Systems and Software*, 72(1), pp. 49-57.

Disaggregation of Appliances Based on Power Profile

Michal Kovalčík, Peter Fecilak, František Jakab

Department of Computers and Informatics
Technical University of Košice, Letná 9, 042 00 Košice, Slovakia
michal.kovalcik@tuke.sk, {peter.fecilak, frantisek.jakab
}@cnl.sk

Abstract. This poster deals with the disaggregation of appliances in household, and further with subsequent identification of specific appliance based on the energy profile of the household. The principal focus of research is the design and implementation of algorithms, which are to serve as the basis for the recognition of individual devices. An important element is the analysis, based on which, it is necessary to decide which disaggregation method is the most suitable for the design of algorithms and for the future improvements. Developed application of reliable and functional algorithms, embedded in simple graphical interface, ensures easy operation and understanding of the problem. As a part of the research, the experiment was carried out to verify the functionality of the various algorithms. Algorithms capable to disaggregate the appliance load profile from the overall household consumption profile were designed. To achieve the hardware functionality, the smart meter for measurement of three phase active, reactive and apparent power from the entire household supply conductors was constructed. Designed and implemented algorithms are able to disaggregate different appliances also enable their identification from the start, to the end, and yet provide information about the duration of their operation.

Keywords: Appliances Identification, Disaggregation Methods, House Power Profile, NIALM, NILM, Non-Intrusive Load Monitoring

Acknowledgment. We support research activities in Slovakia/This project is being co-financed by the European Union. Paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF.

On Estimation of the Real Size from Static Image

Ondrej Kainz, Matúš W. Horečný, František Jakab

Department of Computers and Informatics
Technical University of Košice, Letná 9, 042 00 Košice, Slovakia
ondrej.kainz@tuke.sk, horecny.matus@gmail.com,
frantisek.jakab@tuke.sk

Abstract. Traditionally, the estimation of object's real dimensions in the field of computer vision requires utilization of two cameras or calibration of a single camera in a static environment. In this poster, the novel approach for measuring the object's real dimensions from the single camera is presented. Prior to measurement itself, some specific conditions have to be met and several image processing algorithms are being utilized. The first essential condition is the information about camera's height from the reference plane. Next, two images are required in order to subtract the background while keeping the object, the first with the object of interest and the second without, both presume to be taken under the same conditions. The output image containing solely the object of interest is then processed using Canny edge detector and the selection of the object itself is done by bound box or bounding circle. Software solution is built on OpenCV library. The width and the height of the object is obtained following the created reference model based on the linear equations that were established as a result of testing. Separate tests were carried out for both, the width and the height. The final set of equations follow the basic trigonometric rules and are separate for both dimensions. Several experimental procedures with various objects being in a diverse heights proved the accuracy of the measurements to be higher than 90 percent.

Keywords: Computer Vision, Image Processing, Object Size, Size Estimation, Static Image

Acknowledgment. We support research activities in Slovakia/This project is being co-financed by the European Union. Paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF.

Dynamic Recognition of Hand Gestures

Roman Vápeník, František Jakab

Department of Computers and Informatics
Technical University of Košice, Letná 9, 042 00 Košice, Slovakia
{roman.vapenik, frantisek.jakab)@tuke.sk

Abstract. System described in this poster utilizes methods based on color histograms, while providing precise detection. Method itself computes color histograms of an object and background from initial binary mask of the hand. These histograms denote probability that particular color belongs to the hand or to the background. HSV color model is used. Once the histograms are computed the binary mask of hand is predicted for every frame of the video.
In order to simplify the solution the hand centroid is used as a hand description. Image moments proved to be a suitable solution for the computation of this centroid. These moments represent global characteristics of an image shape and provide information about geometric properties of an image. If the image itself is considered as two dimensional intensity distribution, than the moments of this image provide information such as total image area or image centroid coordinates. Position of centroid of the detected hand is computed from image moments. This centroid is tracked frame to frame and between consecutive centroids, based on this the angle is computed. Following the computed angle the symbol is generated. Overall, there are 12 possible symbols and for each symbol there is a range of 30 degrees. Symbols are generated only in the presence of hand movement. Once the sequences of observed symbols are obtained, there is a need to recognize which gesture was performed.
Hidden Markov Models (HMMs) were selected as the most suitable approach, for every gesture to be recognized there is one HMM. Parameters of these models are estimated with Baum-Welch method. After the gesture is completed, it is being recognized through trained HMMs. For every model that represents particular gesture, a probability based on forward method is computed. Subsequently the model with the highest probability is selected as performed gesture.

Keywords: Gesture recognition, Hand centroid, Hidden Markov Models.

Acknowledgment. We support research activities in Slovakia/This project is being co-financed by the European Union. Paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF.

High-level architecture framework for distributing files across multiple cloud storages services

Abdullah Algarni[1], Daniel Kudenko[1]
[1] Department of Computer Science Deramore Lane, University of York, Heslington, York, YO10 5GH, UK
afha500@york.ac.uk, daniel.kudenko@york.ac.uk

Abstract. Cloud Storage is a storage model for storing data remotely over the Internet. Despite the high demand for this service in which increasing dramatically every year, there are major concerns about security, performance, availability, latency and vendor lock-in. Although distribute data across multiple cloud storage using RAID technology can solve problems with availability and vender lock-in, adoption of several cloud computing solutions can introduce issues of latency time that differs from cloud edges to the end user. In order to solve these problems, I propose a framework for distribution data across multiple cloud storage composed of two machine learning algorithms: Reinforcement Learning and Supervised learning that automatically tune the data distribution across multiple cloud storage from the client site based on their access pattern. This framework aims to reduce the latency and also reduce distribution cost across cloud storage services.

Keywords

Cloud Computing, Reinforcement Learning, Supervised Learning, RAID

Work-Flow Correctness Check Using Optical Similarity

Jaroslav Lámer, František Jakab

Department of Computers and Informatics
Technical University of Košice, Letná 9, 042 00 Košice, Slovakia
{jaroslav.lamer, frantisek.jakab)@tuke.sk

Abstract. The principal aim of this project is to carry out the similarity tests between a set of image content classification algorithms, while having as a main focus the evaluation of workflow correctness based on image content. The first evaluation algorithm include the four steps of the capturing work-flow process on video stream. The initial step includes the video of work-flow process made in the correct order. The second step aims to record the same work-flow process in correct order, however from the different field of view. The goal of the third step is to record the same work-flow, this time in incorrect order. Finally, the tested algorithms was applied on all three captured video. All final results were evaluated and compared each other. Focal point in testing were algorithms - Histogram of Gradient (HOG) and Histogram of Function (HOF). These two approaches with color analysis were used in several image searching services, both produce rather poor results in all tests. The similarity between images is observable, however not usable. The correct video frames of the work-flow have a very similar outputs of similarities when compared to the frames from the incorrect work-flow. In the terms of similarity, significant difference was observed in several frames from video with correct work-flow when compared to the frames from correct work-flow captured from different angle. Based on the obtained results, it can be concluded that HOG and HOF algorithms are not applicable to recognize correctness of work-flow processes.

Due to aforementioned, the complex work-flow evaluation system based on sophisticated evaluation model was designed. The input camera data is corrected for the optical distortion, further the pose estimation is being approximated using several computer machine learning methods. After this, the main known key-points are identified and projected into the virtual linear space, where all the workflow steps or errors can be accurately identified.

Keywords: HOF, HOG, Optical Similarity, Linear Space, Machine Learning, Work-Flow Check

Acknowledgment. We support research activities in Slovakia/This project is being co-financed by the European Union. Paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF.

Exploring Storage Bottlenecks in Linux-based Embedded Systems

Russell Joyce

Real-Time Systems Research Group
Department of Computer Science
University of York, UK
`russell.joyce@york.ac.uk`

Abstract. With recent advances in non-volatile memory technologies and embedded hardware, large, high-speed persistent-storage devices can now realistically be used in embedded systems. Traditional models of storage systems, including the implementation in the Linux kernel, assume the performance of storage devices to be far slower than CPU and system memory speeds, encouraging extensive caching and buffering over direct access to storage hardware. In an embedded system, however, processing and memory resources are limited while storage hardware can still operate at full speed, causing this balance to shift, and leading to the observation of performance bottlenecks caused by the operating system rather than the speed of storage devices themselves.
In this poster, we present performance and profiling results from high-speed storage devices attached to a Linux-based embedded system, showing that the kernel's standard file I/O operations are inadequate for such a set-up, and that 'direct I/O' may be preferable for certain situations. Suggestions are made for potential areas of improvement, in order to reduce CPU load and increase maximum storage throughput.

The Sensor Organism

Naums Mogers
Dr Martin A Trefzer, University of York
Dr Dimitris Lagos, University of York

Abstract. Large-scale, distributed sensor networks are the projected weapon of choice for future pervasive computing applications such as, for example, environment monitoring, surveillance, (big) data mining and patient monitoring. However, state--of--the-art approaches face major challenges: specialized sensors are expensive and require careful calibration. Hardware sensors operating in uncertain, harsh environments eventually suffer from stress, ageing and physical damage, which leads to unforeseen effects that can render the device and the data recorded useless. Highly-tuned data processing algorithms are often not scalable and are not robust against faulty sensors delivering wrong data. Generally, systems can only adapt, if at all, in some predefined limited ways and are not capable of autonomously "inventing" new ways of adapting to unexpected changes in their internal and external environment.

This project follows a different approach to sensor network development drawing inspiration from nature, where biological systems self-organize, self-heal, sense, process information and react, emerging from evolutionary and developmental processes. Exploiting the advantages of a distributed structure of cellular systems this project perceives sensor networks as spatially distributed multicellular organisms where each sensor node represents a cell colony or a DNA bank. It is hypothesized that incorporating artificial cell development mechanisms and gene regulatory networks within each sensor node enable emerging beneficial properties of fault-tolerance, adaptivity and scalability in a sensor network. In order to investigate if this hypothesis holds in the context of a constrained distributed embedded system, the model is optimized for the Arduino microcontroller.

Resonance tuning in Soprano voices

Rebecca Vos

University of York
rebecca.vos@york.ac.uk

Abstract. Professional sopranos change the shape of their vocal tract at very high pitches to project their voices over an orchestra. The exact techniques employed and their effects on the sound are not understood. At high fundamental frequencies (f0) the harmonics of the voice are widely spaced. If they are far from resonances this wastes energy. Vocal tract resonances therefore become difficult to measure. Trained sopranos adjust their vocal tract resonances to be near harmonics, boosting their power. This increases the perceived volume and is known as Resonance Tuning (RT). There are three Research Questions considered in this work:

1. What exactly are sopranos changing when they sing high notes?

This will be answered using MRI scans of professional singers. 3D MRI scans of singers singing vowels at different pitches allow information about resonance tuning to be gathered. The 3D images allow us to see exactly how RT is achieved, and what changes singers make to the shape of their vocal tract. We will also measure the vocal tract resonances in an anechoic chamber, to gain more information about RT.

2. How do these changes affect the vocal tract resonances?

This will be investigated using a new method for measuring resonances; Injecting a noise source into the singers mouth and recording the output allows the frequencies of the vocal tract resonances to be calculated.

3. How do listeners perceive these changes?

Perceptual testing will be used to answer this question; to understand exactly how listeners perceive changes in singing we need to test exactly what changes they hear. The vowels are played to listeners, and they are asked to rate certain acoustic aspects of them.