Team learning in cybersecurity exercises

Conference Paper · June 2019

CITATIONS
0

READS
968

3 authors, including:

Kaie Maennel
Tallinn University of Technology
8 PUBLICATIONS **4** CITATIONS
SEE PROFILE

Stefan Sütterlin
Østfold University College
83 PUBLICATIONS **798** CITATIONS
SEE PROFILE

TAL TECH
DEPARTMENT OF SOFTWARE SCIENCE

PROCEEDINGS OF THE 5th INTERDISCIPLINARY CYBER RESEARCH CONFERENCE 2019

29th OF JUNE 2019
TALLINN UNIVERSITY OF TECHNOLOGY

TAL TECH
DEPARTMENT OF SOFTWARE SCIENCE

PROCEEDINGS OF THE 5th INTERDISCIPLINARY CYBER RESEARCH CONFERENCE 2019

29th OF JUNE 2019
TALLINN UNIVERSITY OF TECHNOLOGY

JUNE 2019

The 5th Interdisciplinary Cyber Research conference is organised by TalTech Centre for Digital Forensics and Cyber Security.

Editors: Dr Anna-Maria Osula, Prof Olaf Maennel
Published by: Tallinn University of Technology, Department of Software Science
Design and layout: Anu Teder

Electronically available at: https://www.taltech.ee/institutes/centre-for-digital-forensics-cyber-security/events-19/interdisciplinary-cyber-research-icr-workshop/icr2019-3/

ISBN: 978-9949-584-15-4 (pdf)

CONTENTS

INTRODUCTORY REMARKS

It is our great pleasure to welcome you in Tallinn, Estonia for the 5th Interdisciplinary Cyber Research (ICR) conference, held at the Tallinn University of Technology on the 29th of June, 2019, and organised by Tallinn University of Technology Centre for Digital Forensics and Cyber Security.

This year we celebrate a mini-jubilee of our conference as the event is taking place already for the 5th time. Within these 5 years, ICR has brought together more than 600 participants throughout the world, we have had the chance to listen to more than 125 presentations from world class researchers as well as young scholars, and published more than hundred abstracts in our annual ICR Proceedings. Furthermore, the interdisciplinary approach of ICR has really paid off as we have hosted successful panels on legal, policy, election, cyber exercises, digital forensics, Internet of Things, etc topics – underlining that cyber security is not only a technical area but involves numerous relevant research domains.

Foremost, ICR has proven itself as a connector of people: we are proud that our events bring together active researchers across different research areas, thereby allowing for the creation of new synergies and interesting research projects. For example, one of the concrete results of ICR is a joint academic article "Time of Signing in the Estonian Digital Signature Scheme", written by Tõnu Mets and Arnis Parsovs from the University of Tartu, combining both legal and technical arguments. The authors have admitted that ICR was the key factor for successfully finding a co-author.

We would also underline the long and fruitful cooperation with the Cyber Security Summer School, University of Adelaide as well as the University of Applied Sciences Ravensburg-Weingarten. In particular, University of Adelaide has throughout the years brought numerous excellent authors to our agenda from the other side of the world.

This year's programme boasts 26 presentations from all over the world. We hope that the presentations will not only be informative about "cyber"-research carried out by other disciplines than your own, but also inspiring regarding your current and future research. We continue to underline the interdisciplinary nature of "cyber" by combining different research fields into common sessions.

Most of the speakers have been hand-picked by our international Programme Committee, and the results of the Call for Abstracts are presented in this publication. This year we received a record number of abstracts, and the Programme Committee had to makes some hard choices. The selected abstracts explain the relevance of the research, outline principle research questions and expected or achieved results.

ICR is very thankful to our sponsors we have the pleasure to work with: NATO Cooperative Cyber Defence Centre of Excellence, Microsoft, Guardtime, Startup Estonia, and Saku Brewery.

Last but not the least, we would like to thank everyone involved in organising this event: the members of the Programme Committee for their efforts in reviewing the abstracts, moderators for guiding the discussions in the sessions, speakers for sharing their great ideas, conference participants for being so engaged in the debates, as well as the staff of the Tallinn University of Technology for providing excellent support.

Dr Anna-Maria Osula, TalTech/Guardtime
Prof Olaf Maennel, Tallinn University of Technology

Chairs of ICR2019
Tallinn, June 2019

SESSION 1: CYBER EXERCISES

Session moderated by Prof OLAF MAENNEL,
Tallinn University of Technology

Mr Muhammad Mudassar Yamin,
"MODELING ATTACK AND DEFENSE SCENARIOS
FOR CYBER SECURITY EXERCISES",
Norwegian University of Science and Technology

Mr Gabor Visky,
"CYBER-PHYSICAL BATTLEFIELD FOR CYBER EXERCISES",
NATO CCD COE

Mr Kieren Nicôlas Lovell,
"CYBER GAME TO CYBER EXERCISE: A NEW METHODOLOGY
FOR CYBERSECURITY SIMULATIONS",
Tallinn University of Technology

Ms Kaie Maennel,
"TEAM LEARNING IN CYBERSECURITY EXERCISES",
Tallinn University of Technology

MODELING ATTACK AND DEFENSE SCENARIOS FOR CYBER SECURITY EXERCISES

Muhammad Mudassar Yamin and Basel Katt
(Muhammad.m.yamin,Basel.katt)@ntnu.no
Norwegian University of Science and Technology

1. INTRODUCTION

Technology is evolving at a rapid rate which makes individual, ranging from security specialists to average citizens, technological skill sets obsolete in a short time. The situation of cyber-security in a technologically evolving world is not ideal. Global IT infrastructure and individual's privacy are under threat all the time. One way to tackle this problem is by providing constant training and self-learning platforms. Cyber-security exercise provides a platform for the training of individuals in cyber-security skills. But due to lack of cyber-security skills, adversarial opponents are not readily available for training exercises. The research project will focus on developing novel techniques for emulating adversarial opponents in a cyber-security exercise using a model driven methodology. The researcher plans to segregate attack and defense scenarios and create a modeling language to scientifically model such. The developed attack and defense models will be used to generate artifacts that will be executed in human v/s machine and human assisted with machine v/s human cyber-security exercises to extract empirical data for evaluation of individuals against performance matrices.

2. RESEARCH BACKGROUND

There are two types of cyber-security exercises tabletop based and operation based cyber-security exercises[1]. Tabletop based exercises focus on decision making at a managerial level while operation-based exercises focus on practical cyber-security skill development. We are currently focusing on operation based cyber-security exercises due to their practical skill development nature. In term of operation based cyber-security exercise these teams include in general[2]:

1. White team: A team that creates or generates a cyber-security exercise environment.
2. Red team: A team that attacks the cyber-security exercise environment.
3. Blue team: A team that defends the cyber-security exercise environment.

These teams are primarily involved in three main types of cyber-security exercises.

1. Cyber-attack exercise: Theses exercises are conducted to train, assess and evaluate the performance of red teams. An environment is created by a white team, in which red teams need to achieve specific objectives to compromise the exercise environment in a particular interval of time.
2. Cyber-defense exercise: Theses exercises are conducted to train, assess and evaluate the performance of blue teams. An exercise environment is created by a white team, in which blue teams needs to investigate and prevent a cyber-attack on the exercise environment by red teams under a particular interval of time
3. Cyber-attack/defense exercise: These exercises are conducted to assess and evaluate the performance of red and blue teams at the same time. A white team creates an exercise environment on which active engagement between a red and blue team occurs to attack and defend a exercise environment simultaneously.

3. RESEARCH QUESTION

We are arguing that if the role of white, red and blue team can be modeled then cyber-security exercise can be executed in an efficient and adaptable manner[3]. Therefore we are

proposing three RQ(research question) were formulated for modeling attack and defense scenarios in cyber-security exercises which are given below:

1. How can an efficient and adaptable active offensive opposition process execution be modeled against a given cyber-security exercise defense scenario?
2. How can an efficient and adaptable active defensive opposition process execution be modeled against a given cyber-security exercise attack scenario?

The findings of RQ(1) and RQ(2) will be used as a basis for the modeling of the exercise environment in which proposed RQ(1) and RQ(2) will be executed. Hence

3. How can an efficient and adaptable cyber-security exercise environment be modeled with respect to attack and defense scenarios?

4. RESEARCH METHODOLOGY

Based upon our research findings[4] we identified that automation can assist in reducing time requirements for cyber-security exercises. For this we identified that gamification can assist[5], gamification of cyber-security exercises is a recent trend in which participants are divided into teams for achieving a specific objective, like flags. The strategies that the participants apply to solve the problem, e.g. capture the flag in cyber-security exercise scenario is very difficult to model due to real time decision making of exercise participants, which makes the decision tree involved very complex. To tackle this problem, we are proposing the development of a real time cyber-security strategy game in which players will have the ability to play as an attacker or as a defender in a real time multiplayer environment. Resources are assigned to attacker and defenders based upon the scenario requirement and their actions are recorded and observed by and observer. A detailed scenario creator will be developed in which the scenario is modeled by experts. This will result in a dynamic generation of attack and defense trees, which will be generated during the real time cyber-security strategy game exercise execution. The attack and defense tree model will then be used to execute attacker and defender actions in a real cyber-security exercise environment as an active adversary against human opponents.

5. A SAMPLE SCENARIO

We developed a POC of multiplayer attack and defense game in which a scenario creator creates a scenario. The scenario has an internet facing website for defenders to defend. The website uses multiple APIs to fetch data and present it, the defender responsibility is to ensure the availability of website in case of cyber-attacks. In order to ensure the security of the website the defenders implement a WAF on the website as a security measure. The attacker tries to exploit the website and identifies that one of the API that the website is used to fetch data from is vulnerable to DoS attack, so they attack the vulnerable API to compromise the availability of the website. Created scenarios and their expected attack defense strategies are to be saved and used for future training exercises. The scenario developed using our proposed cyber-security strategy game can be seen in the figure below:

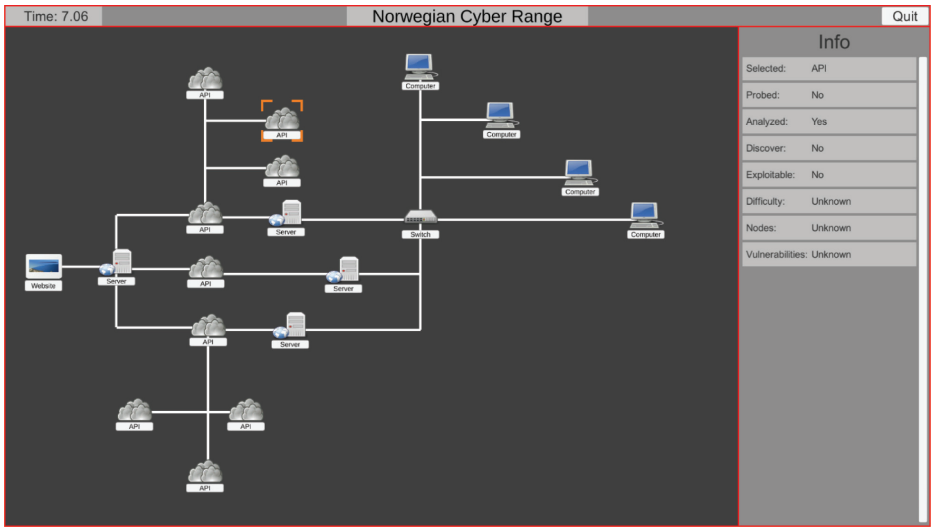
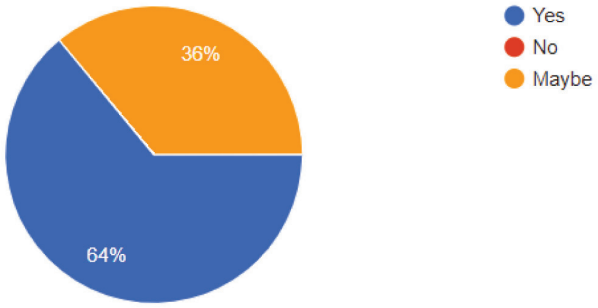


Figure 1. Sample scenario created by scenario creator

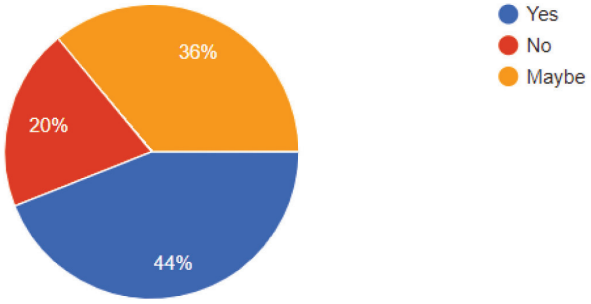
EVALUATION

We evaluated the developed game during NCSC (Norwegian Cyber Security Challenge) 2019[6]. The test subjects consisted of 25 participates who qualified the initial CTF NCSC. We collected important research data through surveys. Multiple questions were asked after the participants played the game results of which are given below:

1. Do you think playing/practicing cyber-security exercise scenarios in a simulated/ modeled game is an efficient way for conducting cyber-security exercises?



2. Do you think current game can be useful for cyber-security education?



We collected additional data as well but due to word count limitations details are omitted.

CONCLUSION

The developed game is a first step in developing autonomous attack and defense agents. Data generated from the game will be useful in developing complex decision trees that an autonomous agent need for executing red or blue team roles.

Keywords: Cyber Security, Exercises, Scenarios

REFERENCES

[1] R. Gurnani, K. Pandey, S. K. Rai, A scalable model for implementing cyber security exercises, in: Computing for Sustainable Global Development (INDIACom), 2014 International Conference on, IEEE, 2014, pp. 680–684.

[2] J. Vykopal, M. Vizv´ary, R. Oslejsek, P. Celeda, D. Tovarnak, Lessons learned from complex hands-on defence exercises in a cyber range, in: Frontiers in Education Conference (FIE), IEEE, 2017, pp. 1–8.

[3] Yamin, M. M., & Katt, B. Inefficiencies in Cyber-Security Exercises Life-Cycle: A Position Paper. AAI Fall Symposium 2018

[4] Yamin, M. M., Katt, B., Torseth, E., Gkioulos, V., & Kowalski, S. J. (2018, September). Make it and Break it: An IoT Smart Home Testbed Case Study. In Proceedings of the 2nd International Symposium on Computer Science and Intelligent Control (p. 26). ACM.

[5] Hendrix, M., Al-Sherbaz, A., & Victoria, B. (2016). Game based cyber security training: are serious games suitable for cyber security training?. International Journal of Serious Games, 3(1), 53–61.

[6] Norwegian Cyber Security Challenge – NCSC. (n.d.). Retrieved from https://www.ntnu.no/ncsc

CYBER-PHYSICAL BATTLEFIELD FOR CYBER EXERCISES

Maj. Gabor VISKY
NATO Cooperative Cyber Defence Centre of Excellence
gabor.visky@ccdcoe.org

INTRODUCTION

A cyber-physical system (CPS) is an implement intertwining physical processes, hardware, software and communication networks[1]. Examples include energy production and distribution facilities, water treatment plants, and traffic and transportation control systems. The number of security incidents affecting CPS has increased over the past years[2] as has their impact on society[3]. Operation Technology (OT) and Information Technology (IT) can now be monitored, controlled and configured remotely via a private or public network like the internet.

From an engineering and availability perspective, the controlling systems are usually well designed and tested; however, cyber-security considerations seem to be missing in the majority of cases. Prevention measures[4] and well designed and configured[5] systems can reduce the risk of cyber attacks, but the education and practice of the responsible personnel are also important since in the event of service dropout they have to handle the situation. This is challenging in the case of critical infrastructure elements such as nuclear power plants, since loss of control could be dangerous. This issue can be solved by using a special, isolated, safe and secure environment, a so-called cyber battlefield or cyber range, where methods can be tested and personnel can be trained and drilled under controlled conditions.

A cyber exercise offers a good opportunity for testing the CPS and its applied measures, checking the configurations and analysing the implemented mechanisms for cyber personnel practising defending activities, without jeopardising the real critical infrastructure. Because of this, the cyber exercise is usually conducted on a cyber battlefield which contains critical infrastructure control elements such as Programmable Logic Control (PLC), physical or virtualised hardware elements and simulated environment.

The main objective of this article is to describe the design considerations and the construction of a cyber-physical battlefield, containing several processes controlling CPSs and an environment (process) simulator, that can be used as a scenario-independent critical infrastructure element during operations-based cyber exercises for fully isolated participant teams. The platform, since it contains an environment simulator subsystem, can support complex scenarios, scoring and real-time status checking as well. This unique platform can be used by exercise participants to focus on the specialities of critical infrastructure which can be crucial for preparation and training.

CYBER-PHYSICAL BATTLEFIELD

For educational, training and system testing purposes during cyber exercises, cyber battlefields are used as a playground by cyber security staff to practise real-world incident management scenarios. Depending on the scale of the exercise, the complexity of the battlefield can become extremely high, so it must be carefully designed and run to meet the requirements[6]. A cyber-physical battlefield should provide a safe and secure infrastructure developed and managed by the Green team for participants who are at least partly isolated from the real cyber world. It should contain a monitoring, controlling and scoring system which is independent of the attacker (Red) and defender (Blue) teams, so its status cannot be influenced by the participants. In practice, cyber exercises are often visited by politicians, decision-makers and the media[7], so the system should contain a demonstra-

tion element displaying very limited and easily understandable information, such as the value of the parameter regulated by the controller of the defender team.

The construction of a universal CPS simulator platform has become a critical goal for the Technology Branch of NATO CCD COE, since this kind of compilation of various devices and technological solutions can be used for different purposes – courses, exercises, research and demonstration – with different scenarios. The main objective of the project was constructing a scenario-independent mobile tool, reusable in different cases with different PLC software but without hardware modification. It needed to have a relatively large number of independent channels and environment simulation parts, with some visual elements such as displays that show the current status of the controlled process and a firecracker that explodes when the simulated critical infrastructure is irreversibly damaged. These requirements were based on the experiences from previous exercises such as the number of the realised independent channels. The universal platform was scaled to be able to provide service for 28 participant teams. To meet this requirement, the platform contains 28 similar CPS instances and one environment simulator device.

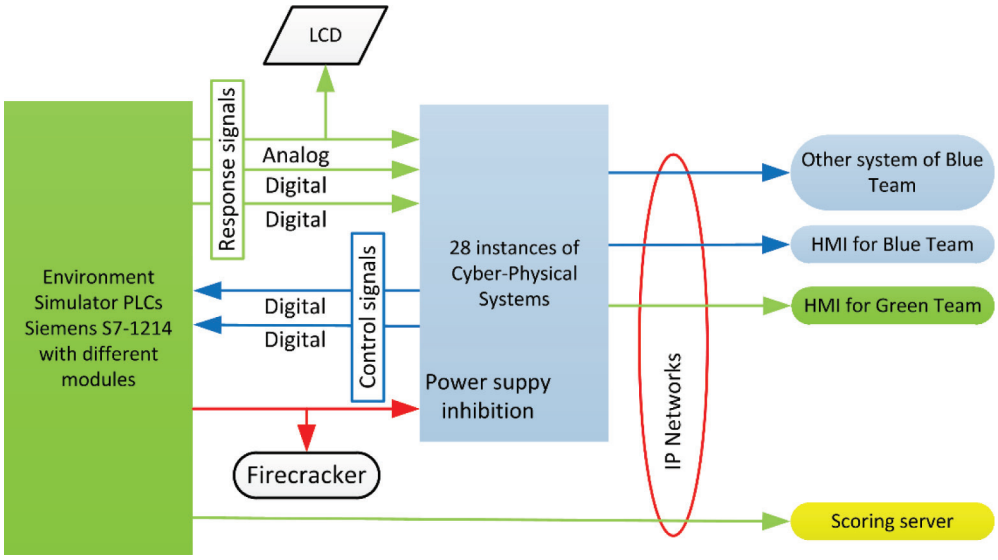


Figure 1. Architecture of CPS Platform

The architecture of the CPS Platform is shown in Figure 1. The green parts represent the environment (process) simulation. This part simulates the environment and sends data regularly to the scoring server according to the status of the environment which is controlled by the CPS. The blue part shows one instance of the process controller CPS which the Blue teams are responsible for; this is repeated for each Blue team.

Each simulated CPS can influence the simulated environment through two digital signals. The environment simulator can send back one analogue and two digital signals as the response of the environment. This setup enables the build-up of a simple and clear closed control loop, which is very commonly used for process control.

Both the environment and the process controller CPSs can be connected to other systems with different communication protocols, which gives an easy integration capability. One CPS instance can be connected to two different Human Machine Interface (HMI) devices. One of them is controlled by the Blue team, the other is installed for demonstration and support purposes and controlled by the Green team.

CONCLUSION

The cyber-physical battlefield was first introduced during Locked Shields 2019, when the platform was successfully integrated into the infrastructure of the exercise and provided a good practising environment for the Blue and Red teams. In this case the simulated critical infrastructure was a power plant, and the process controller regulated power production according to the power consumption that was sent via S7 protocol.

The key result of the research is the constructed battlefield that will be used in future exercises and training, since it can be reprogrammed according to the use case and can

be moved to external premises. Although the project was closed successfully, we faced difficulties with the mechanical construction since the platform has to be rugged enough to move while being compact and light.

Since the process controller CPS contains exactly the same program except when an external connection is established (in which case the IP address of the PLC differs), possible further research should be the realisation of fast PLC content multiplication with the same content but different addresses.

Keywords: Cyber-Physical System, Cyber-Exercise, Critical Infrastructure

REFERENCES

[1] E. A. Lee, 'Cyber Physical Systems: Design Challenges,' in *11th {IEEE} International Symposium on Object-Oriented Real-Time Distributed Computing {(ISORC} 2008), 5–7 May 2008, Orlando, Florida, {USA}*, 2008.

[2] G. Loukas, Cyber-physical attacks: A growing invisible threat, Butterworth-Heinemann, 2015.

[3] M. J. A. T. C. Robert M. Lee, 'Analysis of the Cyber Attack on the Ukrainian Power Grid,' Electricity Information Analyzing and Sharing Center, Washington, March 18, 2016.

[4] E. E. O. S. O. O. Oludele Awodele, 'Vulnerabilities in Network Infrastructures and Prevention/Containment Measures,' Department of Computer Science, Babcock University, Ilishan-Remo, Ogun State; Nigeria, 2012.

[5] French Network and Security Agency, 'Managing Cybersecurity for Industrial Control Systems,' 03 06 2012. [Online]. Available: https://www.ssi.gouv.fr/uploads/2014/01/Managing_Cybe_for_ICS_EN.pdf. [Hozzáférés dátuma: 06 04 2019].

[6] K. Kukk, 'Mapping The Best Practices For Designing Multi-Level Cyber Security Exercises in Estonia – Master's thesis,' Tallinn University of Technology, Tallinn, 2017.

[7] 'President Kaljulaid at the visit of the largest cyber defence exercise Locked Shields: politicians should make full use of its opportunities,' 26 04 2018. [Online]. Available: https://www.president.ee/en/meedia/press-releases/14269-president-kaljulaid-at-the-visit-of-the-largest-cyber-defence-exercise-locked-shields-politicians-should-make-full-use-of-its-opportunities/index.html. [Download: 08 04 2019].

[8] R. G. P. T. V. S. a. A. Z. A. Ogee, "The 2015 Report on National and International Cyber Security Exercises',' ENISA, [Online], 2015.

[9] Cyber Storm V: National Cyber Exercise, Cybersecurity and Infrastructure Security Agency (CISA), Online, 2016.

CYBER GAME TO CYBER EXERCISE: A NEW METHODOLOGY FOR CYBERSECURITY SIMULATIONS

Kieren Niĉolas Lovell
TalTech University of Technology
kieren.lovell@taltech.ee

1. INTRODUCTION

The increasing role that technology plays within society means our approach has changed from 'how do we stop cyber attacks?' to 'how do we respond effectively?'.

As a result, demand for cybersecurity simulations has increased. These exercises have normally been developed from the Capture the Flag (CTF) framework, have been run by academic & industry institutions for years, and their strategic track normally done separately, if at all (Red/Blue)[1].

This approach produces technically sound exercises. However measuring their impact in improving the cybersecurity posture of their respective responsibilities is very hard to assess, if at all[2]. This means the exercise becomes more a game, where the objective is to 'win', not to cooperate, develop, learn technical skills and practice the C3 skills as required to successfully take charge of an incident.

This document proposes a methodology to adapt the traditional exercises used by NATO to better provide an improved global approach towards cybersecurity simulations[3]. One reason for using NATO / H.M. Government emergency response doctrine as a baseline is their derivation from mature disciplines that have dealt with problems connecting strategic, tactical and operational aspects within a real time environment over their development cycle (FOST, for example, setup by First Sea Lord Louis Mountbatten in 1958)[4] [5] [6]. Generally, cybersecurity exercises only focus on one aspect within this battlespace: technical security; and at only one of the C3 layers: either the strategic pillar in accordance with their established national policy doctrine or within the technical arena[7] [8] [9]. Therefore the conclusion is that a cyber incident is 'the IT department's problem' rather than what it should be: an 'OPSEC problem' that affects everyone[10].

The traditional cybersecurity exercise approach sees the communication issues between the Command and technical teams *not practiced*. One learning outcome from any exercise should be to confirm the operational teams understand the impact of their actions, and the precedence order they should use to resolve incidents. The Command team should understand the technical/operational limitations of any response, and what can/cannot be achieved[11].

A number of simulations have been conducted by TalTech and the University of Cambridge that have attempted to bridge this command divide by combining the real-time elements and technical services to help drive the strategic response plan[9].

This paper is a proposal to document:

A new methodology for cyber exercises providing a framework connecting all aspects of an organization together and truly test their responses; it also provides a feedback loop to improve IT baseline, their command structure and provides lessons learned to be checked against when they conduct follow on exercises, for continuous development[2].

METHOD

The proposed method is in three phases.

Phase one: Conduct a full OSINT profile on the organization that is being exercised[12]. This covers multiple areas; HR data, PR material, and other public material, from leaked

or misconfigured services[13]. This is then supplemented with data that can be sourced through shodan.io and other passive scanning techniques[14] [15]. A full organization search is then completed for any exposed credentials, along with analysing social media footprints at both the individual and organisation level[16]. Data is then taken and used to produce an attack profile.

Phase two: Attack profile. Looking at the organisation's vulnerabilities, what's important to the operation of the entity, and how this could be exploited. This is the same as an impact and risk assessment which is used to implement security controls to mitigate possible threats[17].

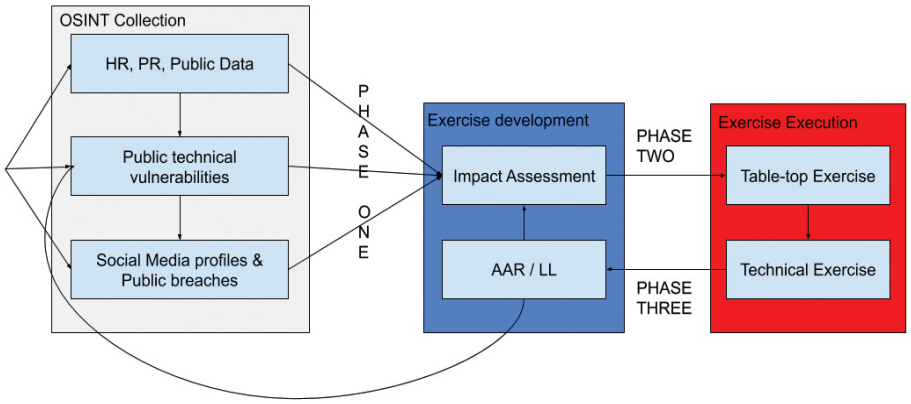


Figure 1.1. Proposed exercise life cycle

This impact assessment is then used to provide the storyline for the table-top[18]. Since this plan is using real information, this bridges the divide between technical and strategic areas; though you are simulating the attack, the real vulnerabilities remain realistic. Once the attack is played out, the strategic element can assess any damage that vulnerability would cause the organization. The IT teams are informed of actual concerns, and dialogue between Command and the operational team must develop to form a plan. The response plan is tested for Command & IT staff, providing a check on existing cybersecurity incident process, that procedures are helping, not hindering your efforts. The work above can also produce & adapt 'Standard Operating Procedure' (SOP) cards for CERTs, being particularly helpful with newly established security teams that have no procedures, informing SOP cards after writing the After Action Review (AAR)[6].

Phase Three: The AAR identifies lessons learned from the exercise. It lists rectifications in order of precedence, providing metrics and projected timelines to improve cybersecurity within the organization[19]. Changing the impact assessment for future exercises because identified exposures are mitigated and measures are put in place, this makes the exercise scenario develop organically to match current security weaknesses. The exercises will move from more generic threats to targeted threats (APT), as issues are mitigated, and the organisation's cybersecurity baseline matures[20].

PARTICIPANTS

The format described has been conducted by one large commercial business in Estonia (October 2018), the University of Cambridge Silver Team (July 2017), and TalTech University of Technology Rector's Office (November 2018). The OSINT techniques have been used for two major exercises, Exercise Neptune (a maritime OSINT exercise tracking and detecting NATO shipping) and Exercise Mercury (Universities across Europe).

ASSESSMENTS AND MEASURES

AAR: The assessment to understand if this approach has worked effectively is centred around the After Action Review[19]. AAR provides a metric to measure the current state of cybersecurity posture in terms of impact and effectiveness of the attack, not just how many compromises are valid. As more follow-on exercises are conducted, this will build metrics to calculate both the improvement in response times and exercises, and to provide a measurement of the organisation's improvement within network & system security, OPSEC, and COMSEC arenas.

Keywords: Cybersecurity, Exercises, NATO, Command, Control, Communications, C3

REFERENCES

[1] A. Cybersecurity, 'CTF Hacking: What is Capture the Flag for a Newbie?' [Online]. Available: https://www.alienvault.com/blogs/security-essentials/capture-the-flag-ctf-what-is-it-for-a-newbie. [Accessed: 21-Apr-2019]

[2] K. Maennel, R. Ottis, and O. Maennel, 'Improving and Measuring Learning Effectiveness at Cyber Defense Exercises', *Secure IT Systems*. pp. 123–138, 2017 [Online]. Available: http://dx.doi.org/10.1007/978-3-319-70290-2_8

[3] International Defence Training-Royal Navy, 'Flag Officer Sea Training', 2015. [Online]. Available: https://www.royalnavy.mod.uk/-/media/royal-navy-responsive/documents/idt/flexible-training-options/ost/fto12_ost.pdf?la=en-gb. [Accessed: 21-Apr-2019]

[4] S. A. P. Smith, 'USS Donald Cook Starts Flag Officer Sea Training'. [Online]. Available: https://www.navy.mil/submit/display.asp?story_id=86199. [Accessed: 21-Apr-2019]

[5] NATO, 'Allied Joint doctrine for Operational-level planning'. [Online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/787263/archive_doctrine_nato_op_planning_ajp_5_with_UK_elements.pdf. [Accessed: 21-Apr-2019]

[6] 'Community Emergency Response Team | Ready.gov'. [Online]. Available: https://www.ready.gov/community-emergency-response-team. [Accessed: 21-Apr-2019]

[7] International Telecommunications Union, 'Guide to developing a national cybersecurity strategy', 2018. [Online]. Available: https://www.itu.int/dms_pub/itu-d/opb/str/D-STR-CYB_GUIDE.01-2018-PDF-E.pdf. [Accessed: 21-Apr-2019]

[8] D. S. Alberts and R. E. Hayes, *Understanding Command and Control*. Ccrp Publication Series, 2006.

[9] 'Definitions and procedures', *College of Policing*, 23–Oct-2013. [Online]. Available: https://www.app.college.police.uk/app-content/operations/command-and-control/definitions-and-procedures/. [Accessed: 21-Apr-2019]

[10] 'OPSEC', *US Navy*, 2019. [Online]. Available: https://www.navy.mil/ah_online/OPSEC/. [Accessed: 21-Apr-2019]

[11] Strategic Studies Institute, S. Tatham, and U. S. Army War College, *U.S. Governmental Information Operations and Strategic Communications: A Discredited Tool Or User Failure?: Implications for Future Conflict*. Lulu.com, 2014.

[12] Department of the Army, 'Open-Source Intelligence – ATP 2-22.9', 10-Jul-2012. [Online]. Available: https://fas.org/irp/doddir/army/atp2-22-9.pdf. [Accessed: 21-Apr-2019]

[13] 'MetaData and Information Security', *Infosec Resources*, 31–Jan-2013. [Online]. Available: https://resources.infosecinstitute.com/metadata-and-information-security/. [Accessed: 21-Apr-2019]

[14] 'Shodan', *Vulnerability scanner*. [Online]. Available: https://shodan.io. [Accessed: 21-Apr-2019]

[15] 'Fingerprinting Methods Avoided by Nmap | Nmap Network Scanning'. [Online]. Available: https://nmap.org/book/osdetect-other-methods.html. [Accessed: 21-Apr-2019]

[16] D. Olenick, 'Data breaches up 400 percent, 15 billion records compromised: report | SC Media', *SC Media*, 08-Mar-2019. [Online]. Available: https://www.scmagazine.com/home/security-news/data-breach/data-breaches-up-400-percent-15-billion-records-compromised-report/. [Accessed: 21-Apr-2019]

[17] C. Biscoe, 'How to write an ISO 27001-compliant risk assessment procedure – *IT Governance Blog*', IT Governance Blog, 11-Jan-2018. [Online]. Available: https://www.itgovernance.co.uk/blog/how-to-write-an-iso-27001-compliant-risk-assessment-procedure. [Accessed: 21-Apr-2019]

[18] Cabinet Office, 'Emergency planning and preparedness: exercises and training', *GOV.UK*, 28-Apr-2016. [Online]. Available: https://www.gov.uk/guidance/emergency-planning-and-preparedness-exercises-and-training. [Accessed: 21-Apr-2019]

[19] J. E. Morrison and L. L. Meliza, *Foundations of the After Action Review Process*. 1999.

[20] Yuqing Li, Wenkuan Dai, Jie Bai, Xiaoying Gan , Member, IEEE, Jingchao Wang, and Xinbing Wang , Senior Member, IEEE, 'An Intelligence-Driven Security-Aware Defense Mechanism for Advanced Persistent Threats – IEEE Journals & Magazine', *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 3, pp. 646–661, 10.1109/TIFS.2018.2847671.

TEAM LEARNING IN CYBERSECURITY EXERCISES

Kaie Maennel
School of Information Technologies, Tallinn University of Technology, Tallinn, Estonia kaie.maennel@taltech.ee

Joonsoo Kim
School of Information Technologies, Tallinn University of Technology, Tallinn, Estonia joonsoo.kim@taltech.ee

Stefan Sütterlin
Faculty of Health and Welfare Sciences, Østfold University College, Halden, Norway stefan.sutterlin@hiof.no

I. INTRODUCTION

Operational work in cybersecurity often takes place in teams and requires effective knowledge sharing and collaboration between individuals, teams and organizations. Therefore team-based cybersecurity exercises (CSXs) are popular training methods. In addition to small-scale CSXs in educational settings (e.g., university course, competition across schools), there are several hundred exercises with learning-related performance objectives[1]. CSXs simulate realistic and complex environments and learning takes place in teams consisting of individuals with differing skillsets who need to perform together. As a learning outcome, the participants should be able to analyse, evaluate, synthesize, and articulate cyberpower effects in relation to geopolitical conditions and multidomain contexts[2].

Learning is a dynamic process in which learning steps, environment, individuals in the group, and group behaviours change as the team learns. When designing exercises, the organizers need to understand how teams learn, and what are the indicators of team learning. There is so far no clear definition and consensus regarding the best possible operationalization and assessment of team learning effects. Team learning can be defined as a process, in which a team takes action, obtains and reflects upon feedback, and makes changes to adapt or improve. The intertwined processes of sharing, storage, and retrieval processes need to take place for team learning to occur[4]. Learning involves collective thinking skills so that groups can develop abilities exceeding those of the sum of individual group member's talents[3].

We aim to analyse team learning in CSXs. As a novel aspect, we explore what characteristics of situational reports (SITREPs) may represent useful operationalizations of information processing associated with group learning. Situational reporting is commonly used in CSXs, and teams report on their collective knowledge of existing situation. Such reports are valuable for post-exercise reconstruction and sense-making of the exercise, as they should capture key incidents and events[5], and thus also show how teams have learned during the exercise. Our research contribution is building a relationship model between words/concepts in SITREPs, and how this can support team learning. Furthermore, we connect our model to other sources of information (such as automated scoring, human feedback, red team information, etc.). The findings can be used for designing technical solution for semi-automated SITREPs scoring to assist unbiased and comparable evaluation (e.g., script providing comparative analysis for human scoring), and provide more detailed feedback to teams.

Assessing learning outcomes based on self-reports comes with considerable shortcomings. As opposed to explicit declarative learning, implicit learning is a rather unconscious process. Behavioural changes can be caused by changed contextual situations or learning may

even be dysfunctional in regard to the mission goals. So far researchers have focused on limited set of learning outcomes, mainly learning of simple concrete facts; however, also cognitive, behavioural, and emotional learning outcomes should be considered[4]. Most common research methods are interviews (e.g.,[6]), surveys and observations (e.g., [4], [7]), and learning maps (e.g., [8]. [9]). However, there might be learned behaviour patterns (e.g., using metaphors) that members are not consciously aware of[4], [6]], and asking group members may not uncover any changes. These measurement methods are applicable for CSXs. However, as activities are conducted on computers/network – observations of behaviours (sitting quietly behind computer screen while mitigating a significant threat) might not provide sufficient information about learning. Observation is rather considered as looking into the "digital footprint" by applying non-intrusive measurements (such review of situational reports, or logfiles). There has been research conducted to measure team performance and effectiveness in CSXs, e.g., [5], [10], [11], [12], [13], but mainly using traditional obtrusive methods.

II. RESEARCH QUESTIONS AND DESIGN

Our research focuses on assessing team learning in absence of objective metrics, considering teams' uniqueness ("flux"), and while avoiding invasive assessment methods in CSXs. Our main research question is the following: what metrics reveal team learning success in CSXs. We hypothesize that metrics for learning effects can be extracted from SITREPs as a source, assuming that such commonly used situational reports are collective repositories of team knowledge. We will consider simpler technical metrics and more complex cognitive measurement constructs (such as cognitive agility index in[14]).

Our approach is analysing SITREPs to evaluate team learning utilizing natural language processing methods. We apply a mixture of quantitative and qualitative research methods for analysis of SITREPs and other commonly collected data. We operationalize information processing required for learning with metrics obtained from SITREPS (such as length, words and expressions used, corrections to previous reporting, consistency/style). We also analyse other relevant exercise data (such as scoring, injects, surveys). We use data of Locked Shields[15] – a team-based red/blue live-fire exercise organized by the NATO Cooperative Cyber Defense Centre of Excellence. The dataset includes over 5 years of exercises, each representing about 20 teams with several hundreds of participants in total.

III. INITIAL RESULTS

We have carried out initial analysis with some preliminary results. For example, our findings indicate that reflections about cyber-related real-world consequences and teams' critical assessment of their control over situation may be a good indicator for metacognitive processes and learning. Namely, the use of control- or off-control-related words is stronger associated with the physical rather than the cyber domain. However, such initial findings need validation by analysing inter-correlations and "semantic proximity"[16].

IV. CONCLUSION AND FUTURE WORK

Assessing learning effects in naturalistic environments is methodologically challenging. However, reliable and valid learning assessment is the precondition for accurate feedback, which in turn is the basis for performance improvement. Understanding learning processes should shape how learning is designed and delivered. At team-based CSXs, team learning dimension needs to be considered, and our research is a step towards evaluating team learning.

We will continue researching how to measure team learning at CSXs using "non-intrusive" methods and how team learning can be effectively transferred to organization. As future work, we apply identified metrics to reflection logs, and formalize reporting structures to enhance reflection on perceived individual and team performance. Such metrics form evidence-based foundation for semi-automated SITREPs scoring to ensure unbiased and comparable learning evaluation and provide feedback to teams.

Keywords: Cybersecurity, Cybersecurity Exercise, Teams, Learning, Measurement

REFERENCES:

[1] A. Ogee, R. Gavrila, P. Trimintzios, V. Stavropoulos, and A. Zacharis, "The 2015 report on national and international cyber security exercises." ENISA, 2015.

[2] B. J. Knox, Ø. Jøsok, K. Helkala, P. Khooshabeh, T. Ødegaard, R. G. Lugo, and S. Sütterlin, "Socio-technical communication: The hybrid space and the olb model for science-based cyber education," *Military Psychology*, vol. 30, no. 4, pp. 350–359, 2018.

[3] P. M. Senge, "The fifth discipline, the art and practice of the learning organization," *Performance+ Instruction*, vol. 30, no. 5, pp. 37–37, 1991.

[4] J. M. Wilson, P. S. Goodman, and M. A. Cronin, "Group learning," *Academy of Management Review*, vol. 32, no. 4, pp. 1041–1059, 2007.

[5] M. Granasen and D. Andersson, "Measuring team effectiveness in cyber-defense exercises: a cross-disciplinary case study," *Cognition, Technology & Work*, vol. 18, no. 1, pp. 121–143, 2016.

[6] A. C. Edmondson, "The local and variegated nature of learning in organizations: A group-level perspective," *Organization science*, vol. 13, no. 2, pp. 128–146, 2002.

[7] D. R. Newman, B. Webb, and C. Cochrane, "A content analysis method to measure critical thinking in face-to-face and computer supported group learning," *Interpersonal Computing and Technology*, vol. 3, no. 2, pp. 56–77, 1995.

[8] M. Uzumeri and, D. Nembhard, "A population of learners: A new way to measure organizational learning," *Journal of Operations Management*, vol. 16, no. 5, pp. 515–528, 1998.

[9] R. Chiva, J. Alegre,and R. Lapiedra,"Measuring organisational learning capability among the workforce," *International Journal of Manpower*, vol. 28, no. 3/4, pp. 224–242, 2007.

[10] D. S. Henshel, G. M. Deckard, B. Lufkin, N. Buchler, B. Hoffman, P. Rajivan, and S. Collman, "Predicting proficiency in cyber defense team exercises," in *Military Communications Conference, MILCOM 2016-2016 IEEE*, pp. 776–781, IEEE, 2016.

[11] K. Maennel, R. Ottis, and O. Maennel, "Improving and measuring learning effectiveness at cyber defense exercises," in *Nordic Conference on Secure IT Systems*, pp. 123–138, Springer, 2017.

[12] P. Legato and R. M. Mazza, "Modeling and simulation of cooperation and learning in cyber security defense teams," in *Proceedings – 31st European Conference on Modelling and Simulation, ECMS 2017*, pp. 502–509, 2017.

[13] Ø. Jøsok, B. J. Knox, K. Helkala, R. G. Lugo, S. Sütterlin, and P. Ward, "Exploring the hybrid space," in *International Conference on Augmented Cognition*, pp. 178–188, Springer, 2016.

[14] B. J. Knox, R. G. Lugo, Ø. Jøsok, K. Helkala, and S. Sütterlin, "Towards a cognitive agility index: the role of metacognition in human computer interaction," in *International Conference on Human-Computer Interaction*, pp. 330–338, Springer, 2017.

[15] "Locked shields," https://ccdcoe.org/exercises/locked-shields/.

[16] T. Slimani, "Description and evaluation of semantic similarity measures approaches," *arXiv preprint arXiv*:1310.8059, 2013.

SESSION 2: DIGITAL FORENSICS 1

Session moderated by Prof HAYRETDIN BAHŞI,
Tallinn University of Technology

Prof Matthew Sorrel,
"EXPLOITING DARK CURRENT FOR FORENSIC IMAGE IDENTIFICATION",
University of Adelaide

Mr Andrew Roberts,
"FRAMEWORK FOR INDUSTRIAL CONTROL SYSTEMS DIGITAL FORENSICS IN THE ENERGY SECTOR",
Tallinn University of Technology

Mr Samuel Henderson & Mr Brian Du,
"A PROACTIVE APPROACH TO IMPROVING THE WAY WE USE MACHINE LEARNING TO DETECT SOCIAL BOTS ON TWITTER",
University of Adelaide

Mr Akim Essen,
"MULTI-MODAL BIOMETRIC SYSTEM SECURITY AND PRIVACY",
Tallinn University of Technology

EXPLOITING DARK CURRENT FOR FORENSIC IMAGE IDENTIFICATION

Richard Matthews
University of Adelaide
Richard.matthews@adelaide.edu.au

Nickolas Falkner
University of Adelaide
nickolas.falkner@adelaide.edu.au

Matthew Sorell
University of Adelaide
matthew.sorell@adelaide.edu.au

ABSTRACT

The state-of-the-art method for fingerprinting digital cameras exploits the non-uniform output of an array of photodiodes due to the distinct construction of the PN junction when excited by photons. This photo-response non-uniformity (PRNU) noise has shown to be effective but ignores knowledge of image sensor output under equilibrium states without excitation (dark current). The dark current response (DSN) traditionally has been deemed unsuitable as a source of fingerprinting as it is unstable across multiple variables including exposure time and temperature. We hypothesise that DSN is responsible for introducing a temperature bias which can contaminate PRNU traces and through proper analysis, can lead to insights regarding the specific temperature at which an individual image under test was taken.

BACKGROUND

A reliable method of linking media to their source camera is through the analysis of sensor pattern noise (SPN) to generate a photo-response non-uniformity (PRNU) trace often referred to as a fingerprint[1]. When tested across the limited range of -7.9°C to 29.5°C it has been observed that this method is not affected by temperature[2]. [1]goes as far as to state that SPN "is not affected by ambient temperature or humidity" since PRNU dominants the SPN. In[3] it was demonstrated that DSN exhibits signal power which adds to the overall correlation energy during the original SPN methods even when using sensors that have DSN correction methods. It is accepted that DSN is heavily temperature dependent due to the dark current density formula:

$$1. \qquad J_d \sim T^2 e^{\frac{(E_t - E_G)}{kT}} \tag{1}$$

We investigate whether the current SPN methods are immune to temperature bias.

METHODOLOGY

We use three Sony IMX219 CMOS digital image sensors. The image sensors have built-in low dark current by design[4] through the use of correlated double sampling both before and after the Analogue to Digital Converter [5]. Through the use of custom-made experimental rig, we vary the temperature between 10°C and 50°C in 5°C increments. Temperature control is managed via a Peltier plate and MCP9808 solid-state thermal sensors. The aperture of the camera is covered with several layers of black electrical tape to ensure no photons are allowed to enter the imaging column. Covering the aperture ensures only dark frames are captured. Using a python script, we set exposure time to a constant t = 1/1008 s, and the effects of internal amplifiers are controlled by setting a predetermined long wait time during the setting up of the camera to allow the gains to reach a stabilised temperature before setting the ISO light sensitivity to 800.

Bayer raw information is appended to the end of each saved JPEG file. This is extracted using DCRAW[6]. To extract the RAW information from JPEG file it is converted to TIFF using DCRAW as per[7] using the command:

dcraw -D -T -4 -W filename

For each temperature interval, an image set of 100 dark frames is taken per camera. We prepare a noise residue for each dark frame by filtering each image using a high-pass filter in the discrete cosine domain to extract the high- frequency components using Matlab.

RESULTS

Using the theory presented in[8] we apply a model based on the dark current density model seen in Equation 1 to the measured results. This model has the exponential form $y = ae^{bt}$. Each model resolves with an adjusted R^2 value of .9449, .9787 and .9523 respectively for camera 1, 2 and 3. These models are shown in Figure 1 and then overlaid against the observed data for Camera 2 in Figure 2. There is a strong indication that the correlation is related to the DSN as hypothesised.

Using this model we can identify that the correlation increases up to an approximately constant value. This constant value occurs when the temperature of the DSN reference pattern matches that of the image. Using this analysis, we can determine an approximate temperature for each image set. This is shown in Table 1.

	Identified (°C)	Forensic Range (°C)	Actual Range (°C)
Camera 01	30.5	26.0 – 35.0	28.0 – 32.0
Camera 02	28.35	23.85 – 32.85	28.0 – 32.0
Camera 03	30.15	25.65 – 34.65	28.0 – 32.0

Table 1. Identified Temperature of Image Sets

To ensure a prohibitively long time was not needed to acquire images, all images were taken over a maximum 4°C range of the target temperature. When averaged this could cause the expected temperature of the image set to be between 28°C and 32°C however, it is more likely than not that the average of the images would be 30°C. Unfortunately, the temperature of the images under test was not recorded in the EXIF metadata meaning the exact temperature could not be independently verified.

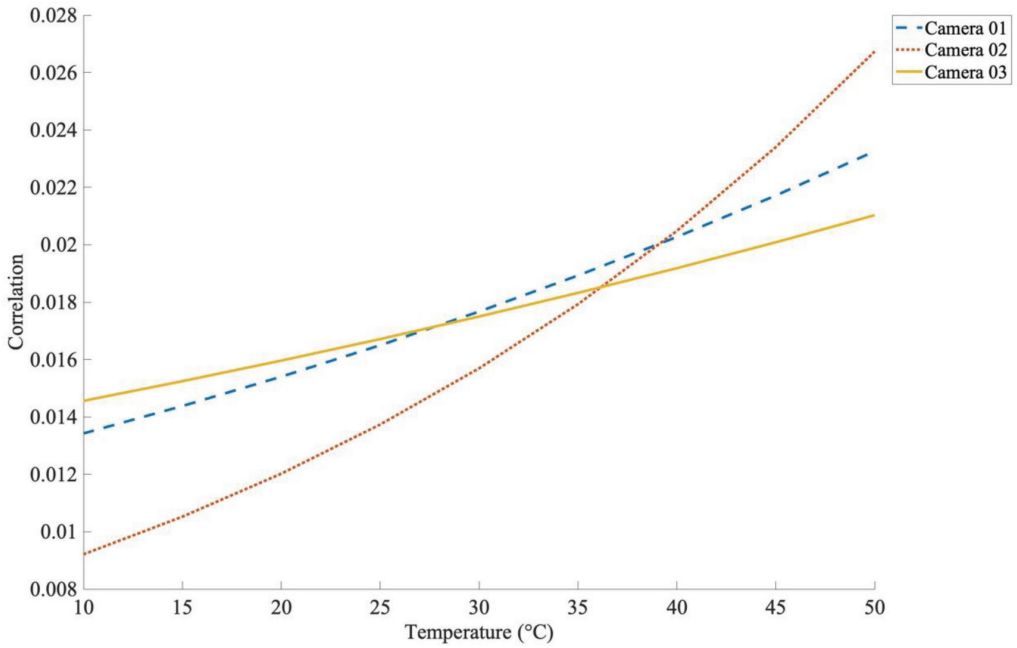


Figure 1. The three theoretical curves plotted against each other enabling an indication of the dopant strength to be determined.

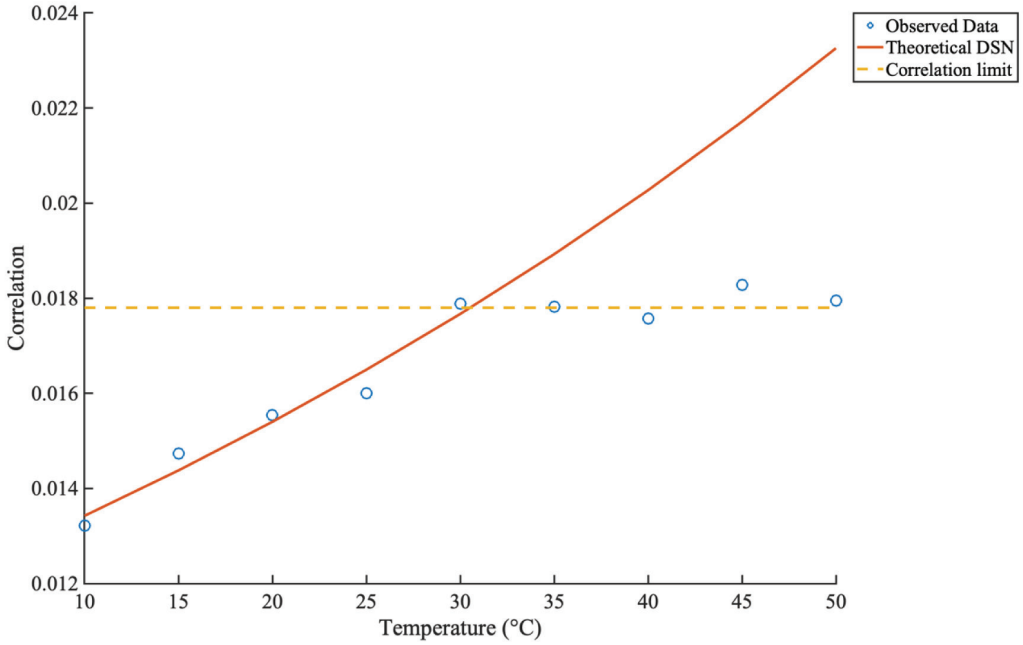


Figure 2. Correlation versus temperature plot for camera one showing the correlation increase in accordance with the theoretical model to a limit which corresponds to the temperature of the image sets under test. The temperature identified here is 30.5°C

CONCLUSION

In this paper, we have demonstrated a temperature bias in the method as first shown in[1]. This temperature bias present relates to the presence of dark current within an image and proves to be a useful forensic trace in its own right. We use this trace to isolate the temperature that an image is taken at independent of other sources such as EXIF metadata. This result is demonstrated across three CMOS image sensors of the same make and model and is experimentally linked to the dark current of the device.

Keywords: Dark Current, Temperature, Sensor Pattern Noise, Image Forensics

REFERENCES

[1] J. Lukas, J. Fridrich, and M. Goljan, "Digital camera identification from sensor pattern noise," *IEEE Transactions on Information Forensics and Security*, vol. 1, no. 2, pp. 205–214, 2006.

[2] Z. Ba, S. Piao, X. Fu, D. Koutsonikolas, A. Mohaisen, and K. Ren, "Abc: Enabling smartphone authentication with built-in camera," in *Network and Distributed System Security Symposium*, San Diego, United States of America, February 2018, pp. 1–15. [Online]. Available: http://dx.doi.org/10.14722/ndss.2018.23107

[3] R. Matthews, M. Sorell, & N. Falkner, "An Analysis of Optical Contributions to a Photo-Sensor's Ballistic Fingerprints," *Digital Investigation*, vol. 28, pp 139–145, 2019. [Online]. Available: https://doi.org/10.1016/j.diin.2019.02.002

[4] Sony Semiconductor, Japan, "IMX219PQ data sheet," https://www.sony-semicon.co.jp/products_en/new_pro/april_2014/imx219_e.html, accessed 7/01/2019.

[5] D. Bessette, "What is sony's exmor technology anyway?" https: //www.framos.com/en/news/what-is-sony-s-exmor-technology-anyway, accessed 27-Nov-2018.

[6] D. Coffin, "Decoding raw digital photos in linux," http://www.cybercom.net/~dcoffin/dcraw/, accessed 18-Nov-2018.

[7] S. Knight, S. Moschou, and M. Sorell, "Analysis of sensor photo response non-uniformity in raw images," in *International Conference on Forensics in Telecommunications, Information, and Multimedia*. Springer, 2009, pp. 130–141.

[8] G. C. Holst and T. S. Lomheim, *CMOS/CCD Sensors and camera systems, 2nd Edition*. SPIE Press, 2011.

FRAMEWORK FOR INDUSTRIAL CONTROL SYSTEMS DIGITAL FORENSICS IN THE ENERGY SECTOR

Andrew Roberts
Tallinn University of Technology
Andrew.Roberts@taltech.ee

I. INTRODUCTION

Safety and stability of critical infrastructure in the energy sector is key to the functioning of society and way of life. Cyber-physical attacks against the energy sector such as the Ukrainian Crash Override and Stuxnet demonstrate the lack of protection and resiliency of industrial control systems (ICS) to advanced cyber threats[1]. If resiliency of the electricity grid from cyber-attack is to be ensured than it is imperative that ICS forensics analysis is improved.

In the last five years, critical infrastructure operations have been impacted by an increase in the commodification of ICS malware, use of ransomware, and supply-chain and vendor-access compromises. The vulnerabilities exposed by ICS cyber-attacks point to a lack of hardening of defensive mechanisms such as intrusion detection, and incident response[2]. The role of digital forensics in providing intelligence feedback to strengthen anomaly detection and improve incident response is critical to reducing the attack surface, and limit the impact of cyber-attacks to systems which sustain human life. The U.S Defense Science Board in their report on resiliency of critical systems to advanced cyber threat noted the benefits of forensics as providing understanding of cyber-attack vectors, persistence within the network, time-to-detect, and time-to-remove[3]. ICS forensics can improve energy critical infrastructure cyber resiliency in three areas:

- Supply Chain Validation: through extracting the original base factory configuration of devices and assessing any changes in the state, before implementation into the environment.
- Root-Cause Analysis: root-cause of industrial events. Examples have included industrial events where cyber was suspected as the main cause, only for data analysis of industrial devices to reveal operator error and plant malfunction where the actual cause.
- Compromise Detection: Host analysis, malware analysis, network traffic analysis, and log analysis. The purpose is to deconstruct the cyber-attack and feedback the intelligence of the tactics, techniques and procedures to strengthen defense[4].

Despite the importance, there are few available models for ICS digital forensics and because of this, there is a lack of proficiency in operational use[5].

How can we define a framework for ICS digital forensics in the energy sector that can be used effectively in dynamic operational environments? The objective of the research is to provide better outcomes for information gathering from the forensic analysis of incidents in ICS that can be used to strengthen cyber resiliency. In this work we will mainly look at the state-of-the-art of ICS forensics.

II. STATE OF THE ART

There are many research projects and papers which focus on ICS digital forensics. Project CRISALIS was a research program focused on examining the vulnerabilities of critical infrastructure to cyber-attack and providing recommendations for detection, and prevention. The CRISALIS deliverable for forensic analysis in ICS defined the role of ICS forensics as providing: compromise detection, clean-state validation and root-cause analysis.

The project delivered a forensic tool, FERRET, which provided a technical platform for forensic analysis. However, the study wasn't focused specifically on the energy sector[4]. Since the release of the CRISALIS report in 2011, ICS malware have increased in sophistication and malwares such as TRISIS have targeted plant specific equipment such as safety instrumentation systems[6].

Studies by Slay et al. and Kniff accurately describe the difficulties translating the traditional model of computer forensics to ICS. Kniff defines ICS forensics as a combination of embedded system analysis, network forensics and intelligent data analysis[7]. ICS forensics is complicated as it must not impact continuity of operations of mission critical systems. Tools used in computer forensics are often not applicable to the unique physical and logical characteristics of ICS devices. Slay et al. concludes that further research is needed to define a framework for ICS forensics that develops forensic readiness in critical infrastructure[5]. Kniff, notes that there is a gap in research into assessing the performance of forensic application of anomaly-based threat detection[7].

De Montfort University and Airbus conducted a study to develop a methodology/toolkit for SCADA forensics. The study identified many of the issues with forensics analysis such as capturing volatile data from PLCs, and analysis of live data. The study didn't focus on a specific critical infrastructure sector and only provided a list of available forensic tools and a technical focused methodology based on literature reviews and reviews into the existing SCADA environments. The recommendations for future research focused on developing a more robust framework for the operational use of ICS forensics, encompassing; technology, organisational and procedural requirements[8].

The U.S Department of Homeland Security and European Network for Cyber Security in their studies and guidelines point to the importance of operational readiness for ICS forensics as one of the key metrics of incident response.[9]. Operational readiness of ICS forensics includes the critical infrastructure operational teams having an understanding of the forensics value of the assets in the environment, the forensics tools that can be used for each device or network component, and awareness and understanding of the digital forensics plan and when this is exercised as part of incident response[10]. The research provides guidelines and recommendations for the applicability of ICS forensics in operational settings, however, due to the age of publication, it omits the importance of forensic output to the intelligence feedback loop and for tuning defensive mechanisms such as the use of intrusion detection signatures.

III. CONCLUSION

There are many studies and approaches to ICS forensics. As cyber-attacks against critical infrastructure have increased in quantity and sophistication the attempts to forensically analyse these events have led to a realization of the uniqueness of ICS forensics and the challenges posed by a lack of tools, models, and methodologies. Future research needs to focus on developing a framework in a specific sector, mainly, due to its prominence, the energy sector, that is flexible enough to scale with the innovation of technologies and the complexity of the threat profile.

Keywords: Industrial Control Systems, Digital Forensics, Critical Infrastructure, Cyber Security

REFERENCES

[1] Sulmeyer, Michael, Military Set for Cyber Attacks on Foreign Infrastructure, Harvard Kennedy School: Belfer Center for Science and International Affairs, Apr 2018. URL: [https://www.belfercenter.org/publication/military-set-cyber-attacks-foreign-infrastructure]

[2] Dragos, Year in Review 2018: ICS Activity Groups and the Threat Landscape, Dragos, December 2018. URL [https://dragos.com/wp-content/uploads/yir-ics-activity-groups-threat-landscape-2018.pdf]

[3] Department of Defense, Task Force Report: Resilient Military Systems and the Advanced Cyber Threat. Defense Science Board Report, 2013, pp.66-67. URL: [https://nsarchive2.gwu.edu/NSAEBB/NSAEBB424/docs/Cyber-081.pdf]

[4] Patzlaff, Heiko, Report on forensic analysis for industrial systems, Project CRISALIS – Work Package 7 Deliverable, 2011, pp.6–54. URL [http://www.crisalis-project.eu/sites/crisalis-project.eu/files/crisalis_deliverable-D7.3.pdf]

[5] Slay, Jill & Sitnikova, Elena. (2009). The Development of a Generic Framework for the Forensic Analysis of SCADA and Process Control Systems. Forensics in Telecommunications, Information, and Multimedia. 8. 77–82. 10.1007/978-3-642-02312-5_9.

[6] Slownik, Joe, & Wylie, Jimmy, TRISIS: The First Safety Instrumented System Malware, Dragos, 2018. pp.4–6. URL: [https://www.sans.org/cyber-security-summit/archives/file/summit-archive-1521570761.pdf]

[7] Knijff, Ronald. (2014). Control systems/SCADA forensics, what's the difference?. Digital Investigation. 11. 10.1016/j.diin.2014.06.007. Developing Cyber Forensics

[8] Stirland, Joe & Jones, Kevin & Janicke, Helge & Wu, Tina & Publications, SDIWC. (2014). Developing Cyber Forensics for SCADA Industrial Control Systems.

[9] European Network for Cyber Security, CrashOverride/Industroyer Analysis, 2017. URL: [https://encs.eu/encs-document/encs-crashoverride-industroyer/]

[10] Department of Homeland Security, (2009). Recommended Practice: Developing an Industrial Control Systems Cybersecurity Incident Response Capability. URL [https://ics-cert.us-cert.gov/sites/default/files/recommended_practices/final-RP_ics_cybersecurity_incident_response_100609.pdf]

A PROACTIVE APPROACH TO IMPROVING THE WAY WE USE MACHINE LEARNING TO DETECT SOCIAL BOTS ON TWITTER

Samuel Henderson, Brian Du, David Hubczenko, Tamas Abraham, Matthew Sorell
The University of Adelaide, Defence Science Technology Group
samuel.henderson@student.adelaide.edu.au, guang.du@student.adelaide.edu.au,
David.Hubczenko@dst.defence.gov.au, tamas.abraham@dst.defence.gov.au,
matthew.sorell@adelaide.edu.au

INTRODUCTION

Social media has profoundly affected the way we acquire and process information. It has been reported that eight in ten Australians use social media[1] and that 52% of social media users utilise it to keep up to date with the news[2]. Furthermore, 17% report that it is used as their primary source of obtaining information[2]. A popular social media platform that is particularly effective at distributing information is Twitter, which will be the focus of this study.

Twitter's Application Programming Interface (API) enables external software to meld with the site and supports users in building bots. Social bots are social media accounts that automatically produce content and interact with humans[3]. Researchers have found that as many as 15% of active Twitter accounts are bots, with bot activity accounting for 50% of the site's traffic[5]. Over the past ten years, there has been an explosion of social bots[3]. While not all are malicious, some social bots can attempt to influence people through the spreading and amplification of misinformation. An example of this was the spreading of misinformation online during the 2016 US election[6]. A recent study by Shao et al. found that a mere 6% of Twitter accounts identified as bots were enough to spread 31% of the low-credibility information on the network[7].

With the increased uptake and usage of social media, it is concerning to consider the impact of these social bots, given their ability to spread and amplify misinformation. Researchers have sought to address this by using machine learning algorithms to detect social bots on social media. For Twitter, the current state-of-the-art classifier is Botometer (formerly known as BotOrNot)[4]. Current classification algorithms have followed a reactive schema, where detection techniques are based on collected evidence of existing bots. Adversaries, therefore, only have to modify the characteristics of their bots to evade detection. This leaves researchers always one step behind in a virtual arms race.

There has been an increased interest in the artificial intelligence community in the vulnerabilities of machine learning models (referred to as adversarial machine learning)[8,9]. In this study, adversarial machine learning techniques will be employed to study how an adversary may evade Twitter bot detection classifiers. Real-world adversaries often have no knowledge about the machine learning models they are trying to attack. Since Botometer is accessed through a public API and the model has not been made available, the most practical way to attack is using a black-box approach[11]. This involves constructing substitute machine learning algorithms to mimic Botometer, from which adversarial examples can be crafted. The purpose of this research is to highlight the vulnerabilities in the existing Twitter bot detection tools and to encourage their further development with adversarial machine learning concepts taken into account.

OBJECTIVES

The main objectives of this research project are to:

- Test the limits and vulnerabilities of a current, state-of-the-art Twitter bot classifier in an adversarial setting.

- Engineer adversarial examples and perform a practical black-box attack against the Twitter bot machine learning algorithm.
- Discuss defensive measures that can be implemented to improve the robustness of these classifier models.

BACKGROUND

BOTOMETER

Botometer is state-of-the-art in Twitter bot detection research. The tool generates more than 1,000 features from Twitter accounts using meta-data and information extracted from interaction patterns and content[4]. These features are then grouped and leveraged to train several different classifiers (one for each group and one for the overall score) using a Random Forest algorithm. These classifiers each output a score. Rather than use the raw score, the Botometer team developed a Complete Automation Probability (CAP) score to provide a better indication of whether an account is a bot or not. A higher account CAP score indicated a higher likelihood that an account is automated. Since the framework provides a continuous bot score, as opposed to a discrete bot/human judgement, an appropriate threshold must be determined to label the accounts. Recent research showed that a threshold of 0.43 maximised accuracy and enabled the classifier to correctly identify more modern and sophisticated automated accounts[4].

ADVERSARIAL EXAMPLES

Machine learning models are vulnerable to adversarial examples; malicious inputs designed to yield erroneous model outputs while appearing unmodified to human observers[9]. These adversarial examples exploit the imperfections and approximations made by the learning algorithm during the training phase. This phenomenon is analogous to the concept of optical illusions to humans. Recent research has demonstrated adversarial examples can be easily crafted with knowledge of either the machine learning model or its training data[8].

A concerning property of adversarial examples from a cybersecurity perspective is that it is possible to generate an adversarial example for any known machine learning model[8]. Another alarming property is that, if an adversarial example is effective against one machine learning model, it will likely be effective against others[10]. This property has been exploited to perform black-box attacks on machine learning models[9]. In a black box attack, the adversary constructs a substitute for the target model and generates adversarial instances against the substitute that can then be used to attack the target[11].

METHODOLOGY

In this study, a black box attack will be performed against Botometer. The following methodology is proposed:

1. Construct substitute models to mimic Botometer's classification algorithm.
2. Craft adversarial examples using the substitute models.
3. Attack the substitute models and Botometer with the same adversarial examples.
4. Evaluate the success and feasibility of the attacks.
5. Discuss defensive strategies that can be applied to current and future bot classification algorithms to defend against these kinds of attacks.

PRELIMINARY AND EXPECTED RESULTS:

The first phase of this study involves constructing substitute models to mimic Botometer's algorithm. To train substitute models, a labelled dataset is required. This was obtained by exploiting the Twitter API platform that allows for the streaming of real-time tweets. A small random sample of all public tweets that were produced in English, as specified in the tweet's language setting, were acquired. The screen names of the users responsible for the tweets were extracted and then passed to the Botometer Python API as input. Botometer output a series of scores for each user, and the threshold of 0.43 was used to label each account[4]. This labelling method was used because it is only necessary to train a substitute capable of mimicking Botometer's decision boundaries, rather than train a substitute model with optimal accuracy. The final dataset was made up of a balanced spread of 5,000 human and 5,000 bot examples.

A set of raw features were obtained by mining the meta-data of each user account. Statistical, sentiment and temporal analysis was performed on the meta-data to engineer a larger number of features. This large sample of labelled accounts and corresponding features was used as a training dataset for the substitute models. A subset of the training data was reserved for testing. The algorithms that were identified as the most suitable for this type of supervised learning were Random Forest, Gradient Boosting and Support Vector Machine. These algorithms were tested with the testing data and obtained accuracies of 88%, 87% and 80%, respectively. This accuracy result describes the similarity between the substitute model and Botometer.

Having obtained substitute models that effectively mimic Botometer, the weighting of each feature can be determined, and this information can be used to craft adversarial examples using existing frameworks[12]. Once the adversarial examples are created, a black-box attack will be conducted against Botometer's classifier. The results will be evaluated to determine which features can be realistically manipulated and hence, determine the feasibility of this type of attack in the wild. The findings of this research can be utilised to provide suggestions on how current and future defensive frameworks of machine learning algorithms can be improved.

Keywords: Bots, Twitter, Adversarial Machine Learning, Fake News, Black-Box Attack

REFERENCES

1. Yellow™, "Yellow Social Media Report 2018", *Yellow*, 2018.
2. Park, S., Fisher, C., Fuller, G. & Lee, J.Y. (2018). Digital news report: Australia 2018. Canberra: News and Media Research Centre.
3. Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). The rise of social bots. *Communications of the ACM*, 59(7), 96–104.
4. Varol, O., Ferrara, E., Davis, C. A., Menczer, F., & Flammini, A. (2017, May). Online human-bot interactions: Detection, estimation, and characterization. In *Eleventh international AAAI conference on web and social media*.
5. Yazan Boshmaf, Ildar Muslukhov, Konstantin Beznosov, and Matei Ripeanu. 2011. The socialbot network: when bots socialize for fame and money. In ACSAC: *27th Annual Computer Security Applications Conference. ACM*, 93–102.
6. Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of economic perspectives*, 31(2), 211–36.
7. Shao, C., Ciampaglia, G. L., Varol, O., Yang, K. C., Flammini, A., & Menczer, F. (2018). The spread of low-credibility content by social bots. *Nature communications*, 9(1), 4787.
8. Szegedy, C., Zaremba, W., Sutskever, I., Bruna, J., Erhan, D., Goodfellow, I., & Fergus, R. (2013). Intriguing properties of neural networks. *arXiv preprint arXiv:1312.6199*.
9. Goodfellow, I. J., Shlens, J., & Szegedy, C. (2014). Explaining and harnessing adversarial examples. *arXiv preprint arXiv:1412.6572*.
10. Papernot, N., McDaniel, P., & Goodfellow, I. (2016). Transferability in machine learning: from phenomena to black-box attacks using adversarial samples. arXiv preprint arXiv:1605.07277.
11. Papernot, N., McDaniel, P., Goodfellow, I., Jha, S., Celik, Z. B., & Swami, A. (2017, April). Practical black-box attacks against machine learning. In *Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security* (pp. 506–519). ACM.
12. Kulynych, B., Hayes, J., Samarin, N., & Troncoso, C. (2018). Evading classifiers in discrete domains with provable optimality guarantees. *arXiv preprint arXiv:1810.10939*.
13. Davis, C. A., Varol, O., Ferrara, E., Flammini, A., & Menczer, F. (2016, April). Botornot: A system to evaluate social bots. In *Proceedings of the 25th International Conference Companion on World Wide Web* (pp. 273–274). International World Wide Web Conferences Steering Committee.

MULTI-MODAL BIOMETRIC SYSTEM SECURITY AND PRIVACY

Akim Essen
Tallinn University of Technology
akimess@gmail.com

Matthew Sorell
University of Adelaide
matthew.sorell@adelaide.edu.au

Olaf Manuel Maennel
Tallinn University of Technology
olaf.maennel@taltech.ee

INTRODUCTION

With current advancements in biometrics and its use in modern technology like mobile phones, people are more inclined to use it with each day as it's simplifies and speeds up the process of authentication and identification. However, there are huge risks by using biometrics scanners that are not fully covered. Those risks are storage security and privacy.

Storage security is by itself a huge risk for users as if their biometric template is stolen, they will have to be deleted from the system because their biometric template is not a password and cannot be changed.

Privacy is an issue of trust. There are several cases of vendors, that provide biometric scanners and software, using the users biometric data for their own commercial purposes.

In this research, we study in-depth how those risks can be countered in the case of multi-modal biometric scanners.

Question behind this research is to find out if it's possible to replace a typical password authentication with multi-modal biometric solution without loss of efficiency, security and privacy.

BACKGROUND

With the past two decades, biometrics have seen a huge increase in its usage across different fields ranging from personal usage and commercial usage to law enforcement[1].This research will focus more on personal and commercial usage, as this is where the privacy issue arises more. Privacy is a significant issue that not only could lead to users private info becoming public but also it will raise people's trust in biometrics, especially for people from the IT field.

Nowadays research focuses on specific biometric scanners, such as ECG, iris or fingerprint. However, there is not much research done to tackle multimodal biometric scanners that are made to work together and share data across each other, which can open a whole new attack vector.

With each specific biometric scanner comes it's own problem or con, for example, face scanners can be easily spoofed if necessary counter-measures are not introduced, so if an additional layer of a different biometric scanner is introduced it can prove to be highly effective to reduce the failure and spoofing rate[2].

The main idea behind using only multi-modal biometrics is to remove the human error factor behind the usage of passwords. Computation power of systems are improving with each year, which leads to an increase in password security by enforcing stricter password policies and with time this will not be enough to remove human error[3]. Biometric scan-

ners entirely remove human error and can be a viable choice to replace passwords. However, there is a substantial issue of why biometrics didn't replace passwords yet, and that's storage security. The biometric template of users needs to be stored for comparison when authenticating, but if this biometric template is stolen, it can not be changed and the person will be simply removed from the system to prevent unauthorised access.

RESEARCH METHODOLOGY

The research has several steps:

1) Data collection methods from biometric scanners.
2) Methods of biometric data fusion.
3) Data mapping.
4) Performance evaluation.
5) Storage security method and evaluation.
6) Experimentation with different attack vectors.
7) Cost and usage evaluation.

The research will focus on only several biometric traits: ECG, pulse, gait and face recognition. These traits can be easily extracted through scanners that are widely available and open-source software. Each of those traits can be used as a separate biometric authentication system, but this research wants to find a correlation between them and use for one system to countermeasure biometric scanners downfalls. The researched system will use different sensors as it will increase reliability. The collected traits data will need to be fused and mapped through an application of classifiers and matching algorithms[4]. This will allow creating a biometric template and model that will enable authentication of a person.

Performance and costs evaluations will be made throughout the whole research to make sure the final result is suited for real-world usage.

With a complete biometric template, the research can move for storage security evaluations and what methods can be used to make sure the biometric template stays secure. Cancelable biometrics and visual cryptography are strong contenders as it not only allows to revoke your biometric template like a password, but also improves privacy, which in turn improves public confidence of biometric scanners[5].

Finally, once the biometric system has been covered, experimentation needs to be done, to find out what attack vectors are not covered, this will allow us to either find the countermeasures or evaluate the failure rate of the system.

CHALLENGES AND EXPECTED RESULTS

Performing data fusion will prove quite difficult as there are several stages to how it's done, that ranges from combining data at either sensor-level or feature-level to running it through classifier and matching algorithms. Even though the biometric data is different from each sensor, the result of fusing will provide a more reliable biometric system than using a single biometric trait and sensor. This will allow reducing matching errors and overall error rates. Expected result from this stage is a reliable biometric template and model that can prevent several spoofing attacks.

Storage security methods and privacy improvements evaluations are the targets of this research, by performing data collection and fusion, we can correctly pinpoint which steps can or needed to be changed to find a balance between performance, reliability, and security. A positive result is a reliable biometric system that can prevent biometric template theft or countermeasure it by using cancelable biometric, but also needs to remain optimised for real-life usage. Performing attacks on all of the steps of biometric data authentication will allow to pinpoint issues in the system and improve it. However, the resulting biometric system could need too much computational power or installation costs for it to prove profitable and serious competition for password replacement.

Each person is unique and if for one a high heart rate is a given, for others it's to high. While authenticating a biometric output some of the readings can be quite different from

each other on each consecutive authentication, which depends not only the environmental changes but also on the person's health. The issue here is the failure threshold. To combat this, on authentication of the biometric sample there are several stages to it, where each can result in an error. Using this errors the system can calculate the error rates and evaluate a match rate.[6] If one of the biometric sensors fails in extreme external conditions, for example ECG sensor results in a false negative result because the wearer is having health issues, then how it should be treated highly depends on the system requirements. This research will also focus to find the middle ground to extreme cases, be it raising the error threshold by verifying the health of the individual with other biometric sensors or instead providing assistance.

Keywords: Passwords, Biometrics, Storage, Security, Privacy, Authentication, Identification, Verification

REFERENCES

[1] Evans, N., Ross, A., Beng Jin Teoh, A. and Marcel, S. (2015). *Biometrics Security and Privacy Protection [From the Guest Editors] – IEEE Journals & Magazine*. [online] Ieeexplore.ieee.org. Available at: https://ieeexplore.ieee.org/document/7192815 [Accessed 12 Apr. 2019].

[2] Dahiya, N. and Kant, C. (2012). *Biometrics Security Concerns – IEEE Conference Publication*. [online] Ieeexplore.ieee.org. Available at: https://ieeexplore.ieee.org/document/6168378 [Accessed 12 Apr. 2019].

[3] Essextec.com. (2015). *IBM 2015 Cyber Security Intelligence Index*. [online] Available at: https://essextec.com/wp-content/uploads/2015/09/IBM-2015-Cyber-Security-Intelligence-Index_FULL-REPORT.pdf [Accessed 12 Apr. 2019].

[4] Kumar Sahoo, S., Choubisa, T. and Prasanna, S. (2012). *54 IETE TECHNICAL REVIEW | VOL 29 | ISSUE 1 | JAN-FEB 2012 Multimodal Biometric Person Authentication*: A Review. [online] Content.ebscohost.com. Available at: http://content.ebscohost.com/ContentServer.asp?T=P&P=AN&K=73183333&S=R&D=a9h&EbscoContent=dGJyMNHr7ESep7I4v%2BbwOLCmr1Gep7RSsa24TLCWxWXS&ContentCustomer=dGJyMPGptEm3rK9OuePfgeyx44Dt6fIA [Accessed 12 Apr. 2019].

[5] Kaur, H. and Khanna, P. (2016). *Biometric template protection using cancelable biometrics and visual cryptography techniques*. [online] springer.com. Available at: https://link.springer.com/article/10.1007/s11042-015-2933-6 [Accessed 12 Apr. 2019].

[6] J. Unar, W. Seng and A. Abbasi, "A review of biometric technology along with trends and prospects", *ScienceDirect*, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S003132031400034X. [Accessed: 06-Jun-2019].

SESSION 3: TECH 1

Session moderated by Prof TOBIAS EGGENDORFER,
University of Applied Sciences Ravensburg-Weingarten

Mr Kieren Niĉolas Lovell,
"EXERCISE NEPTUNE: MARITIME CYBERSECURITY TRAINING USING THE NAVIGATIONAL SIMULATORS",
Tallinn University of Technology

Ms Grethe Østby,
"SIEMS IN CRISIS MANAGEMENT: DETECTION, ESCALATION AND PRESENTATION – A WORK IN PROGRESS",
NTNU Department of Information Security and Communication Technology

Mr Liam Shelby-James & Stefan Norman,
"RELIABILITY AND TRUST IN GLOBAL NAVIGATION SATELLITE SYSTEMS",
University of Adelaide

EXERCISE NEPTUNE: MARITIME CYBERSECURITY TRAINING USING THE NAVIGATIONAL SIMULATORS

Kieren Niĉolas Lovell, Dan Heering
Tallinn University of Technology
kieren.lovell@taltech.ee, dan.heering@taltech.ee

INTRODUCTION

The maritime industry is the backbone of the global economy. In 2017 the volumes of cargo, that was transported with the ships around the world, reached 10.7 billion tons (Asariotis et al. 2018). During 2017 the global tonnage has increased by 42 million gross tons, which is equivalent to a 3.3 percent growth rate. In January 2018, the world fleet reached a carrying capacity of 1.9 billion dead-weight tons (dwt). In light of these numbers, the importance of maritime transportation cannot be overemphasized. The maritime industry has entered the new digital era of its evolution. New technological developments allow shipowners to operate the ships more safely and securely, optimize the sailing routes and save fuel. Smart shipping solutions are supporting crews and are improving the performance of the fleets. One of the biggest changes has also been the rollout of the internet connection onboard ships. The Maritime Labour Convention recommends that "reasonable access to ship-to-shore telephone communications, email and internet facilities should be available to seafarers, with any charges for the use of these service being reasonable in amount" (International Labour Organization 2006). According to the findings of the survey carried out by the Nautilus International, the union for maritime professionals, seafarers are increasingly making employment choices based on the availability of internet access (Nautilus International 2017). Nearly two-thirds of respondents said that they would consider changing the shipping company if it provided better onboard connectivity. The survey included 1,125 people from the UK, 665 from the Netherlands and as well as representatives of 18 companies giving the total sample size of nearly 2,000.

With continuous access to internet resources, social media, and e-mails, the seafarers, ships, and shipowners have become targets for motivated cybercriminals. In general, there are two categories of cyber attacks, which may affect companies and ships: untargeted and targeted (BIMCO et al. 2018). Targeted attacks are more sophisticated and can include the tools and techniques, which are specifically created for targeted shipping company of ships. These tools and techniques may include a distributed denial of service (DDoS) attacks, spear-phishing, subverting the supply chain, social engineering, impersonating a legitimate employee and others. The Port of Antwerp case in 2011 has shown that the collaboration of organized criminals and cybercriminals can lead to dangerous consequences for the community and the ports (Bateman 2013). Untargeted attacks are likely to occur due to the employment of tools and techniques available on the internet (scanning, water holing, phishing, malware, etc.). These types of cyber attacks may cause costly collateral damage for the shipping companies. In June 2017 the world's largest container shipping company, A.P. Møller-Maersk was one of the companies, which was hit by the malware NotPetya (Greenberg 2018).

This paper gives an overview of the exercise developed and carried out in June 2018 at a Cyber Security Summer School, which was organized by Tallinn University of Technology (TalTech). The novelty of this paper is to present a different approach to cybersecurity-related education and training of the seafarers and to point out the threats that emerge from the lack of cybersecurity awareness and cyber hygiene training, and the misuse of social media at sea. All participants were MSc and PhD students.

METHODOLOGY

Simulator-based training is one of the key factors in maritime education and training (MET) institution (Sellberg 2017). The environment created with the simulators allows the cadets to practice the skills and competencies that are needed for their future jobs. Navigational simulators also allow putting the cadets and seafarers in situations and conditions they would normally not encounter during their service at sea. Failures occurring in the simulated environment are incomparable to consequences on the real ship. TalTech Estonian Maritime Academy has a modern Simulator Centre with the navigational, maritime communication, engine room, refrigeration training, marine pollution control, and other simulators. The navigational simulator consists of four bridge simulators imitating the sailing of an actual ship (Figure 1).

Figure 1. Bridge simulator at TalTech Estonian Maritime Academy.

Exercise Neptune was developed to test the security of the legacy systems within the maritime navigational systems and to gather intelligence data of the real target ships sailing at sea during the time of the exercise and look for the possible cyber attack vectors (open-source intelligence (OSINT) exercise) (Rajamäki, Sarlio-Siintola, and Simola 2018).

The equipment and tools used during the simulator exercise:

- 4 Transas bridge simulators (Navi-Trainer Professional Simulator NTPRO 5000)
- 8 laptops with Windows 10 and PC-based chart plotter software Sea Clear II
- wireless network without access to the internet

The participants were divided into two divisions, four ships in each. Each group/ship received the laptop with preinstalled Sea Clear II software.

The aim of the Exercise Neptune is to simulate a threat aggressor in the closest possible way to a realistic terrorist type group. The easiest way to achieve this is to place the students into a cause. In this case, a civil war within Estonia was simulated, with two major factions having been formed. The reason for this kind of scenario is to take the participants out of their comfort zones and to get them to focus on their enemy and the purpose, but in a way where they work closely with other teams, making them exchanging data securely, and advancing their OSINT posture to the whole collective picture. It is simulated to originally place the teams against each other.

As the exercise plays out, it forces the teams to come together into one task force. This achieves two objectives. First, to create a highly focused team that is working on a number of ways to exploit the OSINT data and the vulnerabilities that they have assessed in the

system and then bring that together in one attack plan. When they exchange their data and results with others, it provides the creativity required to exchange their ideas, to adapt and make their respective attacks achievable. This is aided by the "Gamemaster" making the exercise a "high tempo" environment, rather than just a game. In placing constant deadlines, the participants quickly gain the Command, Control and Communication (C3) posture that would normally be present within a state or organised threat actor within a very short timeframe. This is required to understand what the threat landscape really is like. This methodology, while unorthodox, manages to create the results faster than traditional exercises and produces the work ethic normally found in groups that are fighting for a cause. This is an online version of the methodology used in Royal Navy workups during the Flag Officer Sea Training (FOST) training (Soeters, van Fenema, and Beeres 2010). First, focus on your department, then your ship, then on your task force, and then at the end, within the whole task group.

RESULTS

The results of the simulator exercise show, that the divisions were successful in developing cyber attacks against the opposing ships. They were able to breach the Electronic Chart Display and Information Systems (alter the course, manipulate with the chart data), interfere with the Automatic Identification System (AIS) data and compromise the Global Maritime Distress and Safety System (GMDSS).

As a result of the OSINT exercise:

- the teams were able to get hold of 7536 usernames and password used by the employees and crews of NATO warships;
- NATO ships could be tracked using SNAPMAP (map.snapchat.com), Twitter, Facebook and other social media sites;
- daily orders and confidential orders were found on Twitter in photos;
- FITBIT was being utilized by operational troops in exercise areas;
- webcams in ports were utilized to use as intelligence gathering assets (no usernames and passwords were in place);
- public relation departments were just as much to blame as individual sailors for their recklessness;
- it was recognised that mandatory policies are not being enforced.

CONCLUSION

The results of the exercise provided two major learning outcomes. One is that the digital footprint placed by individual seafarer is impacting the whole landscape. All of these individuals are only performing small breaches of data, but when you merge this with the collective intelligence, it provides a full tactical picture that can be then further exploited to provide a full strategic overview of their objectives. This suggests that the way this needs to be taught to seafarers and the maritime industry is in the same way, by demonstrating what the real results are within a real environment. In this way, we take the ownership of IT security from the hands of the IT security specialist and into where it should be: everyone's responsibility within any organisation as a whole. The maritime environment is different from a traditional office; it cannot have the same cyber hygiene approach that is used in this situation, as the threats and approaches are not the same.

It also proves that the hardware used for mission critical services (navigation, emergency communication, engine room software, etc.) can be easily exploited. These exploits, when merged with traditional intelligence gathering and OSINT profiling techniques, provides perfect injection points in where these exploits can be actioned.

In further discussions with the maritime industry, it also found that, like any other organisation, the responsibility for security positions is held across multiple silos. For example, the responsibility of GMDSS security is not held by the same person who is responsible for Desktop security. This means that there are holes in the whole process. This can only be achieved by a unification of the security posture, and ownership of the threats will be taken as one, in respect to the overall risk.

FURTHER RESEARCH

With the results from this exercise, the question is no longer "are ships exploitable" but more "how can we mitigate this threat when it happens, and in a way that the maritime industry can cope with this". Maritime industry can handle flood, fire, engine room and steering gear failures very well. More research is needed to develop bridge and operational procedures (kill cards) that help ship crews to identify possible cyber threats when they happen, and indicate what initial actions are required. Crews need to know how to escalate it to the correct authorities, and to other units in the area. More importantly, we see the need for establishing the drills that are required to make sure that a crew's conduct during an attack in question aids the safety of the ship, and do not hinder the situation, and that they all understand what is going on. With this in question, the research that is required is placing the competent crews within the bridge simulator, arranging possible cyber attacks and following the reactions of the participants. By repeating the exercises within high threat situations, the results will provide a good framework for establishing a good safety net for the shipping industry. It allows providing a more secure approach to cyber attacks for one of the most important industries.

Keywords: Cybersecurity, Navigation, OSINT, Simulator, GMDSS

REFERENCES

- Asariotis, Regina et al. 2018. *Review of Maritime Transport 2018*. https://unctad.org/en/PublicationsLibrary/rmt2018_en.pdf.
- Bateman, By Tom. 2013. "Police Warning after Drug Traffickers' Cyber-Attack." BBC News Europe (October): 2–5. https://www.bbc.com/news/world-europe-24539417.
- BIMCO et al. 2018. *The Guidelines on Cyber Security Onboard Ships*. https://www.bimco.org/products/publications/free/cyber-security.
- Greenberg, Andy. 2018. "The Untold Story of NotPetya, the Most Devastating Cyberattack in History." *Wired*. https://www.wired.com/story/notpetya-cyberattack-ukraine-russia-code-crashed-the-world/.
- International Labour Organization. 2006. "Maritime Labour Convention, 2006, as Amended (MLC, 2006)." https://www.ilo.org/dyn/normlex/en/f?p=NORMLEXPUB:91:0::NO::P91_SECTION:MLCA_AMEND_A3.
- Nautilus International. 2017. *An Investigation into Connectivity at Sea*. https://www.nautilusint.org/en/news-insight/resources/nautilus-reports/connectivity-at-sea-whitepaper/.
- Rajamäki, J., S. Sarlio-Siintola, and J. Simola. 2018. "The Ethics of Open Source Intelligence Applied by Maritime Law Enforcement Authorities." In *European Conference on Information Warfare and Security, ECCWS*.
- Sellberg, Charlott. 2017. "Simulators in Bridge Operations Training and Assessment: A Systematic Review and Qualitative Synthesis." *WMU Journal of Maritime Affairs*.
- Soeters, Joseph, Paul van Fenema, and Robert Beeres. 2010. "Managing Military Organizations: Theory and Practice."

SIEMS IN CRISIS MANAGEMENT: DETECTION, ESCALATION AND PRESENTATION – A WORK IN PROGRESS

Østby, Grethe; Yamin, Muhammad Mudassar; Al Sabbagh, Bilal
Norwegian University of Science and Technology; Stockholm University
grethe.ostby@ntnu.no; muhammad.m.yamin@ntnu.no; bilal@dsv.su.se

ABSTRACT

In this paper we discuss a work in progress to create a socio-technical escalation framework (STEF) to support synchronizing Security Incident and Event Management Systems (SIEMS) and Crisis Information Management Systems (CIMS) for crisis management during cyber crisis's. The process to create the escalation framework starts by first modelling the systems using a socio-technical approach and then using this modelling to outline a defining taxonomy for cyber crisis, escalation relevant information using SIEMS and in finally process have the SIEMS report information input to the CIMS to support the crisis management decision processes.

After this framework has been reviewed by the socio-technical research community we plan to test the model when setting up exercises in the Norwegian Cyber Range (NCR) environment. NCR will be an arena where testing, training, and exercise will be used to expose individuals, public and private organizations, government agencies to simulate socio-technical cyber security events and situations in a realistic but safe environment.

BACKGROUND AND INTRODUCTION

Every incident creates a need for information, both for people dealing with it inside the company as well as outside audiences. CIMS are implemented in many organizations to collect and correlate information during crisis. Mostly this information is submitted by personnel who are involved in dealing with the crisis.

Even though a CIMS will be used for multiple incidents, a crisis should be managed as a single event (Iannella & Robinson, 2007). E.g. a cyber crisis needs to be managed based on the taxonomies of such crisis. Cyber crises are more difficult to manage as the origin of the crisis is difficult to find, and there is a need for provided analysed information escalated from reliable information security sources.

Cyber threats are among the highest scored threats to business operations business reported by CEO's in PWC's annual global CEO report (PWC, 2019). As CEOs turn to what they can actively control inside their organizations, they confront the limitation in their own capabilities, especially the information and skills gaps according to the PWC's survey. Organizations struggle to convert data into useable and actionable intelligence, and the main reason for their frustration is among others 'data siloing' and 'poor data reliability'.

To cope with this situation, we propose an escalation framework by using SIEMS to provide analysed information in CIMS. By providing a clear step-by-step guide to follow, a CIEMS framework motivates companies to a more structured approach in their incident response procedures. For example, the framework can push organizations to make a detailed elaboration of roles and responsibilities within teams dealing with cyber security incidents. That will ensure that every organizations process has its owner and remains under control (Kulikova, Heil, Van Den Berg, & Pieters, 2013).

OBJECTIVES AND RESEARCH QUESTIONS

Our main goal is to develop a socio-technical escalation framework (STEF) that enables Security Incident and Event Management Systems (SIEMS) and Crisis Information Management Systems (CIMS) to synchronize information flow during a cyber-crises.

To better understand the scope and magnitude of the problem two research questions are proposed:

- RQ 1: How can we develop an incident taxonomy most suitable for cyber crisis management?
- RQ 2: What are the suitable methods and tools to escalate and present the cyber crisis information in crisis management systems?

We want to answer the questions by focusing on how SIEMS-systems can be employed in existing information systems, most specific in crisis information systems. Thereby, we want to approach our questions by what can be referred to as a naive inductivist approach. The naïve inductivist approach starts by first observing a phenomenon and then generalizing about the phenomenon which leads to theories that can be falsified or validated (Kowalski, 1994).

RESEARCH APPROACH

In this paper, we approach the cyber security challenges using the design science research in information systems (DSRIS) (Kuechler & Vaishnavi, 2012). Design science research (DSR) is a methodology which can be conducted when creating innovations and ideas that define technical capabilities and product through which the development process of artifacts can be effectively and efficiently accomplished (Kuechler & Vaishnavi, 2012).

How to work on DSR was presented in a thesis written by G. R. Karokola (Karokola, 2012). He visualized this approach as outlined in figure 1. As we are approaching our work in a naive inductivist approach, we modified the logical formalism in the model from abduction to induction.

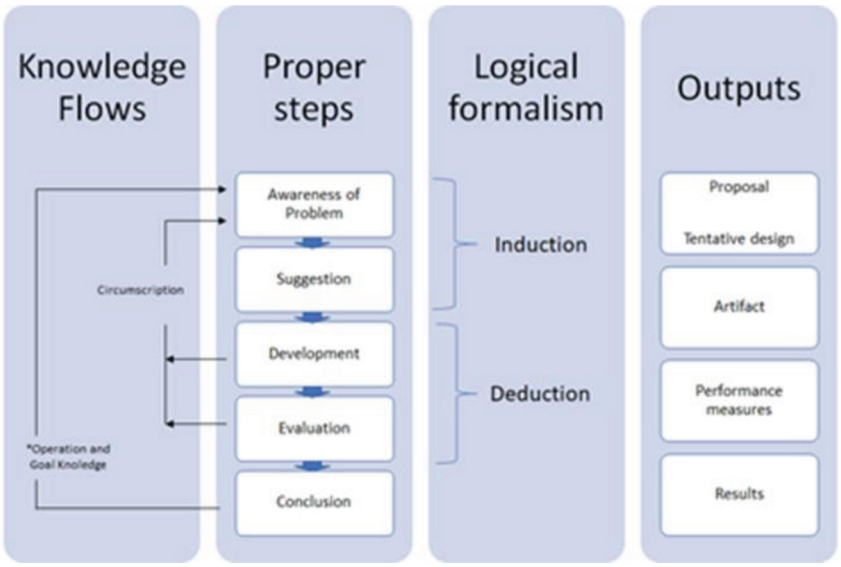


Figure 1. Design research methodology – modified

Our proposed artefact in this work in progress paper is a framework to deal with the problem in cyber crisis management in which analysed data from SIEMS system should be provided to support crisis decisions.

RESEARCH SCOPE

In this work in progress we plan to use the BSE model based on Rasmussen structural hierarchy model to visualize our approach (Cassano-Piché, Vicente, & Jamieson, 2006). We will present this as an acci-map. An acci-map is a systems-based technique for accident analysis, specifically for analysing the causes of accidents and incidents that occur

in complex Socio-technical systems. Different SIEMS-systems might be used on different levels in the model/system to improve information flow between 'data silos' and indicate what relevant information should be provided from a cyber incident to the CIMS systems.

After this framework has been reviewed by the socio-technical research community at the ICR 2019 we plan to test the framework when setting up cyber crises exercises in the Norwegian Cyber Range (NCR) environment. We wish to test the relevance of our framework in different management groups in Norwegian public sector to develop and evaluate our suggestions to provide artefacts that will be manageable during crisis.

Keywords: SIEMS, CIMS, Cyber-Crises, Cyber-Crises Taxonomies; Crises Information Escalation; Cyber Information Escalation, Cyber Decisions

REFERENCES

- Cassano-Piché, A., Vicente, K. J., & Jamieson, G. A. (2006). A SOCIOTECHNICAL SYSTEMS ANALYSIS OF THE BSE EPIDEMIC IN THE UK THROUGH CASE STUDY.
- Iannella, R., & Robinson, K. (2007). TOWARDS A FRAMEWORK FOR CRISIS INFORMATION MANAGEMENT SYSTEMS (CIMS) Olli-Pekka Rinta-Koski. Proceedings of the 14th Annual TIEMS Conference.
- Karokola, G. R. (2012). A framework for Securing a-Government Services, The case of Tanzania. Stockholm University.
- Kowalski, S. (1994). IT Insecurity: A Multi-disiplinary Inquiry. Stockholm University.
- Kuechler, W., & Vaishnavi, V. (2012). A Framework for Theory Development in Design Science Research: Multiple Perspectives. Journal of the Association for Information Systems (Vol. 13).
- Kulikova, O., Heil, R., Van Den Berg, J., & Pieters, W. (2013). Cyber crisis management: A decision-support framework for disclosing security incident information. In Proceedings of the 2012 ASE International Conference on Cyber Security, CyberSecurity 2012. https://doi.org/10.1109/CyberSecurity.2012.20
- PWC. (2019). CEOs' curbed confidence spells caution 22nd Annual Global CEO Survey.

RELIABILITY AND TRUST IN GLOBAL NAVIGATION SATELLITE SYSTEMS

Liam Shelby-James
The University of Adelaide, School of Electrical and Electronic Engineering
liam.shelby-james@student.adelaide.edu.au

Stefan Norman
The University of Adelaide, School of Electrical and Electronic Engineering
stefan.norman@ieee.org

Richard Matthews
The University of Adelaide, School of Electrical and Electronic Engineering
richard.matthews@adelaide.edu.au

Matthew Sorell
The University of Adelaide, School of Electrical and Electronic Engineering
matthew.sorell@adelaide.edu.au

INTRODUCTION

Global Navigation Satellite Systems (GNSS) are a ubiquitous & essential tool across many platforms and systems in the modern era. Many systems, and in fact, industries rely on these satellites for positional or timing data, and for this reason, GNSS requires protection & verification. Due to this reliance, it comes as no surprise that attacks on GNSS are of growing concern in cyber warfare.

For example, in 2013, academics in the Mediterranean Sea took a yacht off course without being detected[1]. In early 2019, also, Russian intelligence has interfered with NATO military training, and elsewhere throughout Russia, Crimea & Syria[2].

Over-the-air attacks come in three main varieties: jamming, spoofing and software attacks.

Jamming refers to a brute force Denial of Service (DoS) attack where noise is generated and broadcast on GNSS frequencies, making it difficult for any receivers to separate and decode any data.

Spoofing is a more targeted attack whereby false signals are transmitted to imitate genuine GNSS signals and can result in erroneous positions being displayed. If the signals are tailored to a specific receiver's location, this can then be slowly adjusted to bring ships & aircraft off course without any obvious indication to the operators[1].

Software attacks are not attacks on GNSS directly, but rather the software implementations of GNSS receivers. For example, sending malformed requests, or exploiting known bugs in the software[3].

Using these attacks individually or in combination may allow an attacker to disable or alter the location returned by a GNSS receiver, which may be intended to disrupt or disorientate the victim.

Keywords: GNSS, Spoofing, Satellite Navigation

OBJECTIVE

Our objective is to formulate a framework that determines how trustworthy the current GNSS location data is and to create a simple user-friendly metric using this framework to display to a user. A secondary objective is to implement the metric in an Android app, utilising low-level GNSS data made available in recent chipsets[4].

RELATED WORK

The focus of our research is primarily the detection of GNSS spoofing. There are several studies investigating these methods of spoofing detection, which range in their effectiveness, depending on the complexity & kind of attack. However, it is not clear if existing receivers are advanced enough to detect sophisticated attacks, as many of these methods are mathematically intensive.

The most fundamental approach is to detect satellites that have suspiciously high-powered signals, which may eliminate the simplest of spoofing attacks from malicious parties that just want their targets to lock on to their signal[5]. This method will fool the most rudimentary receivers that simply prioritise the strongest signals from a single constellation (constellation here means each system of satellites run independently by different organisations, eg; GPS, GLONASS, Galileo, etc.).

More rigorous methods include measuring the incoming signals' phase difference to determine the approximate direction of the signal source, as different satellites will be in different areas of the sky[6]. Some of these methods, such as cryptographic authentication, are computationally intense, as they utilise statistical hypothesis testing[5,7]. Some receivers will prioritise satellites which report being in different locations, as Geometric Dilution of Precision (GDOP) states that satellites in close proximity of each another will not provide as precise location data[8].

METHOD

Google has developed a GNSS Analysis suite targeted at application developers, which consists of an Android app (GNSS Logger) to capture raw GNSS data, and an analysis tool in compiled MATLAB code[4]. This analysis tool provides a wide array of graphs and statistics that can be used to characterise and analyse the captured data, as seen below in Figure 1. This does not demonstrate all of the data collected by this tool however, as additional data such as Automatic Gain Control (AGC) is also recorded and can be useful for spoofing detection.

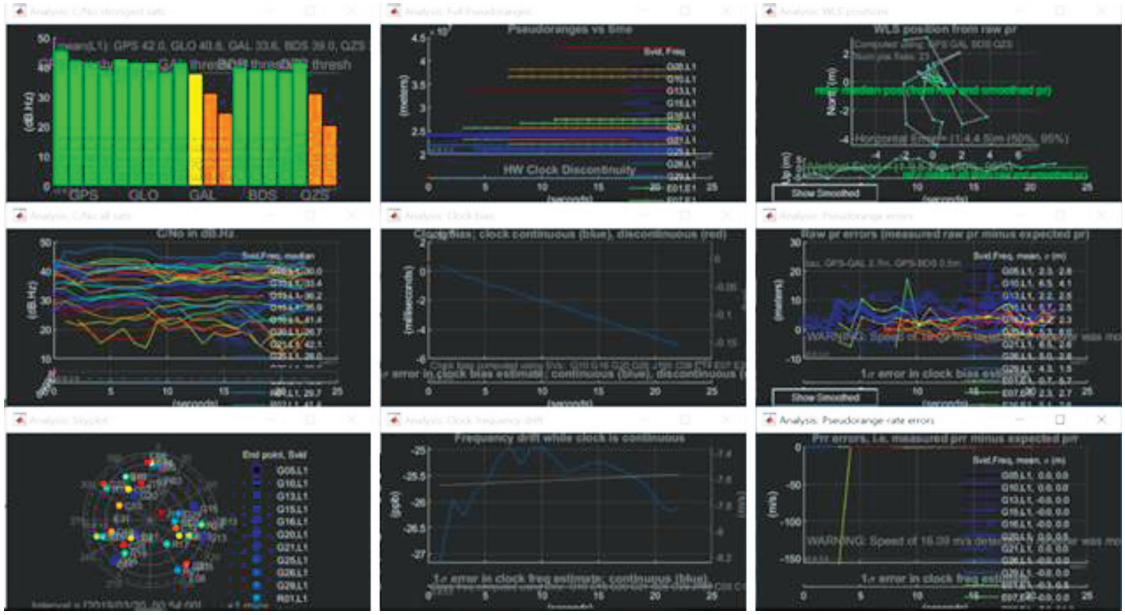


Figure 1. 3 Screen capture from GNSS Analysis (Google.)

The left column shows satellite data, from top to bottom; the strongest satellites from each constellation, each satellite's signal strength over time, and the rough location of each satellite in the sky. The centre column shows clock & timing data, again from top to bottom; the distance from user to each satellite (called "pseudorange"), over time, how much the receiver's clock's frequency must be changed to correct timing over time, and how much it has changed over time. Finally, the third column contains calculated data, including; variation in position, the magnitude of errors in pseudorange over time, and the number of errors in pseudorange over time.

We have used this GNSS Analysis suite as the basis for our data collection and analysis as it requires no translation to implement in an Android app and provides for easy integration into an SQL database for scalable analysis.

However, in order to test our program, a method to broadcast our own GNSS signals is required. Due to strict regulatory requirements and to avoid interfering with the public, these signals must be broadcast in an electrically shielded environment[9]. Shielding the testbed is a non-trivial problem, and we plan to solve this with a rudimentary grounded Faraday cage fabricated from sheet metal, and lined with copper mesh for conductivity.

Several Software Defined Radio (SDR) transmitters have been investigated, but many have a significant expense (from €200 to €1,200) which due to budgetary constraints would limit testing to a single transmitter[10]. After using a high-quality transmitter to provide a reference transmission, we are using several USB to VGA converters (approximately €8 each) which can be hacked with open-source software to become a crude transmitter[11]. This hack utilises higher harmonics of the 165 MHz VGA digital to analog converter, which can create undesirable interference. However, as this will be tested inside an isolated environment, this is not of concern.

In order to provide a realistic testing environment, real-world GNSS signals must be injected into our container in near real-time. There are several GNSS simulator tools capable of transmitting realistic signals for this purpose, as well as simple hardware repeaters to replicate signals received outside the containment, and at the time of writing, we are still investigating them[12].

PRELIMINARY RESULTS

We have identified several key concerns to be analysed from the raw GNSS satellite data in Table 1 below. The table contains descriptions of spoofing detection methods, and what data sets they must be analysed against to detect suspicious results.

"Absolute" here means that a satellite's data is compared to a numerical constant, or set of constants, and not other satellites. The column "Self" refers to whether each satellite's data is compared to its own historical data, as collected on the device. The "System" column indicates the satellite's data is analysed with respect to other satellites in the same constellation. Finally, "All", applies to data that will be compared to all satellites in range.

Table 1. Proposed Data Analysis

Title and Description of Method	Absolute	Relative to		
		Self	System	All
Satellite ID exists: checking to see that every SVID that is received is an actual, in-operation satellite	✓			
Satellite is in the expected location: if the satellite with the given SVID should be reachable in this part of the world	✓			
Duplicate signals from the same SVID: if a single SVID appears more than once from different sources, at least one of these is unauthentic			✓	
Unexpected distance jump: when spoofing occurs, there may be a single, sharp position jump that is obvious		✓	✓	✓
Time variance/jitter: if the timestamp given is delayed, this can indicate a man-in-the-middle attack. Also, if the clock bias is noticeably large, this may indicate the spoofing device is not as reliable as the atomic clocks used by satellites		✓	✓	✓
Time of arrival: if multiple signals are consistently received instantaneously, they are likely from a single spoofer device			✓	✓
Power of signal: if a certain set of satellites' signals are well above others, or if signal strength was slowly increased over time	✓	✓	✓	✓

If there are any malicious signs detected during data analysis, flags will be raised ranging in level with respect to the severity of the observed interference. Based on the number of flags and their impact, the application will utilise our framework to place a value on the level of interference, which will manifest in a certain colour grade, as per the table below. These are the key results of our research, and we expect this kind of algorithms will be most useful for transport companies, but also eventually more widely by the public as well as defence as a rough estimate on the type of interference in the area.

Table 2. Proposed Metric

Purple	●	Encrypted (typically defence)
Blue	●	Authenticated (e.g. GALILEO 2019+)
Green	●	No suspicious activity detected
Yellow	●	Suspicious activity
Red	●	Inconsistent; signal unverifiable, may be heavy interference
Black	●	Device compromised; the device is using a false signal
White	○	No/insufficient signal detected (environmental or malicious)

Keywords: GNSS, Spoofing Satellite Navigation

REFERENCES

1. J. Bhatti, T. Humpfries, "Covert Control of Surface Vessels via Counterfeit Civil GPS Signals", Journal of the Institute of Navigation, 2014
2. *Above Us Only Stars – C4ADS*. [Online]. Available: https://www.c4reports.org/aboveusonlystars. [Accessed: 10-Apr-2019].
3. T. Nighswander, B. Ledvina, J. Diamond, R. Brumley, and D. Brumley, "GPS software attacks," Proceedings of the 2012 ACM conference on Computer and communications security – CCS 12, 2012.
4. *"Raw GNSS Measurements | Android Developers."* [Online]. Available: https://developer.android.com/guide/topics/sensors/gnss. [Accessed: 10-Apr-2019].
5. T. Humpfries, B. Ledvina, M. Psiaki, B. O'Hanlon, P. Kintner Jr, "Assessing the Spoofing Threat: Development of a Portable GPS Civilian Spoofer", *In Radionavigation laboratory conference proceedings*, 2008.
6. Y. Hu, S. Bian, B. Ji, J. Li, "GNSS Spoofing Detection Using Fraction Parts of Double-Difference Carrier Phases", *The Journal of Navigation, 2018*
7. K. Wesson, M. Rothlisberger, T. Humpfries, "Practical Cryptographic Civil GPS Signal Authentication", *Journal of The Institute of Navigation Vol. 59*, 2012
8. "GPS Accuracy: HDOP, PDOP, GDOP, Multipath & the Atmosphere," *GIS Geography*, 24-Feb-2018. [Online]. Available: https://gisgeography.com/gps-accuracy-hdop-pdop-gdop-multipath/. [Accessed: 10-Apr-2019].
9. Australian Communications and Media Authority, "Devices prohibited by the ACMA," 10-Feb-2019. [Online]. Available: https://www.acma.gov.au/Citizen/Spectrum/About-spectrum/What-is-spectrum-and-why-you-need-it/devices-prohibited-by-the-acma. [Accessed: 10-Apr-2019].
10. "HackRF One," *Great Scott Gadgets*. [Online]. Available: https://greatscottgadgets.com/hackrf/one/. [Accessed: 10-Apr-2019].
11. "SDR (Software Defined Radio) » osmo-fl2k," *Steve Markgraf*. [Online]. https://osmocom.org/projects/osmo-fl2k/wiki/Osmo-fl2k, accessed on: 10 Apr. 2019
12. Osqzss, "osqzss/gps-sdr-sim," *GitHub*, 28-Oct-2018. [Online]. Available: https://github.com/osqzss/gps-sdr-sim. [Accessed: 10-Apr-2019].

SESSION 4: LEGAL RESPONSES TO CYBER THREATS

Session moderated by Dr ANNA-MARIA OSULA,
Tallinn University of Technology/Guardtime

Dr Jakub Harašta,
"ROLE OF LAWYERS IN CYBER EXERCISES: QUALITATIVE STUDY",
Masaryk University

Ms Ivana Kudláčková,
"LEGAL CONSTRAINTS ON CYBER WEAPONS",
Masaryk University

Ms Marija Makariūnaitė,
"DECRYPTION PASSWORDS AND BIOMETRIC AUTHENTICATION VS. LAW ENFORCEMENT",
University of Tartu

ROLE OF LAWYERS IN CYBER EXERCISES: QUALITATIVE STUDY

Jakub Harašta
Faculty of Law, Masaryk University / Center for Cyber Law & Policy, University of Haifa
jakub.harasta@law.muni.cz

INTRODUCTION AND RELEVANCE

Lawyers are frequently engaged in cyber exercises in various roles. They take part in drafting scenarios from the legal standpoint, including subsequent scenario escalation through different stages of emergency. Also, they are engaged in the "play" itself, as blue team members (Locked Shields) or members of the target audience (CyberCoalition).

Their inclusion is considered necessary due to their role in real-world decision-making. In spite of that, no study was yet dedicated to mapping the way lawyers are engaged and what is expected from them over the course of different exercises. This paper aims to fill this knowledge gap and map different roles in which lawyers engage in cyber exercises.

I took part in several national and international cyber exercises, and I noted certain level of uncertainty with regard to what to expect from lawyers. Naturally, lawyers play role in real-life decision-making, and because cyber exercises aspire to be realistic, inclusion of lawyers is understood as desirable. That being said, it was often unclear whether lawyers should sharpen their knowledge in response to technical specifications of ongoing scenarios, or whether they should educate others. These uncertainties often led to unsatisfactory experience of both lawyers and IT personnel taking part in the exercise.

RESEARCH QUESTIONS

This submission aims to answer the following two research questions related to the role of lawyers in cyber exercises:

1. What is the role of lawyers in cyber exercises as perceived by (a) lawyers, (b) policy-makers and (c) IT personnel? Is there a common denominator for all three groups regarding the role of lawyers in cyber exercises?
2. What is the expected knowledge to be communicated from lawyers to other participants during cyber exercises?

DATA

Data for this research was collected through semi-structured interviews with 25 respondents who took part in cyber exercises in the past. Semi-structured interview consisted of background questions (position, experience, experience with interaction with lawyers, participation in cyber exercises) and impressions (benefits of exercises, expected role of lawyers in cyber exercises, experience with lawyers in cyber exercises). Semi-structured interview was selected because the group of respondents was diverse – from academics to senior military officials – and given the qualitative nature of the study, it was useful to ask additional (unscripted) questions to frame the topic in more exhaustive way.

Given the inability to identify all the possible respondents and achieve representative sample, snowball method was used to identify specific respondents. Respondents were divided into three categories: lawyers, IT personnel and wide group of people engaged in *policy*. The third category served mainly as the category for respondents who were neither lawyers, not IT personnel.

Every respondent was asked to conduct the interview in her preferred environment. Interview was advertised as 30 minutes long, however because the willingness to share

experience with regard to the topic differed, interviews span between 20 and 60 minutes. Interviews were recorded, transcribed, and send to respondents for authorization. On occasions, this lead to lengthy process of institutional authorization to ensure that no classified or otherwise protected information were disseminated. The process was lengthy, often spanning several months, however it ultimately led to set of authorised transcripts of interviews prepared for qualitative coding.

CONCLUSION/RESULTS

Below are some of the recurring topics that appear regardless of the analysed group of respondents:

- Lawyers serve; they do not lead the exercise. This is often problematic in countries where the use of legal advisors for government is not mature enough for lawyers to understand their supportive role.
- Lawyers are more likely to be listened to, if they understand their supportive role.
- Lawyers are more likely to be listened to because of their personal traits and ability to connect compared to their knowledge and position.
- Pointing out that something is illegal is not enough; explanation in nonprofessional terms is required, as well as ability to modify operational or tactical procedures to achieve the desired outcome.
- Lawyers are required to know more about IT than IT personnel is required to know about law.
- Cyber exercises often lack sufficient evaluation mechanisms; additionally participants often lack both the willingness and the ability to follow through with the non-IT knowledge (legal, policy) after the exercise.
- Engaging lawyers in exercises often lacks realism regarding their different pace of work.

These recurring topics will serve as starting point for further analysis, as some outcomes are still pending. Special attention is to be paid to operationalisation of results. I intend to provide answers to research questions in form of categorisation with sufficient descriptive and analytical value to allow framing of lawyers in exercises.

ACKNOWLEDGMENT

This research was supported by the Center for Cyber Law & Policy established by the University of Haifa in collaboration with the Israeli National Cyber Directoriate. The research was conducted under the auspices of the Minerva Center for the Rule of Law under Extreme Conditions, Faculty of Law and Department of Geography and Environmental Studies, University of Haifa.

Keywords: Cyber Exercises, Preparedness, Legal Advisors, Qualitative Study, Knowledge Development

LEGAL CONSTRAINTS ON CYBER WEAPONS[1]

Ivana Kudláčková
Institute of Law and Technology, Faculty of Law, Masaryk University
ivanaakudlackova@gmail.com

Cyber weapons are getting at the forefront of attention over a period of last years. As the discussions among researchers, military personnel, cybersecurity and IT specialists are intensifying, many pressing issues still remain unsettled. One of those issues is undoubtedly the scope of legal aspects associated with cyber weapons. So far, there are several studies making a contribution to clarification of particular aspects of cyber weapons[2], but there is no literature that would attempt to pay attention to all legal aspects of cyber weapons. This extensive issue is certainly complicated by a fact that a life cycle of a cyber weapon is necessary to be perceived as a long chain of different links. This cycle is not solely comprised of development, export, acquisition and deployment respectively. As presented at the 2016 International Conference on Cyber Conflict (CyCon U.S.), following ten phases of the life cycle of a cyber weapon may be distinguished – project definition, reconnaissance, design, development, testing, validation, intrusion and control, attack, maintenance, exfiltration.[3]

Furthermore, the situation is complicated by the fact that even though cyber weapons *'have been used, and indeed are rapidly proliferating across state arsenals'*[4], neither international consensus on a definition of cyber weapons has been reached[5], nor states have agreed to any specific regulations as to cyber weapons.[6] Even though that some scholars introduce types of cyber weapons, such as namely website defacements or vandalism, distributed denial of service, intrusions (including Trojans and trapdoors or backdoors) and infiltrations[7], those classifications are not significantly helpful. Cyber weaponry is updated constantly in order to ensure a successful attack[8] because once a cyber weapon is launched, the target then becomes aware of the issue and corrects the problem.[9]

The above-mentioned circumstances lead to the following questions. How should states react in order not to be falling behind? How should a legal framework look like to meet a requirement of sufficient flexibility?

RESEARCH

This upcoming research is motivated by the paper Cyber Weapon Reviews under International Humanitarian Law: A Critical Analysis written by Colonel David Wallace. He presented his contribution to be *'a clarion call for greater research and study in this critically important area.'*[10] In order to cover the widest scale of possible cyber weapons, this research inclines to a result-oriented approach to cyber weapons. Inspired by Thomas Rid and Peter McBurney, a cyber weapon is understood as *'a computer code that is used, or designed to be used, with the aim of threatening or causing physical, functional, or mental harm to structures, systems, or living things.'*[11] Further inspired by the paper Challenges and opportunities in cyber weapon norm construction written by Jacqueline Eggenschwiler and Jantje Silomon, this qualitative research will attempt to analyze contemporary legal norms of international law. As cyber weapons could be used both during armed conflict and in peacetime, the research will focus on two significant legal regimes (namely law of armed conflict and human rights law) and examine whether those regimes are adequate and flexible. At this initial stage, the research will focus on deployment of a cyber weapon. At its later stages, the research will further track back to preceding stages and analyse how legal constraints on deployment of cyber weapons will impact their development. At those later stages, the research will also consider domestic regulation. This

upcoming research will not aspire to introduce a legal regulation, but its main objective will consist in clarification of legal aspects that should be taken into account.

Keywords: Cyber Weapons, Legal Analysis, International Law, Law of Armed Conflict, Human Rights Law

REFERENCES:

[1] This research is supported by ERDF 'CyberSecurity, CyberCrime and Critical Information Infrastructures Center of Excellence' (No. CZ.02.1.01./0.0/0.0/16_019/0000822).

[2] See the list of references.

[3] Maathuis, Clara; Pieters, Wolter; Berg, Jan van der (2016). Cyber weapons: a profiling framework. 2016 International Conference on Cyber Conflict (CyCon U.S.).

[4] Anderson, Kenneth (2016). Why the Hurry to Regulate Autonomous Weapon Systems – But Not Cyber-Weapons. Temple International & Comparative Law Journal, Vol. 30, No. 1, p. 27.

[5] See Dr. Sana Saleh's talk 'The Critical Need for International Consensus on Cyber Weapons' at the Workshop on Ethics and Policies for Cyber Warfare. Available at: https://ccdcoe.org/library/publications/the-critical-need-for-international-consensus-on-cyber-weapons/

[6] Wallace, David (2018). Cyber Weapon Reviews under International Humanitarian Law: A Critical Analysis. P. 22. Tallinn Paper no 11. Available at: https://ccdcoe.org/library/publications/cyber-weapon-reviews-under-international-humanitarian-law-a-critical-analysis/

[7] See Valeriano, Brandon; Maness Ryan C. (2015). Cyber war versus cyber realities: cyber conflict in the international system. Oxford: Oxford University Press. pp. 34–35.

[8] See Bartos, Christopher A. (2016). Cyber Weapons Are Not Created Equal. U.S. Naval Institute Proceedings, Vol. 142, Issue 6.

[9] See Valeriano, Brandon; Maness Ryan C. (2015). Cyber war versus cyber realities: cyber conflict in the international system. Oxford: Oxford University Press. P. 36.

[10] Wallace, David (2018). Cyber Weapon Reviews under International Humanitarian Law: A Critical Analysis. P. 22. Tallinn Paper no 11. Available at: https://ccdcoe.org/library/publications/cyber-weapon-reviews-under-international-humanitarian-law-a-critical-analysis/

[11] Rid, Thomas; McBurney, Peter (2012). Cyber-Weapons. The RUSI Journal, Vol. 157, Issue 1, p. 7.

DECRYPTION PASSWORDS AND BIOMETRIC AUTHENTICATION vs. LAW ENFORCEMENT

Marija Makariūnaitė
University of Tartu
makariunaite@gmail.com

Law enforcement has to face new challenges created by encryption. Appropriate legal framework regarding orders to disclose decryption passwords or perform biometric authentication is essential, so that it would be possible to investigate offences and prosecute the guilty while still respecting the rights of individuals (including the accused), and acting within the rule of law. Currently, in most jurisdictions compelling to disclose passwords or open systems by biometric authentication is not clearly regulated. This should not be regarded as optimal neither for the law enforcement nor for ensuring the human rights. Firstly, without the appropriate legal framework law enforcement officials cannot be sure that their actions will not be regarded as acting in bad faith and unjustly set the guilty persons free and not punished. And secondly, without proper rules it can be unclear for the individuals how to assert their rights.

Right to protect oneself against self-incrimination is recognized in most legal systems, however, ways how forced disclosure of decryption passwords and biometric authentication is interpreted and dealt with in the context of this right varies among countries quite greatly. For example, in Estonia or Lithuania there are no specific rules regulating this issue, thus, no- one can be lawfully forced to disclose a password or perform biometric authentication. However, in jurisdictions with no special rules for this issue, password disclosure and biometric authentication usually is a matter of negotiation between the law enforcement and persons in question (usually the suspects in criminal proceedings). It is important to note that biometric authentication could be circumvented by technical means. For instance, via using fingerprint marks collected from various surfaces or via centralized state level data bases created when issuing identification documents (passports, etc.), which include certain biometric data. Facial recognition also could be regarded as a method of authentication that could be easily overcome, for instance, just by pointing the device in question towards the suspect.

Norway, however, quite recently has made an amendment allowing police to order anyone to open data-processing systems by biometric authentication and to perform such authentication by force if persons refuse to comply with the order. This change has been made because the Norwegian Supreme Court has decided that existing provisions for taking fingerprints, DNR or blood samples by force do not cover forcing someone to put their fingers on phone for opening it[1]. In my opinion, this amendment is a good addition to the existing regulation since it provides clarity both to the law enforcement and to society in general regarding this issue.

It is important to note that compelling to reveal decryption passwords differs from compelling to open systems via biometric authentication. Orders for biometric authentication is somewhat similar to the orders of taking fingerprints, DNR or blood samples by force because, for example, fingerprint exists objectively. If password is written down then it exists objectively, but what if the password is only in person's mind? For instance, should the password be considered neutral, like a document, that could be ordered to disclose, or testimonial – fully protected by the right against self-incrimination?

Some jurisdictions, for example, United Kingdom, France have laws stating that failure to disclose decryption key could result in imprisonment and fine[2][3]. These laws are inter-

preted as requiring to prove that the person in question knows the password, for example, has recently used it[4]. However, person could honestly forget the password, even if he/she has recently used it, or person could still know the password even if he/she has not used it recently, so proving this might raise a lot of doubts. It is also very probable that punishment for not disclosing a password can be viewed as relatively mild compared to the punishment that could be executed in relation to the data accessed by disclosure of the encryption passwords. This, for example, would be especially true, if encrypted data would be related to terrorism crimes, so that is something that should be taken into account as well.

There also is the question of deniable encryption[5], meaning that even if a person complies with the order to disclose a password, it still would not be possible to know whether the password opened all the encrypted data or if there still is a hidden partition/volume. For these reasons, in my view, the non-disclosure of passwords that exist only subjectively in person's mind should not be criminalized.

Ordering to disclose decryption passwords to border control also can become problematic, and not only because of the already above mentioned issues but also due to the existing attorney- client privilege or journalists' right to protect their sources, so these issues might require some special rules, if non-disclosure (whether existing objectively or subjectively) is regarded as an offence.

To sum up, the topic of decryption passwords and biometric authentication vs. law enforcement is a very complex one. It requires comprehensive comparative analysis between different legal systems, good understanding of technical issues, and looking for balance between the interests of law enforcement and compliance with human rights, so that a reasonable legal framework for this issue could be established.

Keywords: Encryption, Decryption, Password, Biometric Authentication, Self Incrimination, Law

REFERENCES:

[1] Ingvild Bruce, Forced biometric authentication – on a recent amendment in the Norwegian Code of Criminal Procedure, in Digital Evidence and Electronic Signature Law Review, Volume 14, 2017, available at http://journals.sas.ac.uk/deeslr/article/view/2429/2391 [27/05/2019]

[2] French Penal Code, as amended by the Act of 3 June 2016, article 434-15-2, available at https://www.legifrance.gouv.fr/affichCodeArticle.do?cidTexte=LEGITEXT000006070719&idArticle=LEGIARTI000006418646 [27/05/2019]

[3] Regulation of Investigatory Powers Act of the United Kingdom, article 53, available at https://www.legislation.gov.uk/ukpga/2000/23/contents [27/05/2019]

[4] Orin Kerr, Amicus Brief of Professor Orin Kerr on Standards for Compelled Decryption Under the Fifth Amendment (October 11, 2018). Massachusetts Supreme Judicial Court, No. SJC-12564; USC Law Legal Studies Paper No. 18-29; available at https://ssrn.com/abstract=3264866 [27/05/2019]

[5] Veracrypt, Documentation, Plausible Deniability, available at https://www.veracrypt.fr/en/Plausible%20Deniability.html [27/05/2019]

SESSION 5: TECH 2

Session moderated by Mr AYKAN INAN,
University of Applied Sciences Ravensburg-Weingarten

Mr Jaan Priisalu,
"ANALYSIS OF THE IMPACT OF POISONED DATA ON TWITTER CLASSIFICATION MODELS",
Tallinn University of Technology

Mr Matthew Theiley,
"RISC-V ISA CUSTOM EXTENSIONS FOR USE IN CRYPTOGRAPHY",
University of Adelaide

Mr Charlie Tran & Mr Stefan Smiljanic,
"UTILISING A VEHICLE TESTBED ENVIRONMENT TO DEVELOP DECEPTIVE CAN BUS ATTACKS",
University of Adelaide

Mr Ahmad Amine Loutfi,
"DE-HYPING BLOCKCHAIN-BASED CROSS BORDER PAYMENT SOLUTIONS: A QUANTITATIVE COMPARATIVE STUDY OF DECENTRALIZED BLOCKCHAIN INFRASTRUCTURES VS. SWIFT GPI",
Norwegian University of Science and Technology

ANALYIS OF THE IMPACT OF POISONED DATA WITHIN TWITTER CLASSIFICATION MODELS

Kristopher Price, Sven Nõmm, Jaan Priisalu
TalTech University
pricekr1221@gmail.com

INTRODUCTION

Many online communities today face growing problems of group polarization, radicalization, and fake news. Social networks such as Facebook tend to recommend users to connect with people they already know and share similar values and beliefs with. In this homogeneous setting, online users become less tolerant and willing to accept information that does not confirm their pre-existing views. It is within this context that fake news has been able to find a mass audience[1]. Fake news not only confirms but also reinforces and strengthens people's beliefs. According to J. Ratkiewicz et al., "when politically active individuals can avoid people and information they would not have chosen in advance, their opinions are likely to become increasingly extreme"[2]. Political researcher Cass Sunstein refers to this phenomenon as group polarization, and it has been a large factor in the radicalization of terrorists[3],[4].

These issues are exacerbated by bots – automated accounts that pretend to be real people on social media. Because of how often they post content, bots may be viewed as more trustworthy and be better at influencing people[5]. Data-scientists have sought to address this issue by using machine-learning to detect bots on social media. Many models have been developed to classify online accounts as bots or real people. These models are 'trained' by giving a set of accounts labeled as 'genuine' or 'bot' to several algorithms. These algorithms generate a model that is able to detect bots based on the common features of all the bot accounts given to it. For a simple example, if every bot account in a training data-set had a default profile picture, the resulting model would tend to classify similar social media accounts as bots. Real-life classification models use a greater range of features, however, and the algorithms used to train them are much more complex.

While much research has been done into detecting bots, not much focus has been put into how bots might avoid being detected. According to Zhouhan et. al., data-scientists who study automated social media accounts on and the people who create those accounts are engaging in a "virtual arms race"[6]. The behavior of social-media bots changes almost as quickly as researchers learn how to detect them. Rather than focus on how the behavior of social-media bots changes over time, this research is concerned with how a poisoning attack may affect the models used to detect bots. According to Xiao et. al., a poisoning attack occurs when an attacker is able to manipulate the data used to train a model. If the training data representing bot accounts has been altered in some way, the accuracy of the resulting model in detecting bots may be reduced[7].

METHODOLOGY

This research uses the Cresci-2017 data-set, which consists of over 14,000 Twitter accounts labeled as 'genuine', 'spam-bot', 'social-bot', or 'amplification-bot'[8]. This data is used to train several models for detecting each different kind of bot. In this research, several methods of poisoning the data are tested. This abstract will show the results of poisoning one model trained to detect spambots.

INITIAL ANALYSIS

Before poisoning a classifier model, however, the features in the data-set must be analyzed to determine which ones most influence the classifier result. The importance of the fea-

tures is calculated in two different ways. First, every feature is ranked in descending order of their Fisher score. The Fisher score results can be seen in Table 1.

Table 11. Fisher Score of Each Feature.

Feature	Fisher Score
Created_at	0.69
Lang	0.42
Statuses_count	0.2
Favourites_count	0.16
Time_zone	0.14
Utc_offset	0.069
Geo_follow_protect_verify	0.058
Friends_count	0.0011
Listed_count	0.00035
Followers_count	0.000032

Next, the F1 Score of a model is calculated based on excluding every feature. Features that are associated with a lower F1 Score when excluded are considered more important. The results of this method can be seen in Table 2.

Table 2. Calculated F1 Score for Excluding Each Feature.

Feature	F1 Score without
favourites_count	0.8945241
Utc_offset	0.9224079
Statuses_count	0.9261423
friends_count	0.9352764
Followers_count	0.9357285
lang	0.9362008
Listed_count	0.9362119
Created_at	0.9365591
Geo_follow_protect_verify	0.9371561
Time_zone	0.9481067

After average ranked position of importance is taken for each feature in both tables. The features are listed in this order in Table 3. For example, followers_count is the tenth most important feature in Table 1, but the fifth most important feature in Table 2. Therefore, in Table 3 it is considered the eighth most important feature.

Table 3. Features in Descending Order of Importance.

Feature
favourites_count
statuses_count
Lang
utc_offset
created_at
friends_count
time_zone
followers_count
Geo_follow_protect_verify
listed_count

POISONING METHODS

The first method of poisoning the data involves 'flipping' the labels of a random set of accounts in the data. For example, if an account is labeled as a 'spambot', its label is changed to 'genuine', and vis-versa. The second method is altering the feature in the data that is most influential in determining the classifier outcome. Table 3 shows that the most important feature when classifying spam-bots is the favourites_count, or number of tweets an account likes. Twitter accounts that like a very low number of tweets tend to be spambots. Because of this, the poisoning attack focuses on incrementing the value of favourites_count.

The results of these two methods are measured by how much the accuracy is reduced and what percent of accounts are poisoned. For the second method, a third metric is taken of how much the favourites_count is incremented when poisoning the data.

PRELIMINARY RESULTS

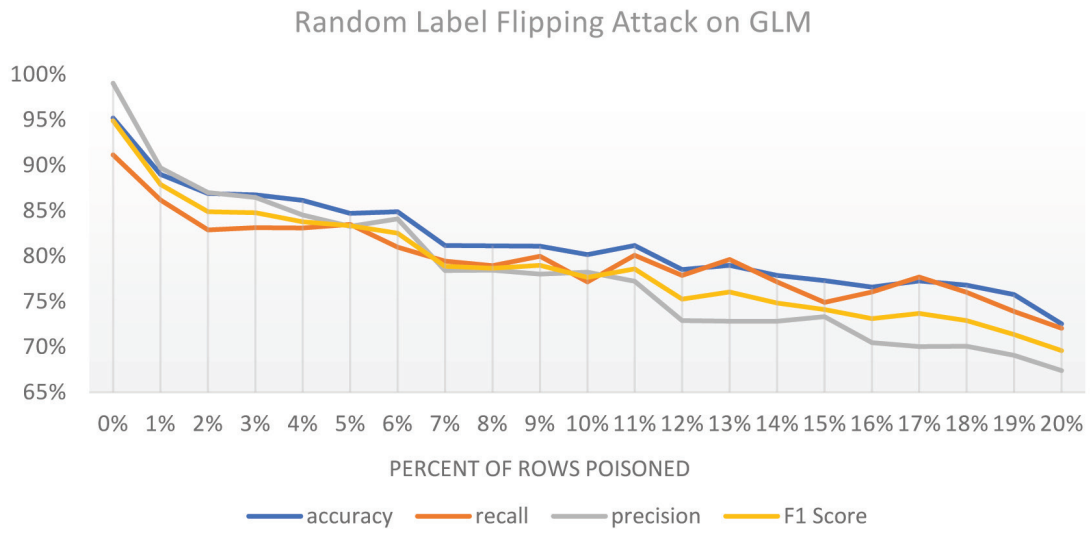


Figure 1. Results of Random Label Flipping Attack on GLM-trained Model for Detecting Spambots.

Figure 1 shows the results of the Label Flipping method on model trained using linear regression to detect spambots. This graph straightforwardly illustrates that as a greater percent of accounts have their labels flipped from 'spambot' to 'genuine' or vis versa, the accuracy, recall, and precision go down more.

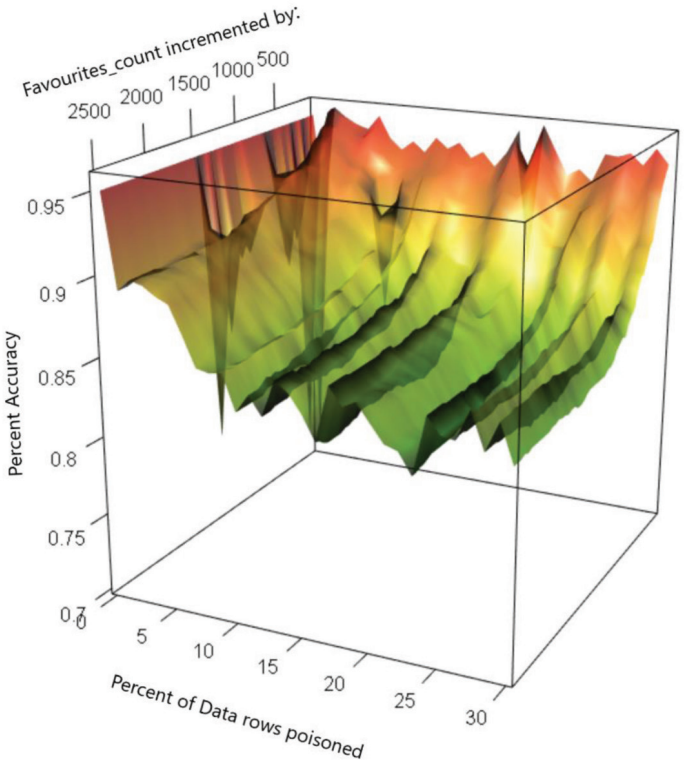


Figure 2. Accuracy of Poisoned Spambot GLM Classifier.

In Figure 2, a model for detecting spambot Twitter accounts has been attacked using the second poisoning method described in the methodology, where the favourites_count is incremented for a random set of Twitter accounts. This model is trained using linear regression. The accuracy of the model in detecting spambots is represented on the Y axis, with values ranging from 70 to 100%. The percent of accounts poisoned are represented on the X axis, ranging from poisoning 0% to 30% of all Twitter accounts. The value that favourites_count is incremented by for each poisoned account is represented on the Z axis, and ranges from 1 to 2500. The results show that there are diminishing returns in reducing the accuracy of the model.

Table 4. Local Minima for Accuracy in Figure 1.

Favourites_count incremented by number between	Percent of rows poisoned	Accuracy (1st model)	Accuracy (2nd model)	Accuracy (3rd model)
2400:2450	23%	81%	83%	84%
1050:1100	16%	81%	83%	83%
100:150	1%	72%	89%	89%

Table 4 shows three local minima for the results in Figure 2. The accuracy is reduced to 81% if 23% of all rows in the data-set are poisoned and the favourites_count is incremented by a number between 2400 and 2450. The accuracy is also reduced to 81% if 16% of all rows are poisoned and the favourites_count is incremented by a number between 1050 and 1100. Interestingly, incrementing the favourites_count by a number between 100 and 150 for only 1% of the rows reduces the accuracy to 72%. To determine if these results are consistent if different sets of rows are poisoned, two new models are generated based on a different random set of rows. The new results are consistent with the first two minima, but not with the third. This does show, however, that there is an optimal set of data-points that, if altered only a little, can drastically reduce the model's accuracy in detecting spambots.

SUMMARY

These initial results seem to show that models for detecting Twitter bots are vulnerable to poisoning attacks. However, the scope is limited to a specific model trained to detect a specific kind of bot. These methods must be repeated on several different models before a firm conclusion can be drawn. After analyzing the overall results, we can determine which attack method is more optimal and figure out how to implement it in real life. These analyses are covered more in-depth in the master's thesis of the same title which this abstract is based on. Hopefully the results of this study will help scientists mitigate the issue of online bots and help people deal with the many social issues that are prevalent in today's online communities.

Keywords: Adversarial Machine Learning, Social-Media, Bots, Fake News

BIBLIOGRAPHY

[1] D. M. J. Lazer et al., "The science of fake news: Addressing fake news requires a multidisciplinary effort," *Science* (80-.)., 2018.

[2] J. Ratkiewicz, M. D. Conover, M. Meiss, B. Gonc, A. Flammini, and F. Menczer, "Detecting and Tracking Political Abuse in Social Media," *Icwsm*, pp. 297–304, 2011.

[3] C. R. Sunstein, "The Law of Group Polarization," *J. Polit. Philos.*, vol. 10, no. 2, pp. 175–195, 2002.

[4] J. Johnson, "The Self-Radicalization of White Men: 'Fake News' and the Affective Networking of Paranoia," *Commun. Cult. Crit.*, vol. 11, no. 1, 2018.

[5] L. M. Aiello, M. Deplano, R. Schifanella, and G. Ruffo, "People are Strange when you're a Stranger: Impact and Influence of Bots on Social Networks," in *Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media*, 2014, pp. 10–17.

[6] C. Zhouhan, R. S. Tanash, R. Stoll, and D. Sabramanian, "Hunting Malicious Bots on Twitter: An Unsupervised Approach," *Lect. Notes Comput. Sci. Soc. Informatics*, pp. 501–510, 2017.

[7] H. Xiao *et al.*, "Is Feature Selection Secure against Training Data Poisoning?," vol. 37, 2015.

[8] S. Cresci, R. Di Pietro, M. Petrocchi, A. Spognardi, and M. Tesconi, "The paradigm-shift of social spambots: Evidence, theories, and tools for the arms race. e," in *26th International Conference on World Wide Web*, 2017, pp. 963–972.

RISC-V ISA CUSTOM EXTENSIONS FOR USE IN CRYPTOGRAPHY

Matthew Theiley, a1668977@student.adelaide.edu.au
Vu (Kelly) Hoang a1684248@student.adelaide.edu.au
Dr Matthew Sorell matthew.sorell@adelaide.edu.au
Dr Yuval Yarom yval@cs.adelaide.edu.au
The University of Adelaide, South Australia.

ABSTRACT

The focus for this paper was researching RISC-V ISS and compiler behavior to make it viable for use with cryptographic algorithms. A customized version of a well-known cryptography algorithm AES was produced to be compatible with RISC-V. Compilation of the algorithm showed it to be inefficient and insecure when built for RISC-V. The next steps for this project will be test the algorithm on the ISS ensuring that it works as expected and researching how to improve its security and efficiency. This paper aims to show that just because something is secure in software it does not imply that it will be secure when compiled to run on hardware.

DEFINITIONS

ISA	Instruction Set Architecture – Functional Computer Design.
ISS	Instruction Set Simulation – Simulation of ISA
RISC-V	Processor Name – Reduced Instruction Set Computer
ARM	Processor Name – Advanced RISC Machine
AES	Cryptography Algorithm – Advanced Encryption Standard
RSA	Cryptography Algorithm – Rivest-Shamir-Adleman
PGP	Cryptography Algorithm – Pretty Good Privacy
GNU	Compiler Toolchain – GNU's Not Unix
SIMD	ARM Extension – Single Instruction Multiple Data
Galois Field	A matrix used for calculating mixing of bytes for AES
S Box	A matrix used for calculating substitution of bytes for AES
C++	C++ is a programming language used to create software
Assembly	Programming language used to run software on ISS/hardware
Library	Adds additional functionality to a programming language
Crypto++	A C++ library providing support algorithms including AES and RSA
Vectorization	Turing concurrent events into events that run in parallel
Compiler	Converts a programming language into software
Toolchain	Includes tools for creating compatible software such as compilers

1. INTRODUCTION

1.1 SCOPE

This paper investigates how the RISC-V compiler responds to constructing cryptographic algorithms in C++. The paper shows how the RISC-V Toolchain was acquired along with how the algorithms used for testing were created. The paper also will look at some prelimi-

nary findings about how the RISC-V compiler responds to AES encryption and decryption algorithms and what this could mean for the RISC-V ISS and hardware implementations. Potential improvements that could be made to the RISC-V will also be explored.

1.2 PROJECT BACKGROUND

1.2.1 WHAT IS RISC-V

RISC V is an ISA, which is a design for how a processor operates on a functional level. This means that exact circuitry is not specified, but does reflect traits about the resulting hardware such as available operations that can be performed and memory structures[4], [5]. An ISS is a simulation of an ISA, in the case of this paper the ISS will be written in C++.

1.2.2 WHY RISC-V A GOOD CANDIDATE FOR THIS ANALYSIS

Currently RISC-V does not have a standard cryptography extension according to its ISA unlike ARM[3–5]. This means that there should be security flaws present which can be explored. It is ideal to work with RISC-V as it was designed to be easily extendible and simple to learn[4].

1.2.3 AES ALGORITHM

AES uses several rounds of encryption to obscure data. Each round uses a key which is derived from an initial key. The data is broken up into manageable chunks. Each round the data chunks are operated on using a combination of XORing, Substitution, Cycling and modular operations. AES is a symmetric key algorithm[2].

1.2.4 FUTURE ALGORITHMS

Compatible implementations of RSA and PGP are planned to be made later on in this project along with respective analysis of these additional algorithms[1], [11].

2. METHOD

2.1 ACQUIRING RISC V TOOLCHAIN

The RISC-V Toolchain was acquired by cloning the RISC-V Toolchain Repository[6]. This can be seen in figure 2.0.

```
git clone --recursive https://github.com/riscv/riscv-gnu-toolchain
```

Figure 2.0. Cloning RISC V Toolchain Repository

The toolchain was then configured and built using the pre-existing make file included in the repository[6]. Configuration is seen in figure 2.1.

```
./configure --prefix=/opt/riscv
make
```

Figure 2.1. Configure and Build Toolchain

2.2 CREATION OF CRYPTOGRAPHY ALGORITHMS

RISC-V GNU toolchain was found to be incompatible with existing cryptography libraries like Crypto++ so a customized version of AES was developed in C++[10]. This version of AES used only standard C++ libraries and functionality to ensure its compatibility with RISC-V.

2.3 RISC-V – GNU TOOLCHAIN ASSEMBLY CREATION

The RISC-V Toolchain includes tools for building C/C++ code into RISC-V assembly language and executables which will run on the RISC V ISS and hardware implementations. The Toolchain is also able to print out the assembly language and source code into human readable text as seen in figure 2.2. This text was used for compiler and ISS analysis.

```
riscv64-unknown-elf-g++ -g -c -Wa,-alh'
riscv64-unknown-elf-objdump -d -S'
```

Figure 2.2. RISC-V Assembly and Source Printout Commands

3. PELIMINARY RESULTS

Upon observing the human readable text output for the customized AES algorithm, it can be seen that there is a lot of loading a storing occurring. This is seen in figure 3.0. In particular many constant components in the system such as Galois fields and S boxes are being setup in memory when they could just be hardwired in the system. The assembly code for each of the steps in the algorithm are broken down into the basic instructions available in RISC-V like add, multiply, load and store[4]. Data and instruction composition are exposed throughout the setup, encryption and decryption steps of the AES algorithm.

```
                mat[i][j] ^= key_schedule[round_num][i][j];
9da:   fd843683          ld      a3,-40(s0)
9de:   fec42783          lw      a5,-20(s0)
9e2:   fe842703          lw      a4,-24(s0)
9e6:   070a              slli    a4,a4,0x2
9e8:   9736              add     a4,a4,a3
9ea:   97ba              add     a5,a5,a4
9ec:   0007c703          lbu     a4,0(a5)
9f0:   fd043603          ld      a2,-48(s0)
9f4:   fec42783          lw      a5,-20(s0)
9f8:   fcc42583          lw      a1,-52(s0)
9fc:   fe842683          lw      a3,-24(s0)
a00:   058a              slli    a1,a1,0x2
a02:   96ae              add     a3,a3,a1
a04:   068a              slli    a3,a3,0x2
a06:   96b2              add     a3,a3,a2
a08:   97b6              add     a5,a5,a3
a0a:   0007c783          lbu     a5,0(a5)
a0e:   8fb9              xor     a5,a5,a4
a10:   0ff7f713          andi    a4,a5,255
a14:   fd843603          ld      a2,-40(s0)
a18:   fec42783          lw      a5,-20(s0)
a1c:   fe842683          lw      a3,-24(s0)
a20:   068a              slli    a3,a3,0x2
a22:   96b2              add     a3,a3,a2
a24:   97b6              add     a5,a5,a3
a26:   00e78023          sb      a4,0(a5)
```

Figure 3.0. Example Snippet from RISC V Toolchain Assembly/Source Output

4. ANALYSIS

Despite the algorithm in C++ being designed to hide information and secure it through the use of symmetric keys, it can be seen in figure 3.0 that all code is broken down into a form where data flow and algorithm structure are visible. The RISC-V compiler has responded by building up the algorithm using instructions available from RISC-V. With no specialized cryptography instructions in the RISC-V ISA the compiler has no choice but to use many standard instructions rather than a few specialized ones. The increased amount of instructions should result in more clock cycles which would result in a longer time to run the algorithm[14]. The visibility of data flow and instruction decomposition would make this system leak out sensitive information about how the algorithms are operating. Resulting hardware based off the ISA using the AES algorithm would be vulnerable to side channel attacks as they could use this leaked information to either break the cipher or access the data directly[7–9].

Based on what ARM has currently done with its cryptography extensions it can be seen that they use a single instruction for encryption and decryption for AES specifically (AESE and AESD). They also have specific instructions for some of the modular operations like mixing (AESMC and AESIMC) [13]. ARM also uses SIMD also known as vector instructions[12]. Following in the footsteps of ARM new specialized instructions should be made to make RISC-V more compatible with AES minimizing the amount of instructions required. However, these new instructions should also aim to hide as much information about the data and internal subsystems as possible.

At the moment the biggest challenges that face this project are making sense of the assembly and source code and how to test the AES algorithm on the RISC-V ISS. Only an initial look at the human readable output has been done. It is very possible that not all parts of the assembly and source code will be able to be understood and it is almost certain that there will be challenges in being able to get the compiled AES algorithm to run and be tested on the RISC-V ISS.

5. CONCLUSION

To conclude, when a system designed for security purposes relies on secure software and neglects to design secure hardware it is opening itself up to security risks. Seeing how a secure software algorithm is compiled when on an insecure hardware system clearly shows where security risks lie. For the next part of this project it will be investigated what can be done to improve the security and performance of the RISC-V ISA and in turn hardware implementations of it.

Keywords: Hardware, Security, RISC V, Cryptography, AES, ISA, Compiler, Encryption.

REFERENCES

[1] M. Calderbank, "The RSA Cryptosystem: History, Algorithm, Primes", Math.uchicago.edu, 2007. [Online]. Available: http://www.math.uchicago.edu/~may/VIGRE/VIGRE2007/REUPapers/FINALAPP/Calderbank.pdf. [Accessed: 09-Mar-2019].

[2] "Announcing the ADVANCED ENCRYPTION STANDARD (AES).", Nvlpubs.nist.gov, 2001. [Online]. Available: https://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.197.pdf. [Accessed: 11-Mar-2019].

[3] "ARM Architecture Reference Manual", Static.docs.arm.com, 2018. [Online]. Available: https://static.docs.arm.com/ddi0487/da/DDI0487D_a_armv8_arm.pdf?_ga=2.120333960.1489961675.1551855755-943608715.1551855755. [Accessed: 13-Mar-2019].

[4] A. Waterman, K. Asanovi´ and S. Inc, "The RISC-V Instruction Set Manual, Volume I: User-Level ISA", Content.riscv.org, 2018. [Online]. Available: https://content.riscv.org/wp-content/uploads/2017/05/riscv-spec-v2.2.pdf. [Accessed: 14-Mar-2019].

[5] A. Waterman, K. Asanovi´ and S. Inc, "The RISC-V Instruction Set Manual, Volume II: Privileged Architecture", Content.riscv.org, 2018. [Online]. Available: https://content.riscv.org/wp-content/uploads/2017/05/riscv-privileged-v1.10.pdf. [Accessed: 16-Mar-2019].

[6] J. Wilson, "RISC-V GNU Compiler Toolchain", https://github.com, 2019. [Online]. Available: https://github.com/riscv/riscv-gnu-toolchain. [Accessed: 12-Mar-2019].

[7] Y. Yarom and K. Falkner, "FLUSH+RELOAD: a High Resolution, Low Noise, L3 Cache Side-Channel Attack", Eprint.iacr.org, 2013. [Online]. Available: https://eprint.iacr.org/2013/448.pdf. [Accessed: 20-Mar-2019].

[8] P. Kocher, J. Jaffe and B. Jun, "Differential Power Analysis", cryptography.com, 2015. [Online]. Available: https://42xtjqm0qj0382ac91ye9exr-wpengine.netdna-ssl.com/wp-content/uploads/2015/08/DPA.pdf. [Accessed: 20-Mar-2019].

[9] P. Kocher, "Timing Attacks on Implementations of Diffie-Hellman, RSA, DSS, and Other Systems", cryptography.com, 2015. [Online]. Available: https://42xtjqm0qj0382ac91ye9exr-wpengine.netdna-ssl.com/wp-content/uploads/2015/08/TimingAttacks.pdf. [Accessed: 21-Mar-2019].

[10] W. Dai, "Crypto++® Library 8.2", https://www.cryptopp.com/, 2019. [Online]. Available: https://www.cryptopp.com/. [Accessed: 12-Mar-2019].

[11] J. Callas, L. Donnerhacke, H. Finney, D. Shaw and R. Thayer, "Open PGP Message Format", Ietf.org, 2007. [Online]. Available: https://www.ietf.org/rfc/rfc4880.txt?fbclid=IwAR2Z1Sanzjiwk-KfH4hTLrvqP-ZZiNdzBfKrS6Z50TkvA0GU4kqGvjMJtxO0. [Accessed: 28-Mar-2019].

[12] "ARM Cortex – A53 MPCore Processor Cryptography Extension", http://infocenter.arm.com, Revision: r0p3, 2014. [Online]. Available: http://infocenter.arm.com/help/topic/com.arm.doc.ddi0500e/DDI0500E_cortex_a53_r0p3_trm.pdf. [Accessed: 11-April-2019].

[13] "ARM Cortex – A57 MPCore Processor Cryptography Extension", http://infocenter.arm.com, Revision: r1p3, 2014. [Online]. Available: http://infocenter.arm.com/help/topic/com.arm.doc.ddi0514g/DDI0514G_cortex_a57_mpcore_cryptography_trm.pdf

[14] M. Olivieri, A Cheikh, G Cerutti, A Mastrandrea, F Menichelli "Investigation on the optimal pipeline organization in RISC-V multi-threaded soft processor cores" https://www.researchgate.net, Available: https://www.researchgate.net/publication/320091191_Investigation_on_the_Optimal_Pipeline_Organization_in_RISC-V_Multi-threaded_Soft_Processor_Cores. [Accessed: 21/05/2019].

UTILISING A VEHICLE TESTBED ENVIRONMENT TO DEVELOP DECEPTIVE CAN BUS ATTACKS

Stefan Smiljanic, Charlie Tran, Aaron Frishling, Bradley Cooney, Daniel Coscia, Matthew Sorell
The University of Adelaide, Defence Science and Technology Group
stefan.smiljanic@student.adelaide.edu.au, charlie.tran@student.adelaide.edu.au, aaron.frishling@dst.defence.gov.au, bradley.cooney@dst.defence.gov.au, daniel.coscia2@dst.defence.gov.au, matthew.sorell@adelaide.edu.au

1. INTRODUCTION

The controller area network (CAN) serial communications protocol has been widely used in the automotive industry for almost three decades. Within automotive vehicles, the CAN bus standard facilitates communication between all the electronic control units (ECUs) that manage a vehicle's many functions. Access to the CAN by an adversary's unauthenticated device allows it to listen, broadcast and intercept communications.

Through message manipulation and injection, subtle and deceptive attacks capable of physical, social and financial damage can be developed to highlight the vulnerabilities of the CAN protocol. The victim's dashboard could suggest that the vehicle requires servicing, is malfunctioning, or is travelling at a higher speed than displayed. This research will aim to develop and analyse these subtle and long-term attacks, rather than overtly disabling the engine or brakes.

The attack architectures that will be considered are man-on-the-side (MOTS) and man-in-the-middle (MITM), where the attacking device is physically connected to the CAN bus. The MOTS architecture attaches the attacker's device to the CAN bus directly to read and insert new messages on the network (see Figure 1). In MITM, the attacking device is inserted between an existing ECU and the CAN bus, allowing it to listen, broadcast and intercept messages (see Figure 2).

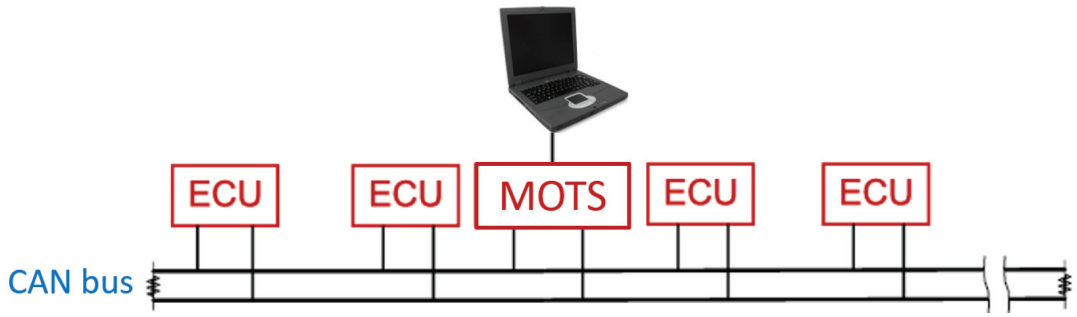


Figure 1. Man-on-the-side attack architecture

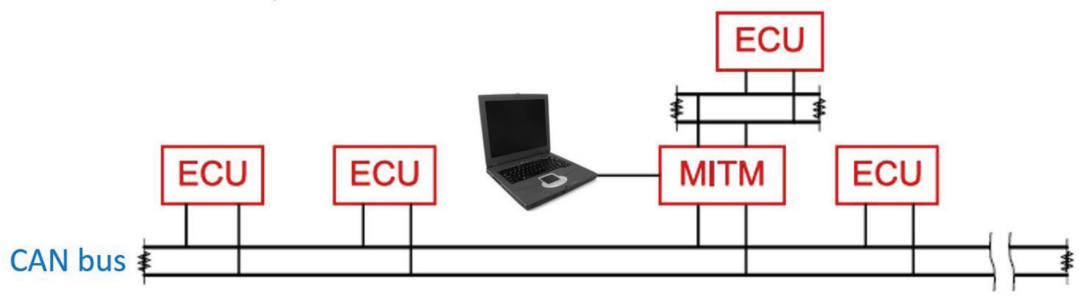


Figure 2. Man-in-the-middle attack architecture

The research presented in this paper will evaluate the effectiveness of subtle attacks as an undocumented cybersecurity threat for automotive CAN systems. These attacks will be developed in both MOTS and MITM architectures in a vehicle testbed environment. The research intends to highlight new threat scenarios and foster the development of new system solutions for automotive cybersecurity.

2. RESEARCH QUESTIONS

- What are the benefits and limitations of a testbed environment for research in automotive security?
- How can the vulnerabilities of the CAN bus protocol be exploited in a testbed environment to perform targeted and deceptive attacks?
- What are the advantages and disadvantages of MOTS compared to MITM attack architectures?

3. ABBREVIATIONS AND DEFINITIONS

- CAN Controller Area Network
- ECU Electronic Control Unit
- MITM Man-in-the-Middle
- MOTS Man-on-the-Side
- OBD On Board Diagnostics

4. RELATED WORK AND MOTIVATION

The exploitation of CAN vulnerabilities is a well-researched field. Articles [1, 2] illustrate the ability to control the dashboard, engine and other systems in a MOTS architecture via the OBD-II port of a working vehicle. Some MOTS attacks can remotely access the CAN bus through Bluetooth and cellular radio [3]. Article [4] analysed the remote attack surface of a 2014 Jeep Cherokee, which demonstrated remote engine and brake control of the vehicle. This was possible due to a vulnerability that allowed malicious messages to be injected onto the vehicle's CAN.

Encryption and authentication protocols that would invalidate MOTS attacks and address the security flaws of existing vehicle CAN networks have been researched in the past [5, 6]. These protocols, however, impact the timing and safety robustness of the CAN standard and are not widely adopted by vehicle manufacturers [7].

The MITM architecture was discussed in article [8], where a security auditing platform for OBD-II devices was created. Using its capability of blocking, forwarding and modifying CAN messages in real-time, the platform could be tested in an adversarial MITM attack between an ECU and the CAN bus.

Previous research has not sufficiently evaluated the impact of deceptive, targeted and subtle attacks on the vehicular CAN bus networks. This paper will address these concerns through threat modelling of both MOTS and MITM architectures in a testbed environment.

5. OBJECTIVES

The research aims to:

- Discover and exploit CAN vulnerabilities to create attacks that are subtle and deceptive
- Demonstrate and evaluate the attacks' abilities to deceive or financially burden the victim
- Implement these MITM and MOTS attack architectures on the testbed
- Evaluate the vulnerabilities, threat scenarios and defence mechanisms
- Highlight the usefulness of the testbed environment in developing CAN bus attacks

The extended objectives are to:

- Create an attack framework for the implementation of the research on other vehicle models and manufacturers
- Weaponise the attack in a small standalone hardware form factor

6. METHOD

The research objectives will be achieved through the following method.

1. Identify the CAN message IDs associated with a vehicle dashboard function by reverse engineering the CAN bus message dumps collected from real-world data.
2. Create a systematic set of experiments to determine effective use of the CAN message in a deceptive attack.
3. Implement and demonstrate the attack in MOTS and MITM architectures in the testbed environment.
4. Create a device capable of executing all of the attacks by connecting:
   a. To the OBD-II port for the MOTS architecture
   b. Between the dashboard and wiring loom for MITM architecture

7. CURRENT AND EXPECTED RESULTS

The testbed used in this research was created by former honours students of the University of Adelaide in 2018 and consists of four main ECUs and a dashboard from a 2016 Mazda2 [9]. By playing back real CAN data onto the testbed, gathered from logging CAN communications while driving, the CAN message ID that related to the dashboard's odometer reading was identified. This is an undocumented finding that was used to develop a targeted MOTS attack by increasing the odometer.

To develop a MITM attack, the real-time message handling architecture presented in [8] shall be used. The aim will be to incorporate speedometer and dashboard indicator attacks in both architectures and develop frameworks for attacking vehicles of other manufacturers. The results will be verified in a more realistic environment using two devices: one to transmit simulated driving messages and the other to read and send malicious messages. A potential complication is how the simulated driving messages will be competing on the CAN bus with the malicious injected messages.

Keywords: Automotive, CAN, Communications, Network, Testbed, Man-in-the-middle, Man-on-the-side, Deception

8. REFERENCES

[1] K. Koscher, A. Czeskis, F. Roesner, S. Patel, T. Kohno, S. Checkoway, D. McCoy, B. Kantor, D. Anderson, H. Shacham, and S. Savage, "Experimental security analysis of a modern automobile", *2010 IEEE Symposium on Security and Privacy*, 2010, pp. 447–462.

[2] R. Currie, "Hacking the CAN Bus: Basic Manipulation of a Modern Automobile Through CAN Bus Reverse Engineering", 2017, [Online]. Available: https://www.sans.org/reading-room/whitepapers/threats/paper/37825

[3] S. Checkoway, D. McCoy, B. Kantor, D. Anderson, H. Shacham, S. Savage, K. Koscher, A. Czeskis, F. Roesner, T. Kohno, "Comprehensive experimental analyses of automotive attack surfaces," *Proceedings of the 20th USENIX Conference on Security*, ser. SEC'11. Berkeley, CA, USA: USENIX Association, 2011, pp. 6–6.

[4] C. Miller and C. Valasek, "Remote Exploitation of an Unaltered Passenger Vehicle", 2015, [Online]. Available: http://illmatics.com/Remote%20Car%20Hacking.pdf

[5] A. Hazem and H.A.H. Fahmy, "LCAP – A Lightweight CAN Authentication Protocol for Securing In-Vehicle Networks", 2012, [Online]. Available: http://eece.cu.edu.eg/~hfahmy/publish/escar2012.pdf

[6] A.I. Radu and F.D. Garcia, "LeiA: A Lightweight Authentication Protocol for CAN", in Proc. Int. Conf. Eur. Symp. Res. Comput. Security, 2016, pp. 283–300.

[7] B. Boldt, "Automotive Security in a CAN", 2017, [Online]. Available: https://www.electronicdesign.com/automotive/automotive-security-can

[8] A. Lebrun and J.C. Demay, "CANSPY: a Platform for Auditing CAN Devices", 2017, [Online]. Available: https://www.blackhat.com/docs/us-16/materials/us-16-Demay-CANSPY-A-Platorm-For-Auditing-CAN-Devices-wp.pdf

[9] L. Oliveira, M. Pfeiffer, T. Taziva, A. Frishling, B. Cooney, D. Coscia and M. Sorell, "The challenges of building a testbed environment for security and vulnerability analysis of internal communication networks in vehicles," *4th Interdisciplinary Cyber Research Workshop*, 2018.

DE-HYPING BLOCKCHAIN-BASED CROSS-BORDER PAYMENT SOLUTIONS: A QUANTITATIVE COMPARATIVE STUDY OF DECENTRALIZED BLOCKCHAIN INFRASTRUCTURES vs. SWIFT GPI[1]

Ahmad Amine Loutfi
Professor. Per Bjarte Solibakke
TEFT-lab, Norwegian University of Science and Technology
ahmad.a.loutfi@ntnu.no

Within an increasingly globalized and distributed economy, we often find ourselves in need of performing cross-border payments. The requirements of modern trade mean that cross border payments need to be efficient, affordable, dependable, secure and traceable.

For as long as banking has been conceived, banks have been acting as a trusted third party through which cross border payments are settled. The current infrastructure underlying current inter-bank cross border payments has long been criticized for being slow, expensive and intractable, and banks have been criticized as the bottleneck within the system. Given its very nature, cross border payments can go over multiple intermediaries before reaching the end beneficiary. Furthermore, this space is heavily regulated and backed by a closed infrastructure. Therefore, the cross-border payment ecosystem is expensive, slow, lacks traceability and real time visibility, is prone to inconsistencies and is characterized by a non-fixation of exchange rate until arrival of funds. The primary infrastructure for settling today's cross border payment is the SWIFT banking network.

While many practitioners have long tried to disintermediate banks and perform Peer2Peer distributed payments, the technology to achieve it was simply not available. That is until recently, when Blockchain was invented. Blockchain has finally given us the long sought-after technology to perform distributed Peer2Peer digital transactions that do not require trusted third parties. Its advent has fundamentally disrupted the foreign exchange ecosystem, as it can allow the disintermediation of traditional financial institutions, as well as provide faster more traceable transactions. Blockchain has been disruptive to the cross-border payment ecosystem, to the point that it has generated a response from virtually everyone in the industry.

In fact, as a response to such a radical technology, one can witness two parallel developments: On one hand, an overwhelming number of Blockchain-based systems offered by different vendors such as Bitcoin, Litecoin and Ripple. On the other hand, banking consortiums improving their traditional Swift infrastructure without the decentralized technology. The most noteworthy project in this space is Swift GPI.

As any radical innovation, there is a lot of hype surrounding Blockchain technology for cross border payment. Different reports and white papers claim the supremacy of different Blockchain solutions. And as a whole, Blockchain is also claiming supremacy over improved traditional cross-border payment infrastructures. Furthermore, many of these performance claims were made as part of a prototype version of the product, which has not stood the time of test and real world at the time of claim making. Finally, the metrics of evaluating different solutions are numerous, such as fees, speed, privacy and traceability, are not always referred to and measured within different studies. However, currently,

[1] Please note that this abstract has not gone through the double-blind peer review.

several of the above discussed infrastructure have advanced from a prototype stage to a production stage and have been deployed long enough for us to be able to objectively capture their performance.

Today's cross-border payment space is cumbersome and hyped, with each technology being praised as the ultimate by its proponents. What further complicates this space is the many metrics across which solutions can be compared, e.g.: speed, integrity, confidentiality, fees, traceability, visibility, and integration.

This paper aims to perform a quantitative comparative study between different deployed Blockchain-Based cross-border payment solutions (Bitcoin, Litecoin, Ripple), SWIFT and SWIFT GPI. The comparison will be performed across different performance metrics such as: speed, cost, fees, traceability, visibility and security.

• RESEARCH QUESTIONS

RQ. How do Blockchain based cross-border payment solutions (Bitcoin, Litecoin, Ripple) perform against traditional Swift infrastructure and SWIFT GPI?

This research question is a composed one. This is why it needs to be subdivide. In fact, the supremacy of different cross-border payment solutions can de dependent on the performance metrics. Hence, the two sub-research questions which need to be considered are:

- SRQ1. What are the relevant cross-border payment performance metrics that are most relevant to the end-consumers?
- SRQ2. How do Blockchain based cross-border payment solutions (Bitcoin, Litecoin, Ripple) perform against traditional Swift infrastructure and SWIFT GPI, the improved traditional Swift infrastructure, across the different identified metrics?

• METHODOLOGY AND DATA

This paper will follow a quantitative comparative methodology and will follow a multiphase approach:

Phase 1-Defining the metrics: to answer the first research question, primary research will be conducted in the form on interviews and surveys with different financial service providers (e.g. Banks) to define the metrics that can be used as a base comparison for the different cross border payment solutions.

Phase 2-Data collection: The initial litterateur review conducted reveals that datasets about traditional based cross-border payments is publicly available. We also would like to use data from Norwegian banks that are using traditional SWIFT infrastructure: Sparebanken Møre, and Sparebanken 1 for relevancy and DNB which has improved it traditional infrastructure to SWIFT GPI.

Blockchain performance data will be collected from systems, which are in production and can be readily used. Data will be processed with a python-based framework and analyzed using a time series method to achieve a complete comparison.

Keywords: Blockchain Technology, Cross-Border Payment, Fintech

REFERENCES

- Innopay (2015). Unlocking opportunities in the API economy. Retrieved from https://www.innopay.com/en/publications/unlocking-opportunities-api-economy
- Kakavand, H., Kost De Sevres, N., & Chilton, B. (2017). The Blockchain Revolution: An Analysis of Regulation and Technology Related to Distributed Ledger
- Mathis FJ, Cavinato J (2010) Financing the global supply chain: growing need for management action. J Thunderbird International Business Review 52(6):467–474
- Tasca, Pelizzon, & Perony, (2016). Banking Beyond Banks and Money – A Guide to Banking Services in the Twenty-First Century. Springer

SESSION 6: DIGITAL FORENSICS 2

Session moderated by Prof MATTHEW SORELL,
University of Adelaide

Mr Ben Agnew,
"AN OVERVIEW OF INFORMATION SECURITY CONCEPTS AND THEIR RELEVANCE TO DIGITAL FORENSIC EVIDENCE PROCEDURES",
University of Adelaide

Mr Glenn Walsh & Mr Jimmy Tang,
"FORENSIC APPLICATIONS OF 3D SCANNING",
University of Adelaide

Mr Luke Jennings,
"IDENTIFYING PATTERNS AND ACTIVITIES FROM IPHONE AND APPLE WATCH STEP-COUNT DATA FOR USE IN A DIGITAL INVESTIGATION",
University of Adelaide

Ms Joanna Rose del Mar,
"AUTOMATED PHOTO CATEGORIZATION FOR DIGITAL FORENSIC ANALYSIS USING A MACHINE LEARNING-BASED CLASSIFIER",
Tallinn University of Technology

AN OVERVIEW OF INFORMATION SECURITY CONCEPTS AND THEIR RELEVANCE TO DIGITAL FORENSIC EVIDENCE PROCEDURES

Ben Agnew, Matthew Sorell, and Cate Jerram
University of Adelaide
benjamin.agnew@adelaide.edu.au, matthew.sorell@adelaide.edu.au, cate.jerram@adelaide.edu.au

INTRODUCTION TO DIGITAL EVIDENCE

Electronic evidence is electronic data that "has the potential to make the factual account of either party more probable or less probable than it would be without the evidence"[1]. This data can be stored and/or manipulated on a computer system or electronic device or transmitted by a communications system. Looking at digital forensic evidence, we restrict ourselves to looking at digital (binary) data that is collected at a crime scene, analysed and presented in court. It comes in a number of different forms such as CCTV video, crime scene photography, phone call logs and seized equipment such as computers and hard drives. Like physical evidence, it is collected at a crime scene, transported to a lab where it is stored and then processed. Throughout this process, the chain of evidence needs to be preserved, which means keeping a strong record of who was in possession of the evidence and ensuring that it hasn't been maliciously tampered with. In the context of physical evidence there are procedures to ensure the chain of evidence is preserved[2] [3]. These rely heavily on physical access control systems and paper audit records.

Once digital evidence is at the lab, it can be protected with traditional security methods such as physical access control. However, transporting evidence in the form of digital data requires the use of a physical device such as SD cards and USB drives. The data is transferred to the device at the crime scene and then transported to a lab where it is transferred again to a data storage system. This process can make preserving the chain of evidence difficult.

There are very few standards and guidelines in this area. *The Scientific Working Group on Digital Evidence Framework*[4] simply requires that procedures be established for the handling of evidence. While access controls and logs are required, little detail is given on how this would be achieved or what security it would provide.

ISO/IEC 17025[5] defines the legal requirements for digital forensic labs in the UK. However there is some doubt over how well this standard applies to digital forensic procedures with many experts in the field suggesting that it is not relevant[6].

The *Australia and New Zealand Guidelines for Digital Imaging Processes*[7] describe the procedures used for digital forensic photography. Whilst some technological security measures are used such as write-once memory, it still relies heavily on a paper audit record of who did what with the evidence.

REQUIREMENTS FOR A DIGITAL EVIDENCE CONTAINER

There is a need for a 'secure' digital evidence container/device which can be used to collect and transport digital evidence from the crime scene. By investigating existing procedures for physical evidence[2], these requirements for a digital evidence device were formed:

- Tamper evident. Any attempt to alter or erase the data on the device would be detectable.
- Un-forgeable. Any attempt to swap out the device for a fake one should be detectable.

- Clean. Before digital evidence is loaded onto the device, it needs to be clean and not contaminated with any digital data from a previous use.
- Offline. During evidence collection and transport, there may not be any network connection available.

OBJECTIVES IN INFORMATION SECURITY

We then looked at existing technology used in information security, which revealed many commonalities between the aims and requirements for both digital forensics and information security. The *Handbook of applied cryptography*[8] defines four core objectives of information security:

- Confidentiality. Ensuring the data is only available to those that are authorised to have it.
- Data integrity. Ensuring the data hasn't been altered or deleted.
- Authentication. Ensuring that an entity is who they claim to be and ensuring that a claimed data source is correct.
- Non-repudiation. Preventing an entity from denying a previous action or commitment.

The main objective when dealing with digital evidence is preserving data integrity. However authentication and non-repudiation are also important.

TECHNOLOGY IN INFORMATION SECURITY

Many technologies and methods have been designed to address these aims of information security. By reviewing the field, we have categorised existing technologies into four distinct concepts that underpin how security is achieved.

CRYPTOGRAPHY

While initially aimed at the confidentiality problem, cryptography can be used to address all four objectives[8]. Cryptography relies on a secret key being kept secret from any possible attacker. It is also important to ensure that the secret key cannot feasibly be guessed by an attacker trying all possible keys. The encryption and decryption processes are carried out used specialised mathematical algorithms such as AES, RSA and Elliptic Curve. These algorithms are designed such that the cipher-text cannot be decrypted without knowledge of the key. Cryptography technology has been used in computing and internet applications for several decades. Some well know uses include SSL/TLS, PGP email and digital signatures. As a result of this widespread use, a large amount of research work has been carried out on finding weaknesses and improving the overall security of cryptography systems.

WIDELY WITNESSED

This idea relies on a relatively simple concept: once something has been witnessed by a sufficiently large number of people, it is impossible to change it without someone noticing the change. Implementing this concept on a large scale in information security has only been realized fairly recently with the invention of blockchain[9]. A 'block' (containing the data to be protected) is created and shared with many other nodes in the network who each witness the new block (and if valid, accept it into the blockchain)[10]. If someone tries to tamper with a block, or if multiple competing blocks are created, the nodes will together reach a consensus on which block is the true and correct one. This is done using a consensus protocol, such as Proof of Work, which is designed to make creating false blocks very expensive. However for this process to work there needs to be a network connection to communicate with other nodes, which means blockchain can't be used for offline applications. Blockchain is mainly aimed at protecting data integrity however it also has uses in non-repudiation applications.

HARDWARE/PHYSICAL SECURITY

It is possible to design electronic hardware that is resistant to malicious attackers. Depending on the aim and application, there are many different designs and methods used. The simplest example of this is write-once memory, which by design makes tampering with the data difficult. There are a range of products sold as "Hardware Security Modules" that are designed to securely store cryptographic keys and carry out cryptographic functions in a tamper-proof environment[11].

MANUFACTURING VARIANCES/DEFECTS

When electronic components are manufactured, there are tolerances in the specifications that the component must meet. As a result there are small variations between each component manufactured and these can be a source of useful information for security purposes (in particular, forensic purposes). This concept is most useful in chips with millions of components such as digital camera image sensors. By measuring these manufacturing variations it is possible to create a unique "digital fingerprint" for the device[12]. This concept can be particularly useful for protecting against forged devices since it allows a method to uniquely identify every device manufactured. It is also possible to design chips to exploit these manufacturing variations to create random numbers and random functions such as the concepts used by Physical Uncloneable Functions[13].

CONCLUSION

For many decades, information security has relied of cryptography alone and there are countless examples where this has failed. In recent years the use of widely witnessed technology (in particular blockchain) in conjunction with cryptography has become very widespread and as a result, security has been substantially improved. Hardware security concepts have been used in many specific applications (usually in conjunction with cryptography). Exploiting manufacturing variations however is still a relatively new area and while a reasonable amount of research has been done, there is not a lot of commercial use outside of specialized forensic applications. It appears that the best security is achieved when ideas from several of the four concepts discussed above are combined together.

Keywords: Information Security, Digital Forensics, Digital Evidence

REFERENCES

[1] S. Mason and D. Seng, *Electronic Evidence, Fourth Edition*. University of London, 2017.

[2] *Property Room Standards*, International Association for Property and Evidence, 2007. [Online]. Available: https://web.archive.org/web/20080204172810/http://www.iape.org/Standards 7-03/index.htm

[3] *Western Australia Police Property Management Practices*, Western Australian Police and Corruption and Crime Commission, 2005. [Online]. Available: https://www.ccc.wa.gov.au/sites/default/files/Western%20Australia%20Police%20Property%20Management%20Practices.pdf

[4] *SWGDE Framework of a Quality Management System for Digital and Multimedia Evidence Forensic Science Service Providers*, Scientific Working Group on Digital Evidence, 2017. [Online]. Available: https://www.swgde.org/documents/Current%20Documents/SWGDE%20Framework%20of%20a%20Quality%20Management%20System%20for%20Digital%20and%20Multimedia%20Evidence%20Forensic%20Science%20Service%20Providers

[5] "ISO/IEC 17025 – General requirements for the competence of testing and calibration laboratories," International Organization for Standardization, Standard, 2017.

[6] P. Beardmore, G. Fellows, and P. Sommer, "UK ISO 17025 Digital Forensics Survey April 2017: Results," 2017.

[7] *Australia and New Zealand Guidelines for Digital Imaging Processes*, Australian and New Zealand Policing Advisory Agency, 2013. [Online]. Available: http://www.anzpaa.org.au/ArticleDocuments/282/2013%20Australia%20and%20New%20Zealand%20Guidelines%20for%20Digital%20Imaging%20Processes.pdf.aspx

[8] S. Vanstone, A. Menezes, and P. v. Oorschot, Handbook of Applied Cryptography. CRC Press, 1996.

[9] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," 2008.

[10] I. Grigorik, "Minimum viable block chain," 2014. [Online]. Available: https://www.igvita.com/2014/05/05/minimum-viable-block-chain/

[11] Ponemon Institute LLC, "Hardware Security Modules Global Market Study," 2014. [Online]. Available: https://www.ponemon.org/local/upload/file/HP%20Atalla%20Report%20FINAL%202.pdfg

[12] J. Lukas, J. Fridrich, and M. Goljan, "Digital Camera Identification From Sensor Pattern Noise," 2006.

[13] C. Herder, M.-D. M. Yu, F. Koushanfar, and S. Devadas, "Physical Unclonable Functions and Applications: A Tutorial," 2014.

FORENSIC APPLICATIONS OF 3D SCANNING

*Jimmy Tang**
University of Adelaide
a1687296@student.adelaide.edu.au

Glenn Walsh
University of Adelaide
a1686736@student.adelaide.edu.au

Matthew Sorell
University of Adelaide
matthew.sorell@adelaide.edu.au

Richard Matthews
University of Adelaide
richard.matthews@adelaide.edu.au

2019

1 INTRODUCTION

Digital forensics has been defined by the Australian Federal Police as ”obtaining, analysing and presenting on data recovered from computers, electronic devices and other digital sources”[1]. However, in the context of this paper, we will simply be looking at obtaining crime scene data using digital devices and methods, not the recovery of digital data.

With the advancement of technology, collecting and visualising physical evidence is inhibited by the limitations of traditional evidence collection, for example: photographs, casting footprints using plaster, and performing physical autopsies. Newer cutting-edge technology allows for 3D models to be generated by capturing real world scenes can provide a more convenient method of collecting, and analysing 3D evidence. This can be done through 3D scanning shoe prints, crime scenes, and performing virtual autopsies, leading to more convenient and cost-effective evidence collection. In this paper, we introduce and evaluate 3D imaging technology such as photogrammetry, and laser scanning against traditional methodologies which are used in crime scene evidence collection. We expect, that although photogrammetry encounters limitations that 3D laser scanning doesn’t, it still has its place as a lower cost method of visualising and measuring physical evidence.

2 MATERIALS AND METHODS

2.1 GROUND TRUTH

The ground truth used in these scans are simple objects with known sizes that are placed in the scanning scene to use as a scaling reference for measuring shoe prints. Figure 1 is a prototype ground truth object that can be scaled accordingly and printed off using a 3D printer. The prototype aims to encode the known size as a part of the structural data, rather than reading off a measuring tape or ruler which is encoded in the texture data.

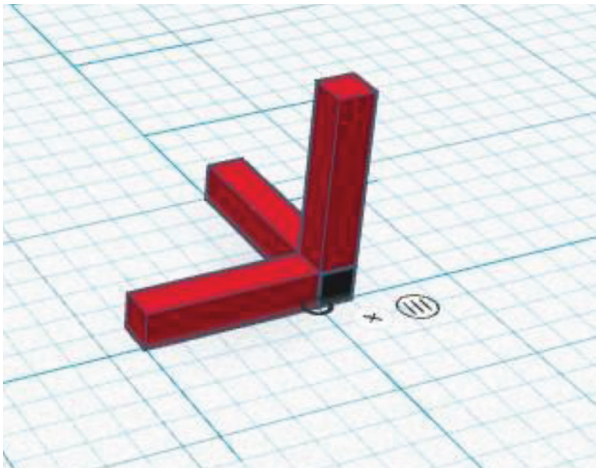


Figure 1. ground truth object with known sizes

2.2 MEASUREMENT TECHNIQUE

There are two major measurement techniques this paper will address. The first is a distance scaling method, where a ground truth object, or a measuring tape will be placed in the scene in order to calibrate measurements. The second is to indirectly measure objects through determining invariant features such as ratios between each dimensional length.

2.3 SHOE AND FOOTPRINT IMPRESSIONS

Shoe print impressions require a surface capable of holding impression; either sand, soil, or snow and anything else that can leave an imprint. Taking a cast of the footprint may introduces foreign material into the scene and is a time-consuming process[2]. 3D scanning and imaging is a means to reconstruct the scene without disturbing the evidence.

3 THE TECHNOLOGY BEHIND 3D SCANNING AND IMAGING

3.1 3D SCANNING AND IMAGING

General 3D imaging data is referred to in literature as a point cloud and is made of 4 components $P_i = (x_i, y_i, z_i, f_i)$[3] where Pi represents the point cloud, (x_i,y_i,z_i represents where a particular point is in 3D space, and f_i represents a value at this point, which could be its translucency, or reflectivity.

3.2 PHOTOGRAMMETRY

Photogrammetry is a low cost, highly computationally expensive technique for creating a 3D model. This method of creating 3D models uses a series of photos to generate a point cloud of which can be texturised and then parsed into a 3D model using software. The model can be rendered on programs on 3D computer graphics software such as Blender, a free and open source 3D creation suite[4]. Figure 2 shows a high-level overview of how photogrammetry works by generating a point cloud[5].

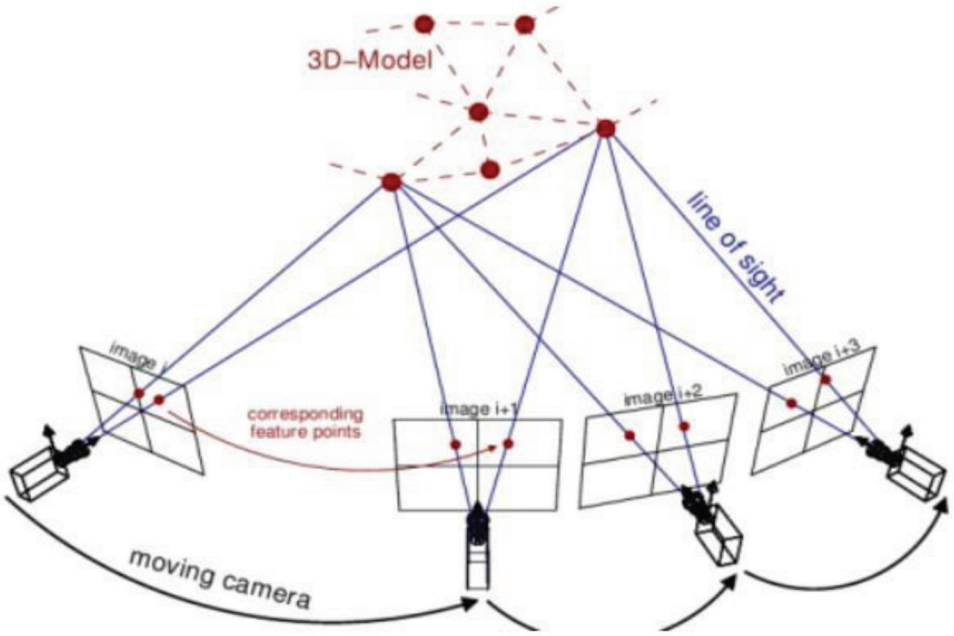


Figure 2. How Photogrammetry Works[6]

Videogrammetry uses the same concepts as photogrammetry, but uses frames of a video to construct an image set parsed to a standard photogrammetry pipeline. To facilitate the slicing of frames, software such as ffmpeg can be used[7]. The ffmpeg command is as follows:

ffmpeg -i ./vid_in.mp4 -vf fps=2 ./frames_out/frame%04d.jpg -hide_banner

Burst shots can also be used to capture the scene at a seemingly faster rate, however, this can very easily blur the photos taken, prolonging the processing stage due to invalid images.

3.3 OBTAINING THE 3D MODEL

3D models have file extensions such as obj, stl, and ply amongst others and contain point clouds and meshes. A model of a box generated using photogrammetry with an image set of 112 images is depicted in Figure 3 using Meshroom, a photogrammetry pipeline front-end to the AliceVision framework[8]. Other commercial or educational methods of capturing 3D Models include Scandy Pro[9] used on the iPhone XR, and dotproduct3d scan[10] for the Intel Real Sense D435i depth camera.

4 PRELIMINARY RESULTS

Tables 1 and 2 both refer data obtained from the 3D model shown in Figure 3. Examining Table 1 we find that the squared difference between the real measurements and the virtual measurements are in the order of negative 6 while the difference in ratios are in the order of negative 3.

coordinates	BU	m	cm	Real (cm)	squared difference
x	0.858811	0.1265615	12.65615	12.5	2.44E-06
y	0.872784	0.1286207	12.86207	12.5	1.31E-07
z	0.699446	0.1030762	10.30762	10.5	3.70E-06

Table 1. Distance Scaling Method

ratio labels	virtual	real	squared difference
x/y	0.983990311	1	0.00025631
x/z	1.227844608	1.19047619	0.001396399
y/z	1.247821848	1.19047619	0.003288524

Table 2. Preservation of Ratios

Measurements made in a virtual environment i.e. Blender's internal measurements (abbreviated as BU in Table 1) are not accurate to the real world due to calibration issues. However, we have found that ratios determined in a virtual environment are preserved well compared to the physical environment. As long as our scanning method can be written to disk as one of these 3D model files, the measurements can be done.

5 DISCUSSION

The techniques in the above subsections are to be applied to forensic objects and scenes of interest such as shoe prints, blood patterns, and bones. Figure 3 is the textured mesh of a box while Figure 4 shows a point cloud representation of a shoe print on kinetic sand. Assuming that we have ground truth objects ready at our disposal, performing the act of taking photos for photogrammetry or using a scanning camera that is already on your person will only take a few seconds for each shoe print compared to creating a casting mixture and applying it to a shoe print.

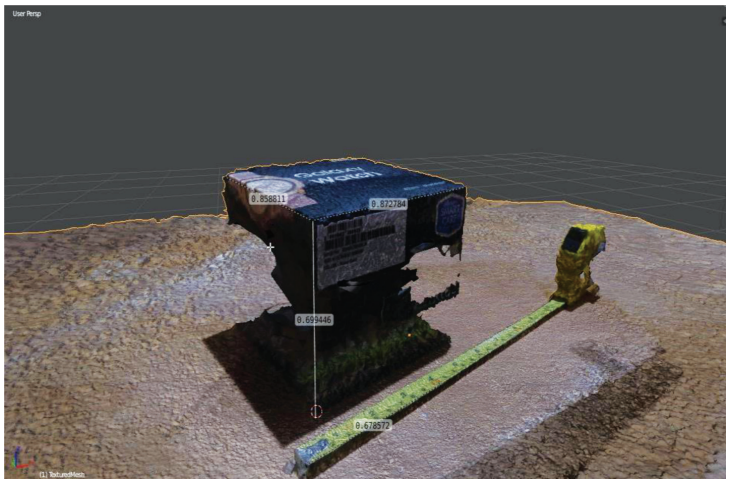


Figure 3. Box scanned using Photogrammetry with measuring tape as ground truth, rendered and measured in Blender

Issues arose with the photogrammetry software Meshroom due to the fact that there is no verification of the dataset before performing the scan. Often there would be issues with the dataset where local cameras aren't able to be found and the entire process would be aborted leaving with the experiment without a usable 3D model. Whenever this happens, the entire process would have to be repeated with a new image set or to cut out problem images. Ways to mitigate this risk for easier capture of forensic scenes and objects is to have a method of verify whether the data is usable or not during time of capture. This is seen in the Scandy Pro software capture run on the iPhone XR of which the system would stop capture if object tracking is failed.

6 CONCLUSION

As the squared difference in both Tables 1 and 2 is significantly small (in the order of at least 10−3), it is reasonable to consider that these methods could be applied to more forensic purposes. These results are preliminary and consist of only one data point under a very controlled environment, but we expect that these methods can be extended into work on shoe prints, blood patterns, and bones. Although preliminary, these methods shall be repeated in order to obtain a statistically sound result and compared against different scanning methods such as photogrammetry, laser scanning, stereoscopic cameras, and structured light systems.

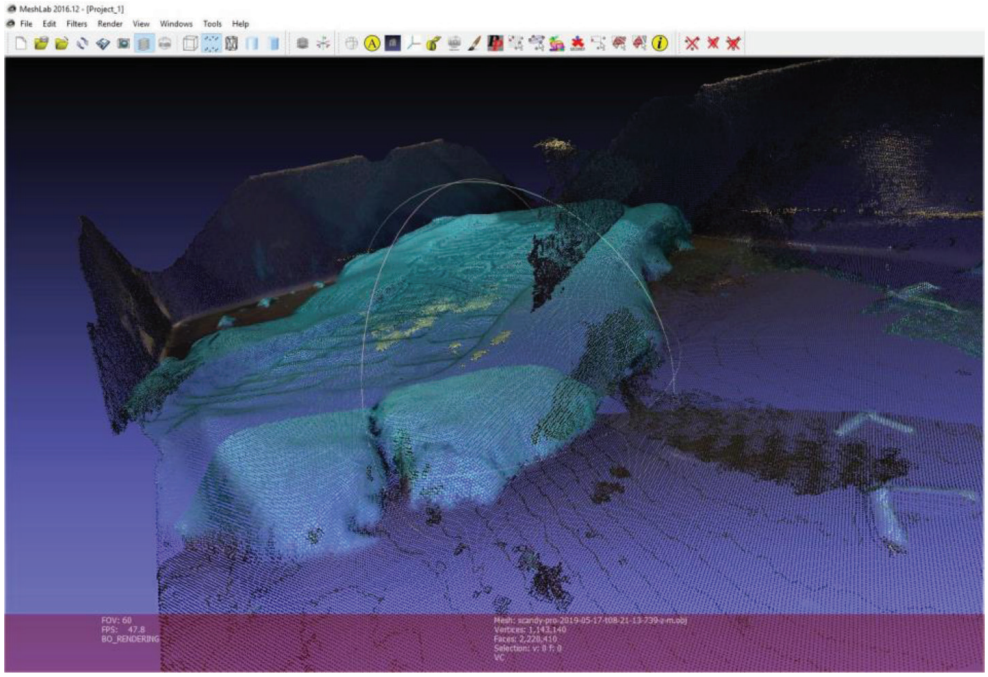


Figure 4. iPhone XR 3D scan using Scandy Software: Point cloud representation as rendered in MeshLab.

Keywords: 3D Scanning, 3D Modelling, Forensics, Photogrammetry, Shoe Casting, Structured Light Systems, Meshroom, Blender, AliceVision

REFERENCES

[1] Afp. Accessed on 22nd May 2019. [Online]. Available: https://www.afp. gov.au/what-we-do/crime-types/cybercrime/digital-forensics

[2] U. Buck, N. Albertini, S. Naether, and M. J. Thali, "3d documentation of footwear impressions and tyre tracks in snow with high resolution optical surface scanning," *Forensic Science International*, vol. 171, no. 2, pp. 157–164, 2007. [Online]. Available: http://www.sciencedirect.com/science/ article/pii/S0379073806006712

[3] J. Geng, "Structured-light 3d surface imaging: a tutorial," Adv. Opt. Photon., vol. 3, no. 2, pp. 128–160, Jun 2011. [Online]. Available: http://aop.osa.org/abstract.cfm?URI=aop-3-2-128

[4] Blender. Accessed on 24 May 2019. [Online]. Available: https://www.blender.org/

[5] S. Colwill, "Low-cost crime scene mapping: Reviewing emerging freeware, low-cost methods of 3d mapping reviewing emerging freeware, low-cost methods of 3d mapping evidence," 2016.

[6] H. H. Bartholomeus. (2019) Msc thesis subject: Importance of camera calibration for uav-based photogrammetry. [Online]. Available: https://www.wur.nl/en/article/MSc-thesis-subject-Importance-of-camera-calibration-for-UAV-based-photogrammetry.htm

[7] ffmpeg. Accessed on 21 May 2019. [Online]. Available: https://ffmpeg.org/

[8] Alicevision. [Online]. Available: https://alicevision.github.io/

[9] Scandy pro. [Online]. Available: https://www.scandy.co/

[10] Dotproduct 3d. Accessed on 24 May 2019. [Online]. Available: https://www.dotproduct3d.com/

IDENTIFYING PATTERNS AND ACTIVITIES FROM IPHONE AND APPLE WATCH STEP-COUNT DATA FOR USE IN A DIGITAL INVESTIGATION

Luke Jennings
The University of Adelaide
luke.jennings@adelaide.edu.au
Matthew Sorell
The University of Adelaide
matthew.sorell@adelaide.edu.au

INTRODUCTION

As technology advances, there is an increasing amount of electronic evidence[1]. iPhones and Apple Watches have in-built technology that record the user's fitness data[2] such as step-counts, which can then be read in the Apple Health app. This research aims to determine if step count data obtained from iPhones and Apple Watches can be utilised to create a user template such that it can be used to identify abnormal behaviours and patterns for use in a Digital Investigation. Similar data is utilised to established if two sets of step-count data are correlated with one another. For the purpose of this research, the step-counts are the primary focus.

BACKGROUND

In[3] health data had been extracted from fitness devices. The aim was to estimate the time of death of a victim in a criminal investigation, as well as using the captured health data as a forensic tool for other purposes[3]. Of particular importance is the data extraction method used for Apple products, which is used in this research. The method details the use of the iPhone Health app and how to export it for use on a computer[3].

The health app on the iPhone can be used to read the steps measured by the in-built pedometer[2]. Additionally, an Apple watch can be paired with the iPhone and the data measured by the watch can also be viewed on the Health app[2]. There are quite a few investigations out there that compare the accuracy of the step count from one smartwatch to the next[4] [5], but none that compare the accuracy of the watch to the paired iPhone.

The Health app is available on all Apple devices[2], and data measured by your device is automatically added to the app[2]. In the app, you can see logs of all your activities such as step-counts, heart rate and other metrics[2] [3]. The information found in the app is backed-up in iCloud and iTunes and is encrypted[2]. This information can be exported in multiple ways: through the Health app (this research's method), through the iCloud or iTunes backup or through chip-off[3]. The different methods require knowledge of the user's phone password, or their Apple ID[3].

The data in question is from two subjects using both iPhone and Apple Watch. Subject 1 has supplied data going back to the 19th April 2017 while subject 2 has supplied data for the period of the 26th to the 28th of December 2018. Both devices are capable of measuring step-counts, distance and flights of stairs climbed. However, the watch has a much larger range of measurement tools and can also measure quantities such as active energy, basal energy and pulse rate.

METHODOLOGY

The health data records were exported from the app as a xml file, converted into csv files through[6] that contain the different measurements from the device. The step-count csv file was then imported into Matlab for analysis.

The step-counts for both users, on both devices for the period between the 26th and 28th of December 2018 were accumulated over 24-hour periods. The results are displayed as plots of the distribution of the accumulation of steps as well as the Kullback-Leibler distance[7], which is defined as:

$$\sum p \log \frac{p}{q}$$

Where p and q are the points being measured.

It is assumed that there are uniformly distributed and accurate timestamps for each step-count period with both devices.

A DISCUSSION OF RESULTS

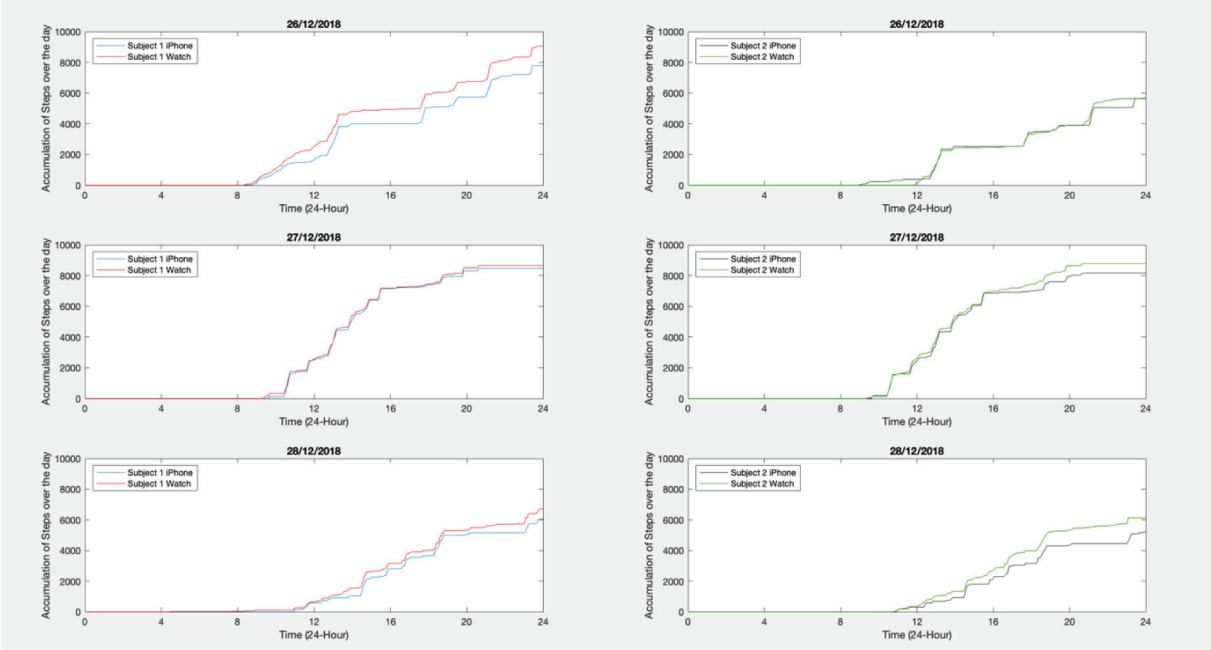


Figure 1. Accumulation of steps comparison

Figure 1 above shows the plots of the accumulation of steps over the day. The left-hand side shows Subject 1's steps, plotting the accumulation of steps between iPhone and Watch, and the right-hand side does the same for Subject 2's data. Although the total steps end up being slightly different, it can be noted that the shapes of the accumulation between Watch and iPhone are very similar. It is also noticeable that the shapes of the accumulation between Subject 1 and Subject 2 on respective days are also similar. For further analysis, let's consider Subject 1 and Subject 2's steps for the 27th, shown below as an example.

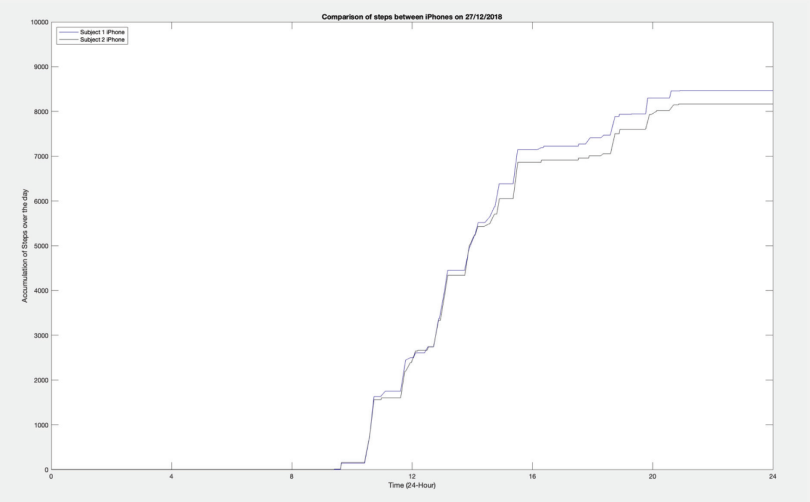


Figure 2. Comparison of users on the 27th

On the 27th December 2018, the accumulation of steps for both users have an extremely similar shape, despite differing steps counts over the day. The differing step counts are due to people having different paces and stride lengths. The similar shape and vertical displacement between both user's steps indicate that these two people were walking together or happened to be walking in the same direction at the same time for most of the day.

Next, the step-counts are compared using the Kullback-Leibler distance. With the distance measurement, a result of 0 indicates the lowest distance between itself, and a large numerical value indicates extremely large distance. As particular examples, we consider the distance matrix for Subject 1's data between iPhone and Watch, and the distance matrix comparing iPhone step-counts between Subject 2 and Subject 1.

	26th iPhone	**26th Watch**	**27th iPhone**	**27th Watch**	**28th iPhone**	**28th Watch**
26th iPhone	0.00	5.51	13.53	3.67	16.26	14.64
26th Watch	3.92	0.00	14.04	14.54	16.05	15.16
27th iPhone	12.44	13.11	0.00	5.03	13.31	13.59
27th Watch	11.82	13.15	2.74	0.00	14.10	13.69
28th iPhone	17.25	16.46	11.59	15.08	0.00	6.22
28th Watch	16.67	16.31	12.51	15.45	4.65	0.00

Figure 3. Distance of Subject 1's step-counts

	Subject 2 26th	**Subject 1 26th**	**Subject 2 27th**	**Subject 1 27th**	**Subject 2 28th**	**Subject 1 28th**
Subject 2 26th	0.00	1.69	14.09	12.83	17.53	16.06
Subject 1 26th	5.25	0.00	14.21	13.53	17.34	16.25
Subject 2 27th	13.09	11.76	0.00	1.87	14.77	13.78
Subject 1 27th	13.03	12.43	2.36	0.00	14.10	13.30
Subject 2 28th	17.31	17.31	12.56	11.60	0.00	3.42
Subject 1 28th	17.30	17.25	12.34	11.58	4.35	0.00

Figure 4. Distance of iPhones between Subject 2 and Subject 1

In Figure 3, there is low distance between iPhone and Watch on the same day indicated by the low numerical values. By contrast, there is a large distance between iPhones on different days, between Watches on different days and between iPhone and Watch on differing days. This result, along with the plots in Figure 1 give us a measure of accuracy needed to confidently compare different users who use multiple devices in order identify patterns and behaviours.

In Figure 4 there is low distance between Subject 1 and Subject 2 on all 3 days, consistent with the earlier discovery from Figure 2 that Subject 1 and Subject 2 were walking together. There is also a large distance between Subject 1 and Subject 2 on differing days, i.e. Subject 1's 26th and Subject 2's 27th.

It is possible that the two subjects may have walked a similar path at the same time, but in other locations. To further verify whether or not the subjects travelled together, additional data such as cell tower records would be needed. In this case, it was already known beforehand that the two subjects were together during the time period of the 26th to the 28th.

As of now, the results are quite consistent with the research question. Certain patterns can be determined from step-count data such as to whether or not, two people happen to be together at a particular time on a particular day.

Since Subject 1 has data dating back over 2 years, the next goal of this research is to create a template which represents their median step-count for a particular day of the week, for example Monday. We can then compare a particular Monday with the template and analyse how much variation is there and identify if any typical or abnormal behaviour has occurred.

Keywords: Step-Count, Electronic Evidence, iPhone, Apple Watch, Health Data, Pattern, Behaviour, Digital, Investigation, Fitness Data

REFERENCES

[1] B. Hooke, 27/Feb/2018, "*On the Horizon: Tech Trends Impacting Law Enforcement Investigations*" [Online], Last Accessed 08/APR/2019, Available: https://www.policeone.com/police-products/investigation/computer-digital-forensics/articles/471631006-On-the-horizon-Tech-trends-impacting-law-enforcement-investigations/

[2] Apple, 30/Nov/2018, *"Use the Health App on your iPhone or iPod touch"* [Online], Last Accessed 08/APR/2019 Available: https://support.Apple.com/en-au/HT203037

[3] Y. Li, *"Forensic Investigation of Fitness Devices"*, Honours Thesis, Dept. Elect. Eng., The University of Adelaide, Australia, 2017.

[4] D. Graziano, 19/May/2015, *"How Accurate is the Apple Watch's Step Counter and Distance Tracking?"* [Online], Last Accessed 9/APR/2019, Available: https://www.cnet.com/news/smartwatch-step-counter-and-distance-tracker-accuracy/

[5] Y. Bai, P. Hibbing, C. Mantis & G. Welk (2018) Comparative evaluation of heart rate-based monitors: Apple Watch vs Fitbit Charge HR, Journal of Sports Sciences. 36:15, 1734–1741, doi:10.1080/02640414.2017.1412235

[6] *"tdda/Applehealthdata"*, GitHub, 2017. [Online], Last Accessed 09/APR/2019, Available: https://github.com/tdda/Applehealthdata

[7] Kullback, S; Leibler, R. A. On Information and Sufficiency. Ann. Math. Statist. 22 (1951), no. 1, 79–86. doi:10.1214/aoms/1177729694. https://projecteuclid.org/euclid.aoms/11777296994

AUTOMATED PHOTO CATEGORIZATION FOR DIGITAL FORENSIC ANALYSIS USING A MACHINE LEARNING-BASED CLASSIFIER

Joanna Rose Castillon del Mar
Tallinn University of Technology
joadel@taltech.ee

Supervisors:
Prof Dr. Hayretdin Bahşi (Tallinn University of Technology, Tallinn, Estonia)
Prof Dr. Leo Mršić (Algebra University, Zagreb, Croatia)
Krešimir Hausknecht (INsig2 d.o.o., Zagreb, Croatia)

INTRODUCTION

Categorizing the content of seized devices for potential evidentiary value, particularly photos, is inherent in a forensic investigation. The increasing amount of data that needs to be processed has outpaced the effectiveness of traditional digital forensic methods. The need for automation becomes even more apparent with limited time, human and financial resources. In a forensic acquisition, thousands of photos are often extracted and analysed for processing. Without access to commercial tools, the forensic examiner must manually review these photos by hand to search for artefacts. The trivial task of categorizing photos manually consumes the examiner's time, especially when there is a substantial number of devices acquired waiting to be processed.

The automatic categorisation use-case can be re-formulated as a machine learning classification problem that can be extended in the forensics context. The non-trivial choice of gun as the output label is implemented in this study as the label must hold significant importance in a typical forensic investigation.

The performance of neural networks in solving image classification is unparalleled; some already exceeding human-level accuracy[1]. These state-of-the-art neural networks are built from powerful computer vision models pre-trained in various categories from benchmarked datasets such as ImageNet[2], and such models are freely available online.

Why design science? As the research study aimed to create a digital forensic prototype that is innovative, purposeful and evaluated in the proper scientific process, the Design Science methodology is applicable and therefore, has been adapted.

Why open-source? Open-source code had the significant advantage of being validated and substantiated[3]. As such, the drive for this research project is fueled by the idea of working with open-source technologies that can be easily cross-examined and validated by software experts.

HYPOTHESIS

This research hypothesizes that an open-source tool can be created that could:

- leverage a pre-trained neural network model with pre-learned capabilities from a non-forensic dataset such as ImageNet, and
- this resulting tool is open-source, usable, easy to use and effective, and
- can be deployed to aid real-life forensic investigations
- using the design science methodology

MODEL EVALUATION AND SELECTION

The study is interested in selecting the best-performing model from a set of pre-trained models downloaded from Keras, i.e., Xception[4], VGG16[5], ResNet50[6] and Inception[7].

These four models are evaluated using the test dataset from Olmos et al.'s Handgun detection paper[8]. One of the advantages of choosing an existing dataset is to remove any possibility of error that we could have introduced due to misclassification during ground truth labelling.

The use of accuracy alone in evaluating performance is insufficient and even misleading. This study prioritises recall rate, or the ratio of correctly classified gun pictures from actual gun pictures, as the study aims to detect as many gun pictures as possible. Also, in forensic applications, unbalanced datasets (the number of true negatives is significantly higher than the true positives) are the norm. With such highly skewed distributions, the Matthew's Correlation Coefficient (MCC) measure is more robust [9], and is, therefore, included in this paper along with precision, false positive rate, processing time, and classification accuracy.

These models are ranked against each other (1 as best, 4 as worst) and the best model with the lowest score, is selected. Tabulated in *priority; only consider models with the highest recall rate

Table 2. Decision Matrix; it shows that the InceptionV3 model, outperforming the rest of the models, is chosen as the final model for the prototype.

• Model	• Recall*	• False Positive Rate	• Precision	• Proc Time	• Accuracy	• MCC	• Results • P * Rank	• Decision
• Priority Ranking (P)	• 1	• 2	• 3	• 4	• 5	• 4	•	•
• InceptionV3	• 1	• 3	• 3	• 1	• 1	• 4	• 41	• ✓
• Xception	• 1	• 4	• 4	• 3	• 2	• 3	• 55	•
• ResNet	• 2	• 2	• 2	• 4	• 3	• 2	• 51	•
• VGG16	• 3	• 1	• 1	• 2	• 4	• 1	• 40	•

*priority; only consider models with the highest recall rate

Table 2. Decision Matrix

PROTOTYPE AND USABILITY TESTING

A prototype for automatic categorization of photos was developed in Python with a Keras-TensorFlow architecture, resulting in a classifier for gun and non-gun categories.

This study focuses measurement on the user's perceived usability, and learnability using the System Usability Scale (SUS) approach[10]. Efficiency is measured by the speed in which tasks were completed both manually and using the prototype. Tasks were designed as a classic test-and-measure approach to elicit a semblance of interaction between the forensic examiner and the prototype in solving an investigation.

RESULTS

The final model is evaluated against two laboratory-generated unbalanced datasets containing only 1% of gun pictures to simulate forensic data. Additionally, the prototype is tested for usability, learnability, and effectiveness with a dataset with manufactured EXIF information to answer typical forensic questions such as GPS Coordinates. There are five respondents: three forensic professionals and two forensic students. The usability and learnability attributes, measured using the System Usability Scale (SUS) approach, resulted in a rating of "Acceptable." Effectivity is measured by comparing the speed of completing tasks between the manual method versus the prototype. The average speed of the manual method is 20.7 minutes, while the tool achieved 19.57 minutes. Hypothesis testing to prove statistical difference between the manual method and the tool's performance was found to be neither worse nor better (no significant difference). However, this

could be attributed to the low number of respondents, resulting in low statistical power. Additionally, alternatives to InceptionV3 (Magnet Axiom software and the Xception model) are also investigated, and the results are promising.

CONTRIBUTION

The growing maturity and the avid support of developers and organisations surrounding the field of computer vision and machine learning make it easier for researchers to transition from theoretical knowledge of machine learning architectures to implementation of applications that are relevant in their specific fields. This study is a demonstration of this possibility in the field of forensics. This research addresses a genuine and important gap in the tools available to forensic examiners in the performance of their forensic activities. This study also demonstrates how the design science methodology can be used in an operationally-focused research.

Keywords: Digital Forensics, Image Classification, Gun Classification, Pre-Trained Neural Networks, ImageNet, Design Science

ACKNOWLEDGMENT

Prof Dr. Hayretdin Bahşi for the continuous hands-on guidance on this topic and the paper

Prof Dr. Leo Mršić for the encouragement in academic writing

Krešimir Hausknecht for the support to pursue this operationally-focused forensics topic

Siniša Urošev for the valuable statistical analysis and computations feedback

Zdravko Kunić, Hrvoje Jerković and *Enes Deumić* for the patience in teaching machine learning concepts and intuition, and the guidance in the practical implementation of these concepts

REFERENCES

[1] Peter Eckersley and Yomna Nasser, "Measuring the Progress of AI Research," EFF AI Progress Measurement Project, 9 9 2017. [Online]. Available: https://www.eff.org/ai/metrics.

[2] J. D. H. S. ,. J. K. ,. S. S. ,. S. M. ,. Z. H. ,. A. K. a. A. K. ,. M. S. B. ,. A. C. B. ,. L. F.-F. Olga Russakovsky, "ImageNet Large Scale Visual Recognition Challenge," *International Journal of Computer Vision*, vol. 115, pp. 211–252, 2015.

[3] E. E. Kenneally, "Gatekeeping out of the Box: Open Source Software as a Mechanism to Assess Reliability for Digital Evidence," *Virginia Journal of Law & Technology*, vol. 6, no. 3, pp. 1–38, 2001.

[4] F. Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," CoRR, vol. abs/1610.02357, 2016.

[5] A. Z. Karen Simonyan, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *3rd International Conference on Learning Representations, {ICLR} 2015, San Diego, CA, USA, May 7–9, 2015, Conference Track Proceedings*, 2015.

[6] X. Z. S. R. J. S. Kaiming He, "Deep Residual Learning for Image Recognition," *CoRR*, vol. abs/1512.03385.

[7] C. S. a. V. V. a. S. I. a. J. S. a. Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," *CoRR*, vol. abs/1512.00567, 2015.

[8] S. T. a. F. H. Roberto Olmos, "Automatic Handgun Detection Alarm in Videos," *Neurocomputing*, vol. 275, pp. 66–72, 2018.

[9] J. Joseph, "The Best Metric to Measure Accuracy of Classification Models," 28 November 2016. [Online]. Available: https://clevertap.com/blog/the-best-metric-to-measure-accuracy-of-classification-models/. [Accessed 9 November 2018].

[10] P. T. K. &. J. T. M. Aaron Bangor, "The System Usability Scale (SUS): an Empirical evaluation," *International Journal of Human-Computer Interaction*, vol. 24, no. 6, pp. 574–594, 2008.

BIOS

KEYNOTE SPEAKERS

Dr Adrian Venables served in the UK Royal Navy for 24 years as a Communications, Warfare, and Intelligence officer. During this period, he was responsible for the management and security of a range of Information Systems worldwide and led specialist teams deployed to operational theatres. Since leaving the Service, he has published a series of journal articles and research papers on the cyber threat landscape and its use by state and non-state actors for espionage, sabotage, and subversion. Adrian joined TalTech, the Tallinn University of Technology in Estonia, as a senior researcher in 2018. He retains his military links by serving as a Commander in the Royal Naval Reserve supporting UK cyber resilience activities in the Baltic region. A Certified Information System Security Professional and Certified Information System Manager, he holds seven computing and cyber security focused degrees. He is a Chartered Information Technology Professional Fellow of the British Computing Society, Chartered Engineer Member of the Institution of Engineering Technology and Fellow of the Chartered Management Institute.

Dr Joanna Kulesza is an assistant professor of international law at the Faculty of Law and Administration, University of Lodz, Poland. She also serves as a member of the Scientific Committee supporting European Union's Fundamental Rights Agency and represents European users at ICANN's At-Large Advisory Committee (ALAC). She has been a visiting lecturer with the Oxford Internet Institute, Norwegian Research Center for Computers and Law, Westfälische Wilhelms Universität Münster and Justus-Liebig-Universität Gießen. Kulesza has served as an expert for the Council of Europe on human rights online (Ukraine 2015, Moldova 2016) and supports with her expertise the Sino-European Cybersecurity Dialogue. She is also a reviewer for the EU COST program and a faculty member of the Program on Cybersecurity and Internet Governance at Indiana University Bloomington. She chairs the Membership Committee of the Global Internet Governance Academic Network (GigaNet). Joanna is the author of numerous publications on international Internet law, including "Cybersecurity and Human Rights in the Age of Cyberveillance" (together with R. Balleste, Rowman and Littlefield 2015) and "Due Diligence in International Law" (BRILL 2016). Her research focus is on the intersection of human rights and cybersecurity.

MODERATORS

Prof Olaf Maennel is a Professor for Cyber Security at Tallinn University of Technology. He graduated from the Technical University Munich in 2005. Since then he has been with the University of Adelaide in Australia and Loughborough University in UK. In July 2014 he has joined TalTech and the Centre of Digital Forensics and Cyber Security. His research interests include network security, network forensics, serious games (focusing on red-teaming and learning aspects), capability profiling & assessment, human factors; and aviation & maritime cyber security aspects. He is also responsible for international admissions and was co-chairing ACM SIGCOMM 2015, ACM IMC 2017, the Global Internet Symposium GI'2017, and also co-chaired jointly with Anna-Maria Osula four Interdisciplinary Cyber Research workshops (ICR). He is also chairing working group 5.2 at the European Cyber Security Organisation (ECSO) on cyber education and professional trainings, and he serves as EU expert evaluator and vice-chair for Horizon H2020 calls.

Prof Hayretdin Bahşi received his PhD from Sabancı University (Turkey) in 2010. He was involved in many R&D and consultancy projects about cyber security as a researcher, consultant, trainer, project manager and program coordinator at Information Security Research Centre of Scientific and Technological Research Council of Turkey between 2000 and 2014. Currently, he is a research professor at Centre for Digital Forensics and Cyber

Security of Tallinn University of Technology. His research interest includes critical information infrastructure security, digital forensics, machine learning and its application to cyber security.

Prof Tobias Eggendorfer is a professor of IT-security at Hochschule Ravensburg-Weingarten, prior he was professor for IT-forensics in Hamburg. He was awarded his PhD by FernUniversität in Hagen in 2007 with a dissertation on e-mail security and spam prevention. He graduated in Munich in Engineering and Business Administration, in Mittweida in Technical Informatics, in Hagen in Computer Science as well as in Law and in Kaiserslautern in Adult Education. His current research interests are in computer security, especially embedded security, and computer forensics.

Dr Anna-Maria Osula is a senior policy officer at Guardtime, a systems engineering company offering data-centric security solutions based on blockchain technologies as well as cyber range exercises. She also serves as senior researcher and lecturer at Tallinn University of Technology (TalTech), and as a research fellow at Masaryk University under the project "Cyber Security, Cyber Crime and Critical Information Infrastructures Center of Excellence". Previously, she worked as a legal researcher at the NATO CCDCOE, focusing on national cyber security strategies, international organisations, international criminal cooperation and norms. She has authored a number of publications, including co-editing the NATO CCD COE book "International Cyber Norms: Legal, Policy and Industry Perspectives". She is also the founder of an annual conference series "Interdisciplinary Cyber Research", which has been organized by TalTech since 2014. In addition to a PhD in law from the University of Tartu, she holds an LLM degree in IT law from Stockholm University.

Mr Aykan Inan is currently a research assistant and PhD student at the University of Applied Sciences in Ravensburg-Weingarten. Before that he finished a combined study program supported by the German Department of Defense at the University of the Federal Armed Forces in Munich with his Bachelor of Engineering in Computer Engineering. Subsequently he worked for the ministry in Bonn for another period of three years. He simultaneously studied IT Security at the Ruhr-University Bochum before he became a research assistant in Ravensburg-Weingarten. In addition, he is also teaching Basics of Computer Science and Operating Systems in a Bachelor's study course. His research activities comprises IT security and forensics in general. But his primary research activity lies on Cryptography and Information Security with focus on attempting to improve attacks on specific algorithms. One of his projects is dealing with the analysis of primes and their hidden characteristics.

Prof Matthew Sorell is Senior Lecturer in Telecommunications and Multimedia Engineering at the University of Adelaide and an Estonian e-Resident. He has coordinated the annual Adelaide-Tallinn-Ravensburg study tour program since 2015, in collaboration with Olaf Maennel and Tobias Eggendorfer. He has received a number of university and national awards for excellence in teaching, including an Australian Learning and Teaching Council citation in 2009. Dr Sorell's primary research and consulting activities are in regulation, security and digital device and multimedia forensic analysis. In 2017 he was appointed to the INTERPOL Digital Forensics Experts Group and earlier this year was awarded a high commendation for contributions to the development of digital forensic evidence by the National Police Chiefs Council in the United Kingdom.

SPEAKERS

Mr Muhammad Mudassar Yamin is a PhD-Candidate at department of information and communication technology in Norwegian University of Science and Technology. He is the member of system security research group and the focus of his research is system security, penetration testing, security assessment, intrusion detection. He is currently working on the modeling of attack and defenses scenarios for cyber security exercises in a cyber range.

Mr Gábor Visky was commissioned in 1998 from János Bolyai Military College as a Radio Reconnaissance Officer. He has Bachelor's degree in the field of telecommunication and Master's degree in Information Engineering in the specialty of industrial measure-

ment From University of Miskolc. During the last 20 years' active duty Gábor served in several different positions in the Hungarian Defence Forces including several years of service abroad as well. In 2018 he has been selected to a researcher position of the Technology Branch of NATO Cooperative Cyber Defence Centre of Excellence, where his main field of expertise is industrial control systems, cyber-physical battlefields used during cyber exercises. Gábor's prior assignments include 15 years designing hardware and software for embedded control systems, and researching their vulnerabilities by reverse engineering. His personal awards including the Officers' Service Sign 3rd and 2nd Class and Merit of Service Medal Bronze and Silver Grade.

Mr Kieren Niĉolas Lovell is a Cybersecurity and Communications specialist and the Head of the Computer Emergency Response Team at TalTech University. In this position, Kieren prevents, protects and investigates attacks on University information systems. In addition to this, he also conducts international OSINT and cybersecurity exercises with 20 universities across Europe, with military institutions, and with commercial companies. Prior to this role, Kieren was Head of the Computer Emergency Response Team for the University of Cambridge. He came to Cambridge after working in the Royal Norwegian Navy, as a Battlewatch Captain and Chief Information Security Officer for Standing NATO Maritime Group One. Whilst he in an operational and research role at TalTech, Kieren is still a lecturer at King's and Pembroke College, University of Cambridge and for TalTech University in Incident Management and OSINT education. He is also an advisor for the Cambridge Science and Policy Group to Her Majesty's Cabinet Office and is also a director of his own security company in Estonia.

Ms Kaie Maennel is PhD student at Tallinn University of Technology (TalTech). She graduated MSc Cybersecurity at TalTech and University of Tartu in 2015. Her research focuses on application of learning analytics in cybersecurity trainings (specifically in cybersecurity exercises), as a way to provide more evidence-based and systematic approach for evaluation of learning impact and enable designing more effective learning. She has participated in the NATO CCDCOE Locked Shields cyber defense exercise as white team member for last 3 years. She is also Audit Learning Leader at Deloitte for Central Europe. In this role she is informed and implements latest trends in corporate sector learning.

Prof Matthew Sorell is Adjunct Professor of Digital Forensics at TalTech and Senior Lecturer at the University of Adelaide. He is a member of the INTERPOL Digital Forensics Experts Group, the Scientific Advisory Board of FORMOBILE, and a consultant to several law enforcement agencies.

Mr Andrew Roberts is a cyber security researcher currently studying at the Tallinn University of Technology and the University of Tartu. He has worked as an IT infrastructure project manager and systems engineer for NTT. His area of research is focused on critical infrastructure protection, national security policy and cyber operations. He holds a Master of Cyber Security Operations from the University of New South Wales, Canberra.

Mr Samuel Henderson is an undergraduate student at The University of Adelaide. He is currently in his final year of a Bachelor of Engineering (Honours) (Electrical and Electronic) with a Bachelor of Mathematical and Computer Sciences (Computer Science Major). This year, Samuel is completing a Defence Science Technology Group sponsored project, investigating the limitations of current Twitter bot detection machine learning algorithms, under the supervision of Matthew Sorell. Samuel is currently working for Lendlease Engineering, a leading infrastructure and services company in Australia.

Mr Brian Du is an undergraduate student at The University of Adelaide. He is currently in his final year of a Bachelor of Engineering (Honours) (Electrical and Electronic) with a Bachelor of Finance. This year, Brian is completing a Defence Science Technology Group sponsored project, investigating the limitations of current Twitter bot detection machine learning algorithms, under the supervision of Matthew Sorell. Brian is currently working for Consilium Technology, an upcoming software development company in Adelaide.

Mr Akim Essen is last year Cyber Security student at Tallinn University of Technology and a Software Developer with 5 years of experience with a Bachelor's degree in Informatics from Tallinn University of Technology. He has a keen interest in various methods of Authentication, Identification and Verification, with other interests laying in the field of Cryp-

tography and Quantum Computing. With a growing need of stronger authentication methods to combat various attacks, biometric scanners can play a vital role in keeping it not only effective, but also user-friendly. His project focuses on evaluating password replacement with several biometric scanners that can work in unison to authenticate a person with ease.

Ms Grethe Østby is a PhD researcher at the Norwegian University of Science and Technology. Her research topic is information security awareness in the society and readiness in public emergency organizations. Grethe graduated as a Siviløkonom/Master of Science in Business in 1998, and she has been working 12 years in sales and in leading implementation and managing teams in operate sales systems, customer-service systems and management systems in private companies, and 7 1/2 years as a crisis manager at The Norwegian Civil Defense national competence center. Grethe started her career at the, at that time, the Norwegian Army Signal Core officers' school, today the Norwegian Army Cyber Defense Academy, and find it lucky that this and her other education and vast experience now can be of value in her PhD-research. Her research interest is in combining crisis management and cyber security and thereby socio-technical research in the society and in public emergency organizations. She likes to describe herself as a long-term (strategic) thinker, is systematic and follow up on responsibilities. She prefers to work in environments that value changes and development, both for the organization, but also on personal levels.

Mr Liam Shelby-James is a final-year honours student at the University of Adelaide, Australia, studying Electrical and Electronic Engineering. His honours research is on the reliability and trust in global navigation satellite systems, which develops a framework to display the level of interference in a set of navigational signals. His research interests include communications, signal processing & RF engineering, with a strong focus on system security & integrity.

Mr Stefan Norman is completing his final year of his Bachelor of Engineering (Honours) in Electrical and Electronic Engineering (Telecommunications) at the University of Adelaide and is taking part in the 2019 Cyber Security Summer School in Estonia alongside a group of students from Adelaide, Australia. Stefan's research interests include signal processing, and the security and integrity of communications. His honours research topic is the trust and reliability of global navigation satellite systems, which aims to develop a framework to determine the level of trust that can be placed in received satellite signals, and produce a user-friendly display metric.

Dr Jakub Harašta (1988) is associate professor at Masaryk University, Faculty of Law. He was Postdoctoral researcher at Center for Cyber Law & Policy, University of Haifa (2018) and Visiting Research Fellow at Minerva Center for the Rule of Law under Extreme Conditions, University of Haifa (2015). Jakub is editor-in-chief of Masaryk University Journal of Law and Technology. In his research, Jakub focuses on various issues of cyber security, cybercrime and personal data protection. He also researches in the field of legal informatics and legal information retrieval, and serves as guest lecturer at the Judicial Academy of the Czech Republic where he co-developed course, which provides training to judges in the field of legal information retrieval. In the past, Jakub was called to take part as legal advisor in both national and international cyber security exercises (Locked Shields, Cyber Coalition, Cyber Czech).

Ms Ivana Kudláčková is a research fellow at the Institute of Law and Technology, Faculty of Law (Masaryk University). She is a member of the research team of the project *CyberSecurity, CyberCrime and Critical Information Infrastructure Centre of Excellence*. She predominantly focuses on use of force in cyberspace, new methods of warfare and relevance of international public law in cyberspace. She graduated from Law at Masaryk University and furthermore attended various courses at the Georgian Institute of Public Affairs and at Ghent University.

Ms Marija Makariūnaitė has graduated from Vilnius University with Master of Laws degree (5 year integrated studies) and from the University of Tartu with Master in IT Law (1 year programme). She is interested in privacy and data protection (especially in California Consumer Privacy Act and EU's General Data Protection Regulation), e-governance, intellectual property protection in the field of IT, and legal issues surrounding cyber security and artificial intelligence.

Mr Jaan Priisalu is researcher in Tallinn University of Technology and Senior Fellow of NATO Cooperative Cyber Defense Center of Excellence. Main research interest in critical infrastructure protection and exercises. He is the former General Director of the Estonian Information System's Authority. He is the co-founder of the Estonian Defense League's Cyber Unit - an organization that defends the Estonian cyber space and unites voluntary IT-specialists - and member of the Estonian Information Systems Audit Association, Priisalu has two decades of practical experience on the matter of preventing and defending cyber-attacks. Before entering the public sector, Priisalu worked at Guardtime, Ühispank, Cybernetica, Hansabank and served as the head of SIRT (Security Incidents Response Team) at Swedbank; having worked through the 2007 cyber-attacks. Jaan Priisalu earned his academic education from the Tallinn University of Technology and Toulouse III (Université Paul Sabatier) in France.

Mr Matthew Theiley is a South Australian student currently studying a double degree of Electrical Engineering and Computer Science at The University of Adelaide. Matthew also is currently working in a paid internship with ASTC (Australian Semiconductor Technology Company) as a software engineering intern. As a child Matthew was fascinated with science specifically in the areas of physics and chemistry being a part of the Young Scientists of Australia and volunteering in roles such as the Science Experience and Science Alive. However, as Matthew grew older he discovered a passion for programming and computers which lead him work teaching small basic and python to students for a company in Brisbane called Junior Engineers. He went on to teach robotics for Sciworld and eventually specialized into the area of computer architecture and design which is how Matthew found his dream job working for ASTC.

Mr Charlie Tran is an undergraduate student at The University of Adelaide, currently finishing his final year in a Bachelor of Engineering (Electrical and Electronic) double degree with a Bachelor of Finance. In 2018, he completed a twelve month internship at Defence Science and Technology (DST) Group, where he researched radar detection and simulation. For his final year project, Charlie's team is researching attacks on the CAN messaging protocol in the automotive industry. In particular, they are developing subtle and deceptive attacks that can take advantage of the security vulnerabilities in the CAN protocol. Charlie is also participating in the 2019 Digital Security study tour in Estonia to learn more about the cybersecurity industry and extend his research.

Mr Stefan Smiljanic is currently finishing his bachelor double degree in Electrical and Electronic Engineering (honours) with Finance at the University of Adelaide. Stefan and his team are utilising a vehicle testbed environment to develop inconspicuous attacks on the widely adopted CAN protocol, as a part of their final year project. His academic highlights include participating in the 2019 Digital Security study tour in Estonia to extend his research, and a Semester Exchange Program in 2018 at the University of Maastricht, the Netherlands. Stefan's professional development includes being involved in MOIREI Electronics, an Adelaide based start-up company, since 2017. Their focus is to create technologies for developing countries in Africa. Stefan is starting in an engineering graduate program in 2020 and looks forward to making a local and global impact throughout his professional career.

Mr Ahmad Amine Loutfi is a PhD Candidate at the TEFT-Lab at the Norwegian University of Science and Technology. He is one of the three PhD candidates that are working on the Dig-Eco project (2018–2022) which targets three aspects of Digitalization: Energy, Health and Finance. His main focus is Financial Technologies (Fintech). The most recent academic experience that is highly relevant to his PhD, is the thesis project he worked on as part of the master degree of international business at NTNU, where he have worked on the applications and challenges of Blockchain technology in the manufacturing value chain at Ekornes, one of Norway's and Europe's largest furniture manufacturing companies. Furthermore, he has worked as a part time "technology and market research analysts" with a tech startup in Oslo. His main duty is to investigate and analyze new emerging technologies and provide applicable business use cases that specific industries/companies can implement. His most recent assignment is related to Blockchain application in the fishing industry. Last by not the least, he would like to touch upon his experiences in supply chain management and logistics, namely by completing a specialized master degree

in this field, and then completing professional placements as a junior consultant at Cummins, Belgium.

Mr Ben Agnew has a bachelors degree in Electrical and Electronic Engineering from The University of Adelaide, Australia and is currently a PhD student at the university. His research involves looking at the analog characteristics of memory cells and the applications this may have in digital forensics. By exploiting manufacturing defects that are usually ignored in digital circuits, forensic information can potentially be extracted from a device. He has presented aspects of this work at ICR in the past, in 2015 and 2017. Today he will be giving an overview of information security technology and how it can be used in digital forensic applications.

Mr Glenn Walsh is in his final year of studying electrical and electronic engineering. He's working with Jimmy Tang on their project, where they are researching the forensic applications of 3D scanning, which is being supervised by Matthew Sorell, and Richard Matthews. Glenn went to Trinity College in Evanston, where he represented the school, and Australia in the F.I.R.S.T. robotics competition held in St Loius, America. He's new to the cyber security school of research, but is hoping to further his knowledge by attending ICR this year.

Mr Jimmy Tang is a Vietnamese Australian born in Victoria but moved to South Australia in his early childhood. He went to a school in Paralowie a suburb north of Adelaide where he had went to Paralowie R-12 school from reception to grade 12 before embarking on a journey into the University of Adelaide North Terrace campus in the year 2015 to study a double degree in Computer Systems Engineering and Computer Science. In 2017 he went on a student exchange to the University of Glasgow in the United Kingdom which lasted for 1 semester and gained international experience while studying high level courses such as Machine Learning and Information Retrieval. Study was not the only thing on Jimmy's mind though, throughout the years he has studied a Vietnamese Martial art, Kendo, and Salsa dancing. Professionally, Jimmy has done worked in embedded systems, HDL such as Verilog, and high performance computing. It has been described before that Jimmy is a tinkerer and loves to learn, play with, test, and experiment with electronics and computer systems and would personally describe himself as adventurous.

Mr Luke Jennings has an Honours degree of Bachelor of Engineering, awarded with First Class Honours in Electrical and Electronic Engineering at the University of Adelaide, Australia. He interned with BHP Billiton as an Electrical Engineer at the Olympic Dam Asset located in South Australia, as a member of the Governance and Technical Stewardship team from November 2018 to February 2019. He is currently undertaking a Master's degree by research at the University of Adelaide in the school of Electrical and Electronic Engineering. His current research focus is on using iPhone and fitness device data to identify behavioural patterns in users for the purpose of aiding digital investigations.

Ms Joanna Rose Castillon del Mar is a Cyber Security student specializing in Digital Forensics at the Tallinn University of Technology and Tartu University. She had her summer internship in INsig2 d.o.o. in Zagreb, Croatia where she pursued the topic of forensics for her master's thesis. She spent the second year of her studies in Algebra University in Zagreb under the Erasmus program where she studied machine learning concepts and its applications, visualization, quantitative analysis and penetration testing. Her interests include software development, data analysis and scuba diving. She comes from the Philippines and had previously worked in Singapore before coming to Estonia.