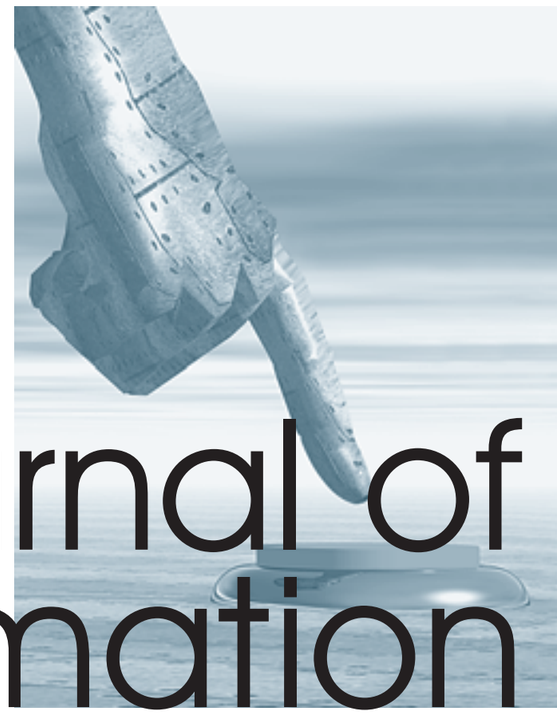
Volume 17, Issue 4, Fall 2018

ISSN 1445-3312 (Printed Journal)

ISSN 1445-3347 (Online Journal)

JOURNAL OF INFORMATION WARFARE

Volume 17 Issue 4 Fall 2018



Journal of Information Warfare

Journal of Information Warfare (JIW)

www.jinfowar.com

Journal Staff

Chief Editor
Dr. Leigh Armistead

Assistant Editor
Dr. William Hutchinson

Deputy Editor in Chief
Dr. Diane Silver

Technical Editor
Dr. Marla Weitzman

Editorial and Technical Advisor
Zachary Hubbard

Administrative and Editorial Assistant
Angel Linzy

Editorial Board

S. Furnell	J. Lopez
J. Slay	P. Williams
H. Armstrong	C. Irvine
C. Bolan	A. Jones
G. Duczynski	W. Mahoney
A. Ahmad	C. Valli
M. Henson	A. Liaropoulos

Advisory Board

Dr. Corey Schou
Idaho State University, Idaho, United States

Professor Matthew Warren
Deakin University, Melbourne, Australia

Dr. Brett van Niekerk
University of KwaZulu-Natal, Durban, SA

Scope

The journal has been created to provide a forum for discussion, information, and interaction between practitioners and academics in the broad discipline of information warfare/operations. It is of interest to professionals from the military, government, commerce, industry, education, and academy.

A full gambit of topics is covered—from the physical destruction of information systems to the psychological aspects of information use. The aim is to provide a definitive publication that makes available the latest thinking and research in the critical area of information warfare.

Submissions

The journal welcomes submissions. To learn more about preparing articles for submission, authors should visit the *JIW* website. Articles may be submitted to Diane Silver at dsilver@gbpts.com or Angel Linzy at alinzy@gbpts.com.

Authors' Responsibilities & Copyright

Authors are to ensure the accuracy of their papers. This journal does not accept any responsibility for statements made by authors in their written papers. Where relevant, authors are to ensure that the contents of their papers are cleared for publication, for example, by their employer, their client, the funding organization, and/or copyright owner of any material that is reproduced.

Copyright of the article is retained by the authors who warrant that they are the copyright owner and have in no way infringed any third- party copyright. In submitting the article for publication, the above warrant is implied as is the grant of a non-exclusive copyright license by the author to the *Journal of Information Warfare* to publish the work as determined by the Editorial Board.

The views expressed by contributors do not necessarily represent those of the editors, advisory board, or the publishers.

Subscriptions

The *Journal of Information Warfare* is published four times per year and is available both online and in hard copy.

Individual; **Individual, Student**; and **Corporate** subscriptions are available. For current pricing, see http://www.jinforwar.com/subscribe/.

Individual
A one-year subscription to the journal for individual subscribers. Both online-only and, online and print subscriptions are available**.**

Individual, Student
A one-year subscription to the journal for students. Evidence of full-time study must be provided. Both online-only and, online and print subscriptions are available**.**

Corporate
A one-year subscription to the journal for corporate/library subscribers. Both online-only and, online and print subscriptions are available**.** A single subscription covers unlimited use for a single campus/geographic location.

Note: Individual print copies of the journal are generally available for purchase only to subscribers and contributors.

All advertisements in this journal are printed free of charge as a service to readers.

Journal cover design, concept, and layout by Laima Croft.

Journal of Information Warfare

Volume 17, Issue 4
Fall 2018

Contents

Journal of Information Warfare
© Copyright 2019

Published by
Peregrine Technical Solutions, LLC
Yorktown, Virginia, USA

Print Version
ISSN 1445-3312

Online Version
ISSN 1445-3347

From the Editor

Fall 2019

Information has always been regarded as an element of power. Too often, however, it has been seen as an enabling or supporting component rather than as the mission-critical element it frequently is in conducting operations. Indeed, the very nature of modern-day operations, with its persuasive and never-ending 24-7 global media coverage, has time and again demonstrated the need for actors and nation-states to utilise all the tools or elements of power at their disposal. To be sure, information must be considered a vital component in any sort of influence type of operations.

However, the factors that make information useful as an element of power are also adding to the difficulties nation-states face in their efforts to conduct information campaigns, or IO, on a successful basis. The shifting of power away from a centralised authority, the loss of control from the federal bureaucracy, and the low cost as well as ease of entry into this domain have combined to create a fundamental change in the ways that information is and can be utilised around the world.

The advent of new information-technology capabilities has certainly benefited elements of society: non-governmental organisations, non-state actors, corporations, individuals, and—unfortunately—terrorists. But it has also radically altered the methods by which the administration and other branches of the federal government can interact with their counterparts around the world. At the centre of the issue is a lack of control—control of content, flow, and paths of communication.

Consequently, despite the heightened expectations understandably fostered by the increased capabilities inherent in IO, governmental entities struggle and often fail to construct successful information campaigns and operations. There is no coherent theoretical construct, definition, or taxonomy—instead, there is a virtual smorgasbord of training classes with varying curricula and content, none of which are integrated or coordinated. In general, too much is expected, and too much has been promised. With no radical changes in funding across these agencies, it is no surprise that progress has overall been disappointing.

While the Journal of Information Warfare cannot solve all of these issues, it can and will continue to provide a forum for voices and theories and research and especially conversations that can and will contribute to the real and significant progress we have every reason to expect.

Join the conversation with us at one or more of the professional conferences slated for the coming year. The next event is the 18th European Conference on Cyber Warfare & Security, which will be held at the University of Coimbra in Portugal in July 2019. The Australians usually host their Edith Cowan University Security Research Institute in Perth, Australia, in early December of each year, with an event that hosts five different security-based conferences over these three days. In March 2020, the 15th ICCWS will convene at Old Dominion University in Norfolk, VA.

Finally, we are always looking for good reviewers for JIW. If you are interested, please contact me (larmistead@gbpts.com) or email our staff at jiw@gbpts.com. Members of our review board help us conduct our double-blind, peer- reviewed assessment; appear on the journal's masthead; and get a free online subscription during each year of their service. Cheers,

Dr Leigh Armistead, CISSP
Chief Editor, *Journal of Information Warfare*
larmistead@gbpts.com

Authors

Burhan Al-Bayati is currently a final-year PhD candidate at the Centre for Security, Communications & Network Research at the University of Plymouth (UK). He holds a BSC in computing from Baghdad University (Iraq), 2002, and an MSC in computing from Pune University (India), 2008-2010. Burhan's research interests include information security, biometric authentication, and cloud security.

Jonathan Z. Bakdash received the PhD degree in psychology in 2010 from the University of Virginia. He is a Research Psychologist with the Human Research and Engineering Directorate, U.S. Army Research Laboratory, South Field Element, at the University of Texas, Dallas. His research interests include human decision-making, human-machine interaction, and cyber security.

Dr. Andrew Blyth, formally the Director of the Information Security Research Group at the University of South Wales, has functioned as an expert witness in computer forensic and data recovery for a wide variety of law enforcement agencies, such as the Home Office, SOCA, and the Metropolitan Police. Dr. Blyth has also published several journal papers in the areas of computer forensic and data recovery, and is one of the leading global authorities on data sanitization and forensic techniques on solid state media. Dr. Blyth is on the ISO advisory board for standards relating to Computer Forensics, is a member of the National IA forum, and works with UK government agencies, including the Defence Science Technology Laboratory.

Dr. Matthew Bovee is the Associate Director of the Computer Science/Computer Security & Information Assurance program at the Norwich University School of Business & Management. As Lecturer there, he teaches general and specialist courses in computer science, digital forensics, and computer security. In addition to cyber security and digital forensics, Dr. Bovee's background includes research, publications, and degrees in accounting and information systems and exercise physiology.

Filipe Caldeira is an Adjunct Professor at the Informatics Department of the Polytechnic Institute of Viseu, Portugal. He obtained his PhD degree in Informatics Engineering in 2014 from the Faculty of Sciences and Technology of the University of Coimbra. He has acted as program director of the Informatics Engineering program since 2014. He is also a researcher at the Centre for Informatics and Systems of the University of Coimbra and at the CI&DETS research center of the Polytechnic Institute of Viseu. He has been recently involved in some international and national research projects.

Professor Nathan Clarke is a Professor in Cyber Security and Digital Forensics at the University of Plymouth. He is also an adjunct Professor at Edith Cowan University in Australia. His research interests reside in the areas of information security, biometrics, forensics, and cloud security. Professor Clarke has over 200 outputs consisting of journal papers, conference papers, books, edited books, book chapters, and patents. He is the Chair of the IFIPTC11.12 Working Group on the Human Aspects of Information Security & Assurance. Professor Clarke is a chartered engineer, a fellow of the British Computing Society (BCS), and a senior member of the IEEE.

Tiago Cruz has been an Assistant Professor at the Department of Informatics Engineering at the University of Coimbra since 2013, where he obtained his PhD in Informatics Engineering in 2012. His research interests cover areas such as management systems for communications infrastructures and services, embedded computing, critical infrastructure security, IoT, and SDN/ NFV. The author of more than 70 publications, including book chapters, journal articles, and conference papers, he has also been involved in various international and national research projects. He is a member of the IEEE Communications Society and an IEEE senior member.

Dr. Petrus Duvenage is a counter-intelligence specialist with extensive experience in various aspects of this field. In the course of his thirty-year career, he served as an officer in the South African armed forces and in the intelligence services. He holds, in his personal capacity, a Senior Research Fellowship at the Academy for Computer Science and Software Engineering, University of Johannesburg. He has a PhD from the University of Pretoria and is currently a PhD candidate at the University of Johannesburg.

Chuck Easttom holds a Doctor of Science (D.Sc.) in Cyber Security as well as three master's degrees (one in applied computer science and one in systems engineering). He is the author of 26 computer science books used as textbooks at over 60 universities. He has also authored dozens of scientific papers on a range of topics and is an inventor with 5 computer science related patents. He is a member of the IEEE and ACM, and a Distinguished Speaker of the ACM. He is also a reviewer for several scientific journals. He is involved in research in a variety of areas including cryptography, cyber warfare, engineering processes, and digital forensics.

Simone Fischer-Hübner has been a Full Professor at Karlstad University since June 2000, where she is the head of the Privacy & Security (PriSec) research group. She received a Diploma Degree in Computer Science with a minor in Law (1988), and a PhD (1992) and Habilitation (1999) degrees in Computer Science from Hamburg University. She has been conducting research in privacy and privacy-enhancing technologies for more than 30 years.

Dr. Lothar Fritsch is currently working as a lecturer and researcher in Information Security at Karlstad University in Sweden. He works and teaches in topics of computer science, information and security, and information privacy. He was the coach of the winning team in the Swedish 9/12 Cyber Challenge 2017. Lothar's work focuses on the analysis of security and privacy requirements in upcoming application areas; particularly, he worked on the deployment of privacy functionality into new systems with respect to requirements engineering and verification, and the assessment of privacy risks and impact.

Associate Professor Paul Haskell-Dowland is the Associate Dean for Computing and Security in the School of Science at Edith Cowan University, Perth, Australia, and is an associate Member of the Centre for Security, Communications & Network Research at the University of Plymouth in the United Kingdom. Paul has delivered keynotes, invited presentations, workshops, and seminars across the world for audiences including Sri Lanka CERT, ITU, and IEEE. He has more than 20 years of experience in cyber security research and education in both the UK and Australia.

João Henriques is a PhD student in Science and Information Technology at the University of Coimbra (UC) and Assistant Professor at the Department of Informatics Engineering at the Polytechnic Institute of Viseu (IPV). His research interests at the Center for Informatics and Systems at UC (CISUC) include forensic and audit compliance for critical infrastructures protection. He also remains a Software Engineer in the private sector.

Dr. Victor Jaquire has been within the field of cyber and information security for over 20 years within government and the private sector focusing on strategy, performance management, and operations. He holds an Honors Degree in Management from Henley University and a master's and PhD in Informatics from the University of Johannesburg--specializing in strategies for cyber counter-intelligence maturity and the security of cyberspace. He has published various academic papers on cyber strategies and cyber counter-intelligence maturity. His professional certifications include CISSP, CISM, and CCISO.

Fudong Li is a lecturer in Cyber Security at the University of Portsmouth, in the UK. Dr. Li is also a visiting lecturer at the University of Plymouth. His research interests are in the areas of biometric authentication and digital forensics; he has over 50 conference papers and journal articles in those domains.

Dr. John McAlaney is a Chartered Psychologist, Chartered Scientist, and Principal in Psychology at Bournemouth University in the UK. His research focuses on the social psychological factors of risk behaviors, including cyber security from the perspective of the attackers, the targets, cyber security practitioners, and other stakeholders.

Glenn Nor has a background in IT network and security, and completed one of Norway's first bachelor degrees that focuses specifically on digital forensics. He is now head of forensic technology services at PwC Norway and pursuing an MPhil/ PhD at the University of South Wales.

Dr. Huw Read is an associate professor at Norwich University and the director for the Centre of Advanced Computing and Digital Forensics (NUCAC-DF). Dr. Read began his academic career in 2004 at the University of South Wales (UK) and has taught several specialist courses in digital forensics and cyber security. For over ten years, he has worked alongside industry as well as government on a number of cyber-related projects, partnering with diverse teams to design solutions to complex security problems. Dr. Read is actively engaged in research and scholarship within the field, having published a number of research articles in journals and spoken at various cyber-related conferences.

Dr. Char Sample is a research fellow employed for ICF International at the US Army Research Laboratory in Adelphi, Maryland, and is also with the University of Warwick, Coventry, UK. Dr. Sample has over 20 years' experience in the information security industry. Most recently, Dr. Sample has been advancing the research into the role of national culture in cybersecurity events. Presently, Dr. Sample is continuing research on modeling cyber-behaviors by culture; other areas of research are information weaponization, data fidelity, and deception.

Paulo Simões is a Tenured Assistant Professor at the Department of Informatics Engineering of the University of Coimbra, Portugal, where he obtained his doctoral degree in 2002. He has been involved in in several European research projects and led several industry-funded technology transfer projects for telecommunications operators and energy utilities. His research interests include Network and Infrastructure Management Security, as well as Critical Infrastructure Protection.

Professor Iain Sutherland is currently Professor of Digital Forensics at Noroff University College in Kristiansand, Norway. He is a recognized expert in computer forensics and data recovery. He has authored numerous articles ranging from forensics practice and procedure to network security. In addition to being actively involved in research, he has acted as a consultant on forensic and security issues for both UK police forces and commercial organizations. His current research interests lie in the areas of computer forensics and computer security.

Helen Thackray is a PhD candidate in Computing and Psychology at Bournemouth University in the UK. Her research examines the social psychological factors of group behavior and social identity within online communities, including the impact on cyber security.

Professor Sebastian von Solms is a Research Professor in the Academy for Computer Science and Software Engineering at the University of Johannesburg in South Africa. He is also the Director of the Centre for Cyber Security at the University of Johannesburg. He specializes in research and consultancy in Information and Cyber Security, Critical Information Infrastructure Protection, Cyber Crime, and other related cyber aspects. He has written and presented more than 130 papers regarding this field—most of which have been presented at international research conferences and/or published in international subject journals.

Automating Aspects of Forensic Case Management

G Nor1,[4], I Sutherland[2,3], AJC Blyth[4]

[1]*Forensic Technology Services*
PriceWaterhouseCoopers
Oslo, Norway

E-mail: glenn.nor@pwc.com

[2]*School of Technology and Digital Media*
Noroff University College
Kristiansand, Norway

E-mail: Iain.sutherland@noroff.no

[3]*Security Research Institute*
Edith Cowan University
Perth, Australia

E-mail: Iain.sutherland@noroff.no

[4]*School of Computing and Mathematics*
University of South Wales
Pontypridd, United Kingdom.

E-mail: andrew.blyth@southwales.ac.uk

Abstract: *The forensics community has invested considerable effort in the development of tools in support of the different stages of a digital investigation. However, the main focus has been on the development of tools to capture data, or to support later forensic analysis in the sifting and sorting of large volumes of data in the search for information relating to specific system or user activities. There has been more limited effort and success in terms of the development of tools to support case management and less still on the organisation of metadata needed for the reporting and formatting of evidence for court.*

This paper reviews some of the current tools for reporting the results of forensic analysis. It outlines a lightweight approach based on the automated creation of folder structures and a related referencing methodology aimed at reducing the possibility of human error. This system, adopted commercially for organising evidence potentially extracted from several different tools, enables multiple investigators to collate and to consistently organise information for reporting and review.

Keywords: *Forensic Analysis, Forensic Case Management, Automation, Digital Forensics, Forensic Report Referencing*

Introduction

Forensic investigations are dealing with expanding volumes of data, often collected from multiple devices and systems. Each distinct piece of evidence requires appropriate metadata recording aspects of the collection and the relevance to the case. The current suites of forensic tools capture a degree of related metadata although, historically, the target of these tools has been on the capture and analysis of the data. The focus has been on the creation of tools to process forensics cases—to capture data for later forensic analysis, or to support forensic analysis in the searching and sorting of large volumes of data for information relating to specific system or specific user activities.

There has been limited success in the development of tools to automate elements of case management and, in particular, in the reporting and formatting of evidence for court. A challenge in creating reporting tools is the wide range of possible requirements for forensic reports dictated by the needs of each case. The wide variation in reporting requirements often results in the adoption of a manual process to consistently organise evidence for reporting and review.

This paper reviews some of the current tools used to support the reporting stage of forensic analysis. It then outlines a lightweight approach to managing artefacts based on the automated creation of folder structures and related referencing methodology aimed at reducing the possibility of human error in collating material for reporting and presentation.

Existing Tools and Methods

Automated processing of data into a storage format or data-store for analysis is an important feature of most forensic tool suites, including both major commercial forensic tools suites, FTK (Accessdata 2018) and Encase (Guidance Software 2018). In the case of many tools, the creation of case metadata begins with the creation of the forensic image, as certain file formats include basic information, such as cyclic redundancy checks and file hashing. Tools may also include a summary file describing the captured image. An example is the text file created when using FTK (Accessdata 2018). This file provides a range of useful case metadata for each forensic image, including bad sectors, hashes, time, and date, in addition to other information, depending on the options selected by the user. For cases in which tool capabilities are limited to generating bitstream copies, metadata used to validate images must be captured by using some other method. Ensuring sufficient case metadata is recorded to describe the evidence is a key part of an investigation. There are other concerns, such as preventing the tampering with evidence metadata and the need to ensure that metadata relating to chain of evidence has not been altered (Gayed, Lounis & Bari 2012). Hardware solutions, such as the Spyrus blockchain-based Hardware Security Module, may offer some solutions in this area (Spyrus 2018). Attempts have been made to generate a digital equivalent to an evidence bag (Turner 2005) as well as to store data (and case metadata) in a common format (DFRWS 2006; Expert witness compression format 2018) and AFF (Garfinkel et al. 2006). However, the exact information recorded depends not just on the specific type of case, but also on each individual case, as each will have specific features. The type of case metadata recorded may also be influenced by the reporting requirements, which may vary depending on charges brought; also, evidential requirements are dependent on the jurisdiction and the variations in the legal system. This may be one of the most significant challenges in developing tools to support reporting for court purposes.

Existing tools store the created metadata in various formats and types of files and would require the investigators to manually go through each metadata file in order to extract information. If a forensic team needs to know who acquired specific data and at what time, it would have to manually go to each file. This paper considers a method to extract metadata relevant to the investigators, while allowing for all metadata in the case to be collated into one location. This master metadata file also contains sections meant for project status and task management, which ensures investigators always know who has done what and when and the current status—thus indicating what actions have been taken on the case.

While there is software capable of extracting metadata from evidence files (Garfinkel 2010), it tends to require a high level of technical understanding and is limited to metadata of the specific evidence files. This paper is focused on providing metadata that is more comprehensible and will give more value to digital forensic investigators.

The requirements for the content and structure of forensic reports are dictated by the needs of the case. These can vary significantly from case to case and from jurisdiction to jurisdiction, which often results in a manual process being used to organise evidence in a consistent manner for review. For instance, each evidence item must be mapped to the correct custodian, with correct item size, correct hash, and correct time. Related attachments must be created and cross-checked to ensure correct content and position in the report. In large commercial cases, the manual process can prove time-consuming and can increase the possibility of human error. Details may be retyped, although best practice may require stringent quality controls, including double-checking by additional personnel which will increase cost and effort.

Reporting the findings of forensic analysis is an essential stage in completing an investigation. And both major tool suites AccessData FTK (AccessData 2018) and Guidance Software's Encase (Guidance Software 2018) have varying degrees of reporting features and abilities to export information found in the investigation as a formatted report. However, investigations rarely use one individual tool, and so investigators may often be required to use multiple tools with different formats. Other examples of similar tools include the database tool MARS outlined by Watson and Jones (2013), which is used to capture data relating to a case for both management and reporting purposes. A comprehensive tool for recording case information, MARS is still in the form of a database and still requires users to manually supply some aspects of the necessary information.

Issues of Case Management and Reporting

This paper proposes a lightweight approach to two areas. The first is collating metadata for forensic case management and the second is the related issue of evidence report referencing. The keys issues of forensic case management and evidence report referencing can be outlined in two comparative scenarios:

Consider a digital forensic case with 50 custodians; each custodian has three electronic items in addition to email accounts, archives of email[,] and network folders. After acquisition is complete the investigators return with 300 forensic images. The case has a short deadline, so five investigators are assigned full time to the case. They have to categorize the forensic images into a folder structure and add metadata for each of the 300 forensic images manually into a spreadsheet that pro-

vides details such as: case name, custodian, acquisition location, investigator who performed the acquisition, size of forensic image, hash of image and all other metadata need to be documented properly and validated. Once the documentation is complete, the work of each investigator would need to be verified as accurate and correct. The number of investigator and work hours required to complete the task is an accepted consequence of the amount of forensic evidence and how much time is available for the case. (Nor, Sutherland, & Blyth 2018)

Consider the same scenario using a different approach:

Investigators return with the 300 forensic images, but this time only one investigator continues the work. The other four investigators continue working on another open case, while one investigator places all 300 forensic images into a single folder, with hash; unstructured. The investigator then makes available some prerequisite information like custodian list, and runs an automated forensic case management program. This organises 300 forensic images, which are placed into custodian folders, based on type. In addition there is produced one master Excel file, which contains all the required metadata for each forensic image, linked to the correct custodian. Quality control consists of appropriate examination and testing of the automation program and a need for a more limited number of quality control checks. (Nor, Sutherland, & Blyth 2018)

Comparing the scenarios demonstrates that the investigator's expertise and efforts should be focused on those activities related to the investigative process, data recovery, and correlation of evidence rather than the manual labour of data entry. The current documentation process in digital forensics is quite extensive (Talib 2015) and could be reduced with appropriate automated processes.

An additional issue concerns how evidence is to be presented in a report. An investigator might use the report function in his or her respective digital forensic software, such as Encase or FTK, and append such auto-generated reports to the main delivery. Then the investigator would subsequently make references to evidence in the auto-generated report in the main delivery. In a more integrated format, others might export all evidence with evidence ID names and attach it to the report, referencing a specific item number. The reader would then have to open an attached folder and find the evidence item with the corresponding number to view the referenced file. At some point the reader needs to check another location to view the evidence referenced. So, in addition to the proposed automated forensic case management tool, a more convenient solution to the evidence report referencing problem is proposed, which automatically links evidence directly into the report. The reader will then only have to click the evidence name inside the report for it to open for immediate review. These two solutions provide an example of one way to streamline this aspect of the forensic process.

Methodology

Two prototype tools were created, one to address case management and one to handle the reporting aspect of the case.

Automated forensic case management

This program was modular in design with the prototype program being implemented in Python. Each module is designed to represent a separate step in the forensic case management progression;

depending on the case, some of the modules may not be relevant and are, therefore, not used. Preliminary information on Case ID, Project Manager, and prerequisite documentation are collated at the start of the process and form part of the documentation produced by the program. It is also transmitted to a database to keep track of projects and case status.

The prototype has the following modules:

Module 1: Takes the initial case information as input and updates the master database. It then creates a pre-defined standardised forensic case folder structure. This empty template will later receive the forensic images in their corresponding folders.

Module 2: Takes a list of custodians as input and updates the master database. It then creates custodian folders into the standardised forensic case folder structure created in Module 1. This means that, if an image belongs to John Doe, there will be created a new folder in the custodian section called 'John Doe', into which all evidence belonging to him will be placed. In addition, it creates a standardised evidence folder structure within each custodian folder.

Module 3: Asks if all forensic images are placed unstructured in a special folder. It will then crawl through all hash files for the hash value for each evidence and store it in a temporary database. It will then create a master list which contains all forensic evidence files (including forensic image

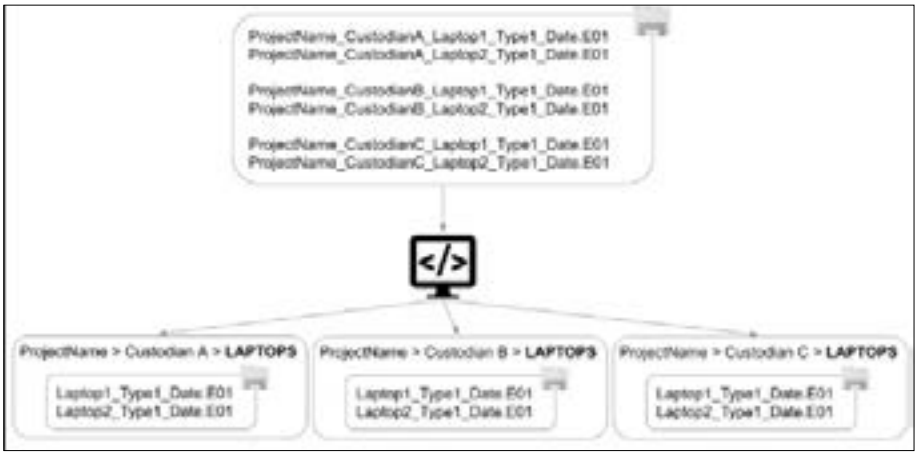


Figure 1: Module 3, automatic sorting of evidence files

segments) in order to start moving all forensic images into their respective custodian folders and into their respective evidence folder structure (see **Figure 1,** above).
Therefore, this module sorts all forensic images into the standardised folder structure automatically.

Module 4: During the operations in Module 3, all evidence metadata/information is stored in a temporary database. When Module 3 is complete, Module 4 starts extracting all information from that temporary database and populates an Excel template. This Excel template becomes the master forensic evidence file which contains all useful information about all forensic evidence. It does this by crawling the information files for each piece of forensic evidence and decoding the information from the specialised naming convention. Is the evidence live memory acquisition? Logical Acqui-

sition? Where was it acquired? Who performed the acquisition? All questions about the evidence are neatly organised into the Excel file.

In addition to the evidence information sheet in the Excel file, there are also several sheets created for the purpose of tracking progress during the case. Who performed OCR? When? What was last worked on? Why was this task skipped? These template sheets make it easy for several digital forensic investigators to work on the same case without having to get a full update directly from the investigator; instead they just get the master evidence file. This file can be provided to clients, judges, and/or managers, and can be read by anyone. (See **Figure 2,** below.) In the long term, a web interface is planned to give access to the same information.

Module 5: This module is designed as the exit module, or case-closing module. If an investigation is complete and the report has been sent, it is very important to have procedures for proper closing of a case, with integrity of evidence preserved. This module will first go through and check the

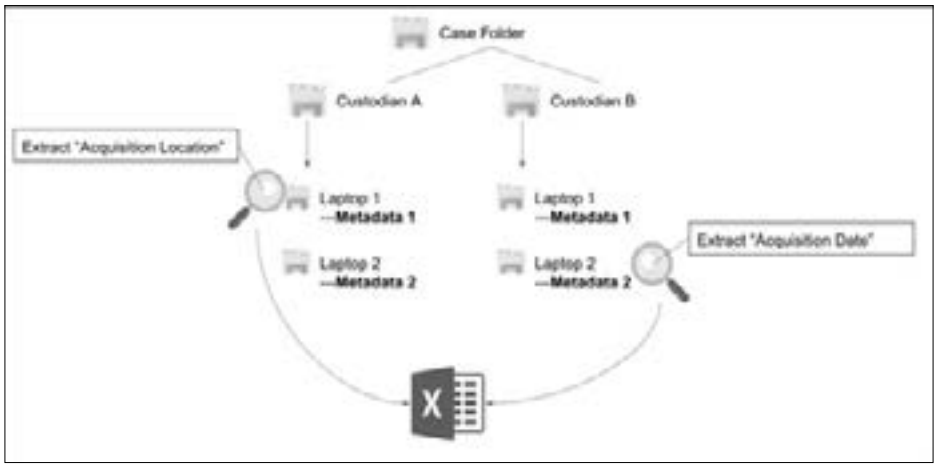


Figure 2: Module 4, metadata extraction

'Forensic Procedure Status' sheet in the master Excel file. If something deviates from mandatory protocol, a warning will be issued saying the case cannot be closed. After that, the module will ask some exit questions, which must be completed in order for the case to be closed.

Evidence Report Referencing

This program is separate from the case management program and is aimed at improving the readability of the reports. There are numerous different evidence report formats, as a result of the final report being assembled from input from a range of expert opinion, including legal and technical. The purpose of this software is to improve readability and scalability for large reports by using a single PDF report with all evidence integrated into the report using hyperlinks to reference additional PDF files. Closing the evidence file, the reader returns to the report as before.

The Python program takes the finished Word document as input, converts all evidence names into interactive links, and saves it in Word format. The next step is to export the document as a PDF after any final edits, with the interactive code embedded into the document. (See **Figure 3,** below.) There are only two required parameters needed for this process to work: 1) the report must be written in Microsoft Word and 2) the evidence must be referenced using a Globally Unique Evidence

Identifier (GUEI). Those who write reports in other formats, such as Google Docs, can easily export their documents into Word at the end of the project. The GUEI is required because the Python software uses a search and replace function to track down an evidence reference. Once found, the reference is substituted with a hyperlink to the referenced evidence with the same GUEI name.

Testing and Results

Automated forensic case management

The prototype tool was tested against a test case. This test case consisted of ten custodians with ten forensic images each. The 100 files were given test acquisition information files, simulating

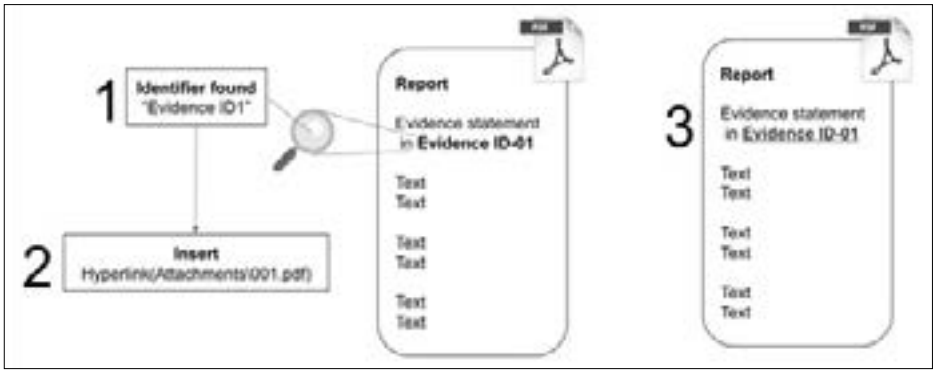


Figure 3: Overview of evidence report referencing

those generated during acquisition, with hash values and general information. The forensic files were all placed in the same folder with no structure or organisation—just 100 files in a folder. After providing the prerequisite information to the program, such as the custodian list, the program ran without incident.

As expected, all 100 files were moved into their respective forensic type folders in their respective custodian folder. This means that each custodian received one forensic file in ten different forensic type folders. For example, one folder might be Logical Laptop Image, while another might be Logical Mobile Image. The database accumulated information about each of the forensic files and the data into the Excel template. The program runtime was less than 15 seconds.

The resulting master Excel file contained all metadata for all evidence in the case, in addition to a blank set of Excel tabs containing the procedure that will follow in the case. This process made not only everything centralised for all involved in the case, but also provided a step-by-step documentation template that showed who did what and when.

Evidence report referencing

During the report testing phase, a Word document was given random data in the form of information found on Wikipedia. It does not really matter what the information is as long as the correct referencing is added to the text. By using markers such as Object ID (OID) or Evidence ID (EID) when referencing evidence (for example, OID5501 or EID5501) the program crawls through the report and replaces the references with actual hyperlink references. The Word file was given to the Python program as input and, after a few seconds of runtime, a new Word document was produced, which contained interactive code on all references in the document.

Discussion

Case management in digital forensics can quickly become demanding. The initial documentation is vital because it forms the initial integrity of the case. For this reason, it would make more sense for the initial documentation to be done automatically.

There are also some additional advantages of using this type of methodology. Accumulating case data into a master database enables investigators to use this information for additional practical purposes. While the generated Excel files are of great use to the cases, the middleware database could serve as a source of live case status. By creating a program that runs as a Web service, the program could keep an up-to-date view on not just the cases in a broad sense, but the program would also show the date, time, and investigator involved. Furthermore, the program could also provide a detailed view of all investigators working on the case, what each is doing or has done, as well as each and every piece of evidence. This information would be available because the master Excel file also includes blank sheets for case progression to be filled out by the investigators themselves. All this information could be shown on a large TV screen in the forensic lab, which would allow all investigators to quickly see how things are going or allow managers and supervisors to keep an eye on progress. This way, the large amount of information accompanying forensic evidence can go from being a slow manual-labour task performed at every project start to becoming an automatic process with the bonus of showing an automated progress with details available.

A key advantage of the methodology described in this paper is the improved accuracy of metadata. The manual collection and re-entry of case information and evidence metadata increases the potential for error. This is particularly true for large cases, in which the manual processing of hundreds of large numbers and names increases the probability of errors occurring despite quality-control measures being in place. Some cases may never get to court, but others might go to court after several years, making it difficult to even identify and to trace errors in the metadata. Finding out years into a case that metadata was incorrectly approved in the beginning can hurt or even contribute to the loss of said case. Having an automatic process to gather, copy, and change case information and evidence metadata adds a level of confidence to the process and provides one less avenue for legal challenges to the evidence.

Summary and Conclusions

The use of automatic programs in digital forensic case management has a number of advantages, in particular improving the efficiency of evidence processing in larger cases. The processing of metadata associated with digital evidence can provide access to a rich amount of information that can be extracted for reporting purposes. Another advantage of using automated programs is that the programs establish standardised file structures and forensic procedure templates. As digital forensics processes become increasingly integrated with people working in different fields, there is a growing need to make digital forensic information more accessible. Metadata spread across multiple files may have limited value for a lawyer or a CEO. This paper outlines a simple step towards unifying information from all aspects of case metadata. An organised overview of the case data, whether it is for a judge or a CEO or for other digital forensic investigators, makes the investigation easier to manage.

Digital Forensic Investigations are often complex during the initial steps in the process. After the

initial work is completed and the data has been prepared for viewing, there is a window of time during which other experts of a non-technical nature (lawyers and accountants, for example) can access and analyse the content. But understanding where evidence came from, its time zone, who handled it, or when certain items were collected compared to others are foundational to obtaining a clear overview of the case. In this paper, a new methodology to improve the accuracy of the management aspects of the entire digital forensic process has been explored.

Acknowledgements

The authors would like to thank the participants of the European Conference on Cyber Warfare and Security in Oslo, June 2018 for useful feedback on an earlier version of this paper: Nor, G, Sutherland, I & Blyth, A 2018, 'Concepts of Automating Forensic Case Management', Proceedings of the 17th European Conference on Cyber Warfare and Security, ECCWS 2018, pp. 338-342.

References

Accessdata 2018, *Forensics tool kit*, viewed on 7 April 2018, <https://accessdata.com/>.

DFRWS 2006, *Survey of disk image storage formats*, viewed 27 March 2018, <http://www.dfrws.org/sites/default/files/survey-dfrws-cdesf-diskimg-01.pdf>.

Expert witness compression format 2018, Github, viewed 9 May 2018, <https://github.com/libyal/libewf/blob/master/documentation/Expert%20Witness%20Compression%20Format%20(EWF).asciidoc>.

Garfinkel, S 2010, AFF and AFF4: *Where we are, where we are going, and why it matters to you,* Open Source Digital Forensics Conference (OSDFCON), viewed on 9 May 2018, <https://www.osdfcon.org/presentations/2010/garfinkel-afflib.pdf>.

Garfinkel, S, Malan, D, Dubec, K, Stevens C & Pham C 2006, 'Advanced forensic format: An open, extensible format for disk imaging', *Proceedings of the 2nd annual International Conference on Digital Forensics, IFIP WG 11.9,* Orlando, FL, US.

Gayed, TH, Lounis H & Bari M 2012, 'Cyber forensics: Representing and (im)proving the chain of custody using the semantic web', *Proceedings of the 4th international conference on Advanced Cognitive Technologies and Applications, COGNITIVE 2012*, Nice, FR.

Guidance Software 2018, *Encase,* viewed 7 April 2018, <https://www.guidancesoftware.com/ encase-forensic>.

Nor, G, Sutherland, I & Blyth, A 2018, 'Concepts of automating forensic case management', *Proceedings of the 17th European Conference on Cyber Warfare and Security, ECCWS 2018,* University of Oslo. Oslo, NO, p. 338-42.

Spyrus 2018, *Other security offerings,* viewed 17 Nov 2018, <https://www.spyrus.com/other-security-offerings/>.

Talib M 2015, 'Studying the documentation process in digital forensic investigation frameworks/ models', *Journal of Al-Nahrain University,* vol. 18, no. 4, pp. 153-62.

Turner P 2005, 'Unification of digital evidence from disparate sources', *Digital Forensics Research Workshop,* New Orleans, LA, US.

Watson D & Jones A 2013, *Digital forensics processing and procedures: Meeting the requirements of ISO 17020, ISO 17025, ISO 27001 and best practice requirements,* Syngress, Waltham, MA, US.

Towards a Literature Review on Cyber Counterintelligence

PC Duvenage, VJ Jaquire, and SH von Solms

Centre for Cyber Security
Academy of Computer Science and Software Engineering
University of Johannesburg
Johannesburg, South Africa

E-mail: duvenage@live.co.za; jaquire@gmail.com; basievs@uj.ac.za

Abstract: *For those connecting the dots, the threat landscape continues to affirm the necessity of having Cyber Counterintelligence (CCI) at the centre of cybersecurity efforts. Concurrent with the growing interest in CCI in corporate boardrooms and the corridors of governments, CCI is evolving from a field of academic enquiry to a distinctive academic sub-discipline. The growing body of CCI-focused literature clearly attests to this evolution. A review of such literature has self-evident academic and practical benefits. This article advances a tentative, selective review of CCI literature that demonstrates the need for a more extensive and in-depth appraisal.*

Keywords: *Cyber Counterintelligence, Cyber Warfare, Cyber Security, Literature, Theory, Denial and Deception*

Introduction

For state and non-state actors with sizable cyber interests, numerous breaches during this decade have affirmed the necessity of having Cyber Counterintelligence (CCI) at the centre of cybersecurity efforts (Prunckun 2018; Stech & Heckman 2018; The Economist 2015). Concurrent with the growing interest in CCI in corporate boardrooms and in the corridors of governments, CCI is evolving from a field of academic enquiry to a distinctive academic sub-discipline. Attesting to the growing interest in CCI is the expanding body of peer-reviewed, academic contributions specifically focused on CCI. These contributions include numerous conference papers (such as Sigholm & Bang 2013; Jaquire & von Solms 2017a-c; Duvenage, von Solms & Corregedor 2015) and several completed post-graduate studies (for example Knowles 2013; Black 2014; Fieber 2015; Putnam 2015; Justiniano 2017; Jaquire 2018; Duvenage 2018). As will be shown in this article, commercial literature on CCI has also been growing sharply in recent years.

A literature review not only has self-evident benefits for CCI's academic progress, but it will also be useful to the increasing number of practitioners specialising or interested in this area. Attempting a comprehensive and representative literature review within the confines of a single journal article would be over-ambitious. Moreover, in the case of a field as young as CCI, such an extensive review would arguably be pre-mature. Therefore, this article's aim is to submit a tentative, selective literature review on CCI. Such a pilot literature review reveals the need for a much more comprehensive and inclusive appraisal of CCI literature.

The rest of the article is structured as follows:

- First, the purpose and benefits of a selective literature review on CCI are discussed in more detail.
- Secondly, the scope and the nature of the selective literature review are defined.
- Subsequently, the literature review is presented with reference to four literature categories, namely (1) peer-reviewed papers and articles, (2) masters' and doctoral studies, (3) books and (4) other literature.
- Finally, the conclusion submits key findings and observations regarding the way forward.

The Purpose and Benefits of a Selective Literature Review on CCI

Within academic research in general, there are various types of literature reviews which serve different purposes (Grant & Booth 2009; Mallett et al. 2012; Kim 2018). Some better-known examples (of review types) include argumentative, integrative, historical, systematic, methodological, and theoretical reviews. From these types, systematic reviews have, for good reasons, been gaining prominence in academic circles (Mallett et al. 2012; Grant & Booth 2009). While a systemic review has many benefits, its compilation is an exhaustive and extensive process.

As suggested earlier, even for an academic sub-discipline as young as CCI, it would have been over-ambitious at this stage to endeavour to create a rigorous systematic review of CCI literature and to present the outcome thereof in a single journal article. In a similar vein, the tentative overview presented in this paper does not purport to adhere to the requirements of one of the other review types cited above. Instead, the article follows a less formalistic and selective approach in its review of CCI literature. This selective approach—further scoped and qualified in the next section ('Qualifying the Nature and Scope of the Selective CCI Literature Review')—provides several benefits:

- It highlights salient contributions to CCI that are of significant practical and/or academic importance;
- It provides some contours of the state of knowledge and the key directions of CCI research;
- It establishes a 'scaffold' for identifying and positioning future research topics;
- It provides a premise for a more comprehensive, systematic CCI literature survey;
- Because it deals with salient research done thus far, it offers an insight into CCI's academic origin, emergence, and development. As is the case with other academic subjects, such a self-awareness of origin and evolution could contribute to consolidating CCI as a distinctive sub-discipline; and
- It identifies research projects/institutions focused on CCI and, by so doing, encourages academic interaction in this field.

Qualifying the Nature and Scope of the Selective CCI Literature Review

This article has thus far emphasised the 'selective' nature and scope of the CCI literature review to be advanced. For the review to be academically credible, the meaning of the word 'selective' needs to be clarified. The review of literature advanced in this article is selective in that it limits its focus in the following five respects:

1) 'Available literature' is deemed as works in the public domain. Due cognisance is taken of the fact that state security structures internationally generate and possess CCI-relevant research and training material, some of which is unclassified but not freely available. The same applies to some corporate entities and cybersecurity vendors that, for various reasons, do not openly share CCI material. Such material is categorically excluded from this review.
2) 'Available literature' is secondly deemed as referring to work published in English. The search which informed the review did not cover untranslated CCI-research possibly published in other languages.
3) The literature review is furthermore 'selective' in that it predominantly focuses on material which explicitly addresses CCI. While overlapping themes (such as cyber denial and deception, insider-threat mitigation, cyber intelligence, and cyber-threat intelligence) are important to CCI, a review of such literature would distract from the article's aim. For purposes of the article, CCI—which constitutes the literature overview's referent object—is defined as that sub-discipline of counterintelligence (CI) "aimed at detecting, deterring, preventing, degrading, exploiting and neutralis[ing] adversarial attempts to collect, alter or in any other way breach the C-I-A of valued information assets through cyber means" (Duvenage, von Solms & Corregedor 2015).
4) The literature review is selective in that it does not purport be an inventory of all CCI-focused work. Instead, in terms of academic works, the review reflects on peer-reviewed, published work featured in selected platforms, namely Scopus, EBSCO, Institute of Electrical and Electronics Engineers (IEEE), Explore, Springer Link, Google Scholar, and Proquest.
5) Lastly, the literature review only covers selected contributions published as of 30 April 2018.

Moving from an overview of the selective scope of this literature review, the next section explains the structural approach to be followed.

Structural Approach to the Selective CCI Literature Review

A literature review should, of course, be structured in a manner optimally achieving its aim and benefits. Given this literature review's earlier discussed aim and benefits, structuring the review per either (a) literature category or (b) chronology of publication was considered. On the one hand, the conventional approach of dividing reviews per literature category (such as articles, masters' and doctoral studies, books) would arguably have been the best suited to plot existing and to provide a scaffold for positioning future CCI research. On the other hand, a chronological literature review would be more effective to convey CCI's academic origin and development. To draw on the advantages both these styles offer, this article opted for a hybrid approach which incorporates a chronological thread with literature type. Practically, this means that the review overall is structured per the literature categories, namely peer-reviewed articles and papers, masters' and doctoral studies, books, and other literature. However, since the bulk of CCI academic work was produced per peer-reviewed articles and papers, this literature category (peer-reviewed articles and papers) is presented chronologically in order to convey CCI's origin and evolution.

This hybrid structural approach to the selective CCI literature review is depicted in **Figure 1,** below.

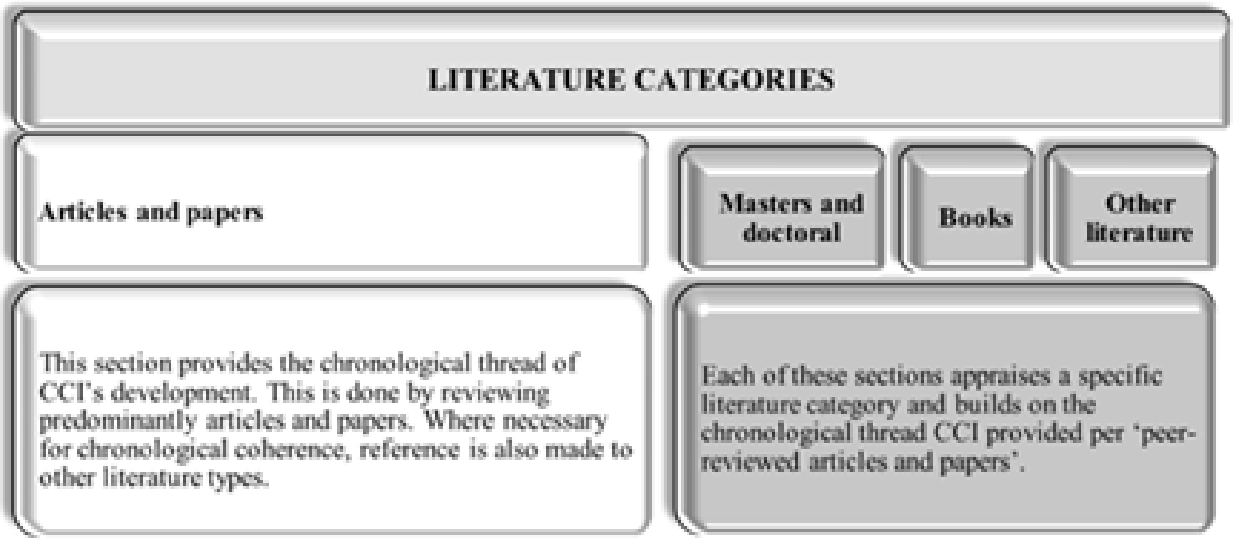


Figure 1: Structural approach to the selective literature review on Cyber Counterintelligence

Peer-Reviewed Articles and Papers

In line with **Figure 1**, this section enumerates CCI's evolution with specific reference to peer-reviewed articles and papers. Although somewhat of an over-simplification, CCI's progression as a distinctive academic sub-discipline consists of the following phases:

- Foundational phase (pre-2009),
- Phase in which Cyber Counterintelligence's emerged as an academic research theme (2009-2012),
- Current stage in which Cyber Counterintelligence crystallised into a distinctive academic sub-discipline (2012-present).

Foundational phase (pre-2009)

As far as could be surmised from available literature, the explicit term 'Cyber Counterintelligence' first emerged in the American statutory security establishment during the early 2000s (see United States 2004; French & Kim 2009). Prior to the 2000s, however, CCI was practiced in the statutory security establishment of the U.S. and the security structures of some other countries. In this regard, French and Kim (2009) rightly assert that "cyber CI has existed de facto since the introduction of IT to intelligence, defence, and national security and has grown as FISs [Foreign Intelligence Services] have embraced cyber tradecraft".

Concurrent with CCI's de facto existence in statutory security circles, a few sporadic academic articles in the 1980s and 1990s expounded key CCI notions—although without using the actual term 'Cyber Counterintelligence'. Such notions included advocating for an integrated CI approach, which not only has defensive and offensive missions, but which also synchronises human and technical resources. The earliest peer-reviewed article found in consulted literature referring to such application of a CI approach to the IT realm is contained in the electronic library of the Institute of

Electrical and Electronics Engineers (IEEE). This item, authored by Stone and Tucker (1988), is entitled 'Counterintelligence and unified technical security programs in security technology'. The authors expound effective CI as a "unified multi-disciplinary concept" consisting of "proactive and defensive" missions. Stone and Tucker (1988) further argue that "advanced technology" is part of the multi-disciplinary CI entirety and thus serves both "proactive" (offensive) and defensive missions. While Stone and Tucker's (1988) paper centres on rectifying perceived deficiencies in the U.S. national CI endeavour thirty years ago, their key contentions regarding an integrated CI effort are still relevant today.

In a related further contribution in the IEEE library, Stone and Bluitt (1993) further expanded on the idea of executing "advanced technological countermeasures" as part of "a pervasive counterintelligence (CI) mandate". Also, Stone and Bluitt (1993) directed their paper specifically at the U.S. statutory CI effort.

No articles or papers of direct CCI-relevance were found in consulted literature for the seven-year period from 1994 through 2001. The first peer-reviewed article that specifically employs the term 'cyber' in conjunction with 'counterintelligence' appeared in a 2002 issue of the Journal of Information Warfare. As suggested by the title of the article, 'Dominating the attacker: Use of intelligence and counterintelligence in cyber warfare', Davey and Armstrong (2002) examined Intelligence and CI's role in augmenting cyber warfare. Cyberwarfare, in turn, is firmly positioned as a subset of Information Warfare. By "employing intelligence and counterintelligence techniques that are superior to those of the attacker", argue Davey and Armstrong (2002), the "cyberwarfare defender" is more likely to prevail. Davy and Armstrong also urge a more "aggressive" posture that includes deception. One such example cited includes allowing the "attacker [to] gain access to information that is actually incorrect, thus providing incorrect intelligence". In respect to CCI's conceptual evolution and especially CCI's relation to cyber warfare, the contribution of Davy and Armstrong (2002) represents a milestone.

Like the work of Stone and Tucker (1988) and Stone and Bluitt (1993), Davey and Armstrong's 2002 article was part of CCI's foundational phase which, if gauged by academic publications, was characterised by only a few sporadic contributions. Insofar as consulted literature goes, no CCI-relevant publications appear for the next five years (2002-2008).

Cyber Counterintelligence's emergence as a research theme (2009-2012)

Following a sporadic foundational phase, 2009 marked CCI's emergence as a specific research theme attracting growing interest. In that year, a seminal article appeared in the launch edition of the National Intelligence Journal (French & Kim 2009). This was the first academic publication (in consulted literature) to use the term 'Cyber Counterintelligence'. In 'Acknowledging the revolution: The urgent need for Cyber Counterintelligence', French and Kim (2009) call on the U.S. intelligence community to move away from the notion that CCI is mostly part of "defensive Information Warfare". Instead, French and Kim (2009) urge the U.S. to be more active and offensive in its approach to CCI. The work's relevance extends beyond the U.S. context. French and Kim (2009) explicitly define CCI, explain CCI's missions within the context of CI, and offer various other insights on aspects useful to the further development within this field. Such aspects include the role of CCI in Information Warfare, critical infrastructure protection, and the CCI process and strategy.

No other peer-reviewed articles and papers were found in consulted literature for the period between 2009 and 2012. It must, however, be emphasised strongly that the absence of academic articles on CCI in consulted literature belies CCI's emergence as a research theme for three reasons. First, there were several CCI contributions during this period in other literature categories (see subsequent section entitled 'Other literature') and in publications not covered by this article's selective review (see, for example, U.S. Naval War College 2018, 'Counterintelligence: Cyber Threat'). Thirdly, the nature and extent of academic contributions regarding CCI from 2013 onward strongly suggest that CCI attracted research interest in the preceding years (2009-2012). Phrased differently, research was done in the 2009-2012 timeframe, but the fruits thereof, in the main, are only reflected from 2013 onward.

Cyber Counterintelligence's crystallisation as an academic sub-discipline (2013-present)

From 2013, a consistent stream of peer-reviewed papers and articles signalled CCI's emergence as an academic sub-discipline with significant contributions in English from researchers in the U.S., Sweden, and South Africa.

The bulk of academic contributions from the U.S. stemmed from Utica College's Master of Science Cybersecurity programme that offers CCI as a specialisation subject. This programme resulted in several 'capstone project' papers (comparable to mini-dissertations in other countries) as well as a thesis with CCI as a specific focus (Knowles 2013; Black 2014; Fieber 2015; Putnam 2015; Justiniano 2017). Since these contributions flow from a master's programme, they are discussed in more detail in a later section, which focuses on masters' and doctoral studies). Suffice it to state here that this Utica research constitutes indispensable contributions to CCI on the conceptual, theoretical, and praxis levels.

Also, in recent years in the U.S., the concept of CCI has attracted interest from researchers at the Mitre Corporation. Branching out from its leading research on denial and deception in active cyber defence, the company subsequently examined "applications of cyber counterintelligence" to "cyber defense" (Heckman et al. 2015; Stech & Heckman 2018). Flowing from this research, Stech and Heckman (2018) contributed a book chapter, which is undoubtedly one of the most incisive and significant works on CCI to date. (This contribution is discussed in more detail under 'Books'.)

Albeit considerably more limited in scope than the research in the U.S., papers delivered at two IEEE-endorsed conferences in 2013 reflected growing interest also outside the U.S. In August 2013, at the European Intelligence & Security Informatics Conference in Sweden, Sigholm and Bang (2013) submitted a paper entitled 'Towards offensive cyber counterintelligence: Adopting a target-centric view on Advanced Persistent Threats'. Coming from a statutory military perspective, the paper is primarily aimed at advancing a "comprehensive process that bridges the gap between the various actors involved in CCI". Sigholm and Bang (2013) present this model to specifically configure the "offensive CCI attribution process". The model essentially consists of all-source information flow and analysis architecture to be employed for attribution purposes.

On the heels of Sigholm & Bang in 2013, Duvenage & von Solms (2013) presented 'The case for cyber counterintelligence' at the 5th International Conference on Adaptive Science and Technolo-

gy in South Africa. The paper defines key CCI concepts and advances conceptual constructs which explain CCI and its relation to CI.

Duvenage and von Solms' (2013) paper formed part of a dedicated CCI research project initiated at the University of Johannesburg's Cybersecurity Centre (UJCC) from which several other contributions would follow (University of Johannesburg 2018). UJCC's website describes the project's aim as establishing CCI as a multi-disciplinary field of academic enquiry within the South African context (University of Johannesburg 2018). To this end, the UJCC project pursues two complementary yet parallel research streams, aimed respectively at

1) Designing an overarching framework for conceptualising and explicating CCI as a distinctive academic field of enquiry; and
2) developing a framework for a CCI maturity model for application by state and non-state actors within developing countries.

Building on Duvenage and von Solms' 2013 contribution, UJCC's first research stream progressively advanced conceptual constructs to explain (in an academic context) what CCI is, how it works, and how it dovetails with other academic disciplines and theory. Such notional constructs include a CCI-posture matrix model and a CCI process model, as well as a taxonomy of CCI Tactics, Tools, Techniques, and Procedures (TTTPs). These notional constructs were submitted per the following peer-reviewed papers and a journal article:

- Duvenage and von Solms (2014), 'Putting counterintelligence in cyber counterintelligence' in Proceedings of the 13th European Conference on Cyber Warfare and Security, Piraeus, Greece.
- Duvenage and von Solms (2015), 'Cyber counterintelligence: Back to the future' in the Journal of Information Warfare.
- Duvenage, von Solms, and Corregedor (2015), 'The cyber counterintelligence process – A conceptual overview and theoretical proposition' in Proceedings of the 14th European Conference on Cyber Warfare and Security, Hatfield, UK.
- Duvenage, Jaquire, and von Solms (2016), 'Conceptualising cyber counterintelligence – Two tentative building blocks' in Proceedings of the 15th European Conference on Cyber Warfare and Security, Munich, DE.
- Duvenage, Sithole, and von Solms (2017), 'A conceptual framework for cyber counterintelligence—Theory that really matters!', Proceedings of the 16th European Conference on Cyber Warfare and Security, Dublin, Ireland.

UJCC's second research stream, to recapitulate, aims to develop a CCI maturity model with emphasis on governments and non-state actors in emerging countries (University of Johannesburg 2018). Peer-reviewed papers presented in this regard are as follow:

- Jaquire and von Solms (2017a), 'Towards a cyber counterintelligence maturity model', in Proceedings of the 12th International Conference on Cyber Warfare and Security, Wright State University, Air Force Institute of Technology, Dayton, OH, U.S.
- Jaquire and von Solms (2017b), 'Developing a cyber counterintelligence maturity model

for developing countries' in Proceedings of the 2017 IST–Africa Conference, Windhoek, Namibia.
- Jaquire and von Solms (2017c), 'Cultivating a cyber counterintelligence maturity model' in Proceedings of the 16th European Conference on Cyber Warfare and Security, Dublin, Ireland.

This section examined CCI's academic evolution via an overview of peer-reviewed articles and papers. The next section explores contributions to the field in the form of masters' and doctoral research.

Masters' and Doctoral Studies

The search term 'Cyber Counterintelligence' (and variations thereof) showed numerous masters' and doctoral studies of possible relevance to a CCI literature review. On closer analysis, however, most of these studies do not have CCI as a primary focus, and CCI is not explored in depth. Instead, CCI is cursorily referred to as part of the broader statutory CI mandate and mostly addressed within challenges faced by the U.S. intelligence community. Ferguson's (2012) thesis entitled Increasing the effectiveness of U.S. counterintelligence: Domestic and international micro-restructuring initiatives to mitigate cyber espionage serves as one such example.

Bucking this trend, masters' studies completed at Utica College from 2013 onwards delivered contributions that are pioneering and invaluable with respect to the academic crystallisation and evolution of CCI. As was noted earlier, these studies are mostly 'capstone projects'. Conducted within the context of U.S. national interests and security, these studies have much broader application and academic relevance than just in the U.S. Overall, important contributions have been made to explicating CCI on the conceptual, theoretical, and praxis levels. The following are some examples:

In his research entitled Applying computer network operations for offensive counterintelligence efforts, Knowles (2013) identifies key aspects of Computer Network Operations (CNO). These aspects are then aligned with the broader intelligence and CI processes. In so doing, "counterintelligence skills and techniques" are leveraged to "assimilate cyber activities" into an organisation's Intelligence endeavour.

- Effective CCI, argues Black (2014), is multidisciplinary and involves unique skill sets. In his thesis entitled The complexity of cyber counterintelligence training, Black proceeds with identifying the implications thereof for CCI training. Black then advances two useful notional constructs, namely (1) a CCI training model and (2) a CCI training proficiency path.
- As suggested by the research title, Putnam's (2015) Digital mirrors casting cyber shadows - The confluence of cyber technology, psychology, and counterintelligence emphasises CCI's multidisciplinary nature. Putnam points out that a successful CI (and thus CCI) programme should consider the opportunities that technology presents as well as certain psychological "principles of persuasions" and motivation. The study details some offensive and defensive CCI applications of these opportunities and principles. Emphasis is placed in this regard on optimizing the CCI targeting and the recruitment processes.
- The interplay between practice and theory which characterises Utica College's research is reflected in Fieber's (2015) commendable contribution: The Iranian computer network

operations threat to U.S. critical infrastructures. Fieber analyses "the Iranian computer network operations (CNO) threat to U.S. critical infrastructures" and proceeds with recommending defensive measures to mitigate this threat. The paper culminates in a handy proposition on a phased, CCI process model "designed to mitigate conditions favorable to the attacker and restore the advantage to the organizational defenders" (Fieber 2015).

- Justiniano's 2017 outstanding and pioneering contribution, entitled Advancing the capacity of a Theatre Special Operations Command (TSOC) to counter hybrid warfare threats in the cyber gray zone, examines CCI's role in the U.S. military milieu with a focus on the hybrid threats posed by Russia and the role of CCI in mitigating and engaging this threat. Justiniano's (2017) research is indispensable reading for examining CCI's role in hybrid warfare more generally. The study identifies critical CCI roles and skillsets before proceeding to propositions on integrating CCI with the U.S. "Cyber Mission Assurance (C-MA)" process in a manner supportive of "Theater Special Operations Command (TSOC)".

Although the bulk of post-graduate CCI studies in consulted literature originated from Utica College's Master's programme, the research project of the University of Johannesburg's Cyber Security Centre (UJCC), mentioned earlier, recently resulted in a master's dissertation and doctoral thesis focussing on CCI. These studies, which mirror UJCC's two CCI research streams (discussed in the previous section), are as follow:

- Jaquire (2018) A framework for a cyber counterintelligence maturity model, Doctor of Commerce thesis, University of Johannesburg, South Africa.
- Duvenage (2018) A conceptual framework for cyber counterintelligence, Master of Commerce dissertation, University of Johannesburg, South Africa.

The preceding two sections focused on academic, peer-reviewed literature—which ranges from papers and articles to masters' and doctoral studies. In the next section, books published on CCI are reviewed.

Books

The past two decades have seen an exponential rise in the number of books from reputable publishers dealing with aspects of cybersecurity. However, until very recently, even outstanding books that address aspects of high relevance to CCI make scant reference to CI and CCI. One such example is Heckman et al.'s (2015) Cyber denial, deception and counter deception – A framework for supporting active cyber defense. Despite the likelihood that this work sets the standard for future works on cyber denial and deception in general, only four sentences in the entire book mention the term 'counterintelligence', and there is no mention of 'Cyber Counterintelligence'.

The first book identified by the survey conducted for this article that has a significant CCI focus was published in 2012 with the title Reverse deception—Organized cyber threat counter-exploitation (Bodmer et al. 2012). Pitched as a practical guide for "IT security professionals", this text is highly significant from an academic perspective. The book comprehensively examines the role of CCI in countering cyber threats through the engagement of hostile actors. In addition to describing CCI Tools, Tactics, and Procedures (TTPs), the authors also explore CCI on a conceptual level. This includes postulations on CI missions as well as CCI's interface with CI and other intelligence

fields. In short, Bodmer et al. (2012) is essential reading for any researcher interested in CCI.

The next book to include a pertinent and significant CCI focus appeared under the editorship of Prunckun (2018) and is entitled Cyber weaponry: Issues and implications of digital arms. While the book has several chapters useful to CCI, Chapter Two is specifically dedicated to CCI. Under the title, 'Human nature and cyber weaponry: Use of denial and deception in Cyber Counterintelligence', Stech and Heckman (2018) make a masterful contribution which anyone serious about CCI should consult. The chapter's primary aim is to advance a "cyber counterintelligence framework in active cyber defences". This system is "referred to as the cyber deception chain, to mitigate cyber spy actions within the cyber espionage 'kill chain'" (Stech & Heckman 2018). To lay a foundation for their CCI framework, Stech and Heckman explain the need for CCI. They proceed by appraising CI definitions, status, and existing frameworks with a view on application to active defence in CCI. The text also observes the existing body of CCI academic research. Proceeding from this basis, Stech and Heckman (2018) present their CCI framework for "active cyber defense". This framework applies and synergises earlier postulations by Duvenage and von Solms (2014) and Prunkun (2014). Stech & Heckman (2018) demonstrate the framework's application by means of a hypothetical case involving the North Atlantic Treaty Organisation (NATO) and the Russian Federation.

Other Literature

In the past eight years, there has been an upsurge in literature dealing with 'threat intelligence', 'cyber intelligence', and 'cyber threat intelligence' (Duvenage, von Solms & Corregedor 2015). Cybersecurity vendors, who are increasingly modelling their products and services on concepts derived from the state security and intelligence realms, in part fuel this upsurge. In contrast to the bourgeoning discourse on, for example, 'threat intelligence' and 'cyber intelligence', contributions to CCI are more limited but are growing. In the main, contributions offer high-level explanations of what CCI is and point to the advantages that CCI practices could have in proactively addressing cyber insecurity. While 'commercial', such works nonetheless contribute to explicating CCI in concrete terms and, in some instances, are consequently also of academic value. In this regard, works by Bardin (2011), Farchi (2012), and Lee (2014) can be singled out.

The following examples of article headlines give a sense of the nature of contributions in commercial online literature:

- 'Cyber counter intelligence' in Defense Tech Magazine (Carrol 2009);
- 'Ten commandments of cyber counterintelligence', first featured on the IDG News Service's online platform CSO Online (Bardin 2011),
- 'Offensive counter-intelligence and cyberwarfare—A paradigm shift in information security' on the Information System Control and Audit Association (ISACA) website (Farchi 2012);
- 'To thwart hackers, firms salting their servers with fake data' in The Washington Post (Nakashima 2013);
- 'Cyber counter-intelligence makes a difference', featured on the South African ITWeb website (von Solms 2014);
- 'Cyber counterintelligence: From theory to practice' first published on the website of the cybersecurity vendor Tripwire (Lee 2014);

- 'Shifting paradigms: The case for cyber counter-intelligence' in InformationWeek (Firestone 2015); and
- 'Counter-intelligence techniques may help firms protect themselves against cyber-attacks' published in The Economist (2015).

While videos are not typically included in literature reviews, CCI's incipient status as well as the merits of a contribution in video format, warrant an exception. This video covers a presentation by Evron (2014), then chairman of the board of the Israeli Computer Emergency Response Team (CERT). This high-level presentation provides a concise, yet incisive and conceptually sharp overview of key CCI fundamentals.

This section reviewed some examples of other literature on CCI. In the section that follows, the article concludes with findings and observations regarding the way forward.

Conclusion

This article advanced a tentative, selective literature review on CCI. This review shows CCI to have evolved, in less than a decade, from a research theme to a distinctive academic sub-discipline. As far as consulted literature is concerned, the bulk of peer-reviewed academic CCI research—documented in papers, articles, and post-graduate studies—was conducted at Utica College (U.S.) and the University of Johannesburg (South Africa). The nature and focus of these institutions' CCI research are inevitably influenced by the respective contexts of a super power (U.S.) and an emerging mid-income country (South Africa). Perhaps because of these differences, the work done is complementary in several respects. Collectively, the research covers diverse topics ranging from general theory and conceptualisation; to CCI training, process models, and maturity frameworks, as well as CCI's application in the military domain. With respect to books and other literature categories, outstanding contributions include works by Stech and Heckman (2018), Evron (2014), Bardin (2011), Farchi (2012), and Lee (2014).

Although CCI is gaining traction internationally, this literature review shows that it is still in its academic infancy and, thus, offers numerous exciting research opportunities. A comprehensive literature review, much broader in scope than this article, would be an invaluable tool for CCI's progression. Such a review would have to cover research in languages other than English and in numerous other databases. Initial research on a comprehensive literature review is being conducted and is already delivering promising results. Those interested in cooperating in this venture are invited to contact the authors of this article.

Acknowledgments

The research presented in this article forms part of a project at the University of Johannesburg's Centre for Cyber Security. More detail can be viewed at <http://adam.uj.ac.za/csi/CyberCounterintelligence.html>.

References

Bardin, J 2011 'Ten commandments of cyber counterintelligence', *CSO Magazine,* 21 June, viewed 7 May 2015, <http://www.csoonline.com/article/2136458/identity-management/ten-commandments-of-cyber-counterintelligence>.

Black J M 2014, *The complexity of cyber counterintelligence training,* Master of Science dissertation, Utica College, New York, US.

Bodmer, S, Kilger, M, Carpenter, G & Jones, J 2012, *Reverse deception—Organized cyber threat counter-exploitation,* McGraw-Hill, New York, US.

Carrol, J 2009, 'Cyber counter intelligence' *Defense Tech,* 9 March, viewed 10 October 2014, <https://www.military.com/defensetech/2009/03/09/counter-cyber-intelligence>.

Davey, J & Armstrong, H 2002, 'Dominating the attacker: Use of intelligence and counterintelligence in cyberwarfare', *Journal of Information Warfare,* vol. 2, no. 1, pp. 23-31.

Duvenage, PC, Jaquire, VJ, & von Solms, SH 2016, 'Conceptualising cyber counterintelligence—Two tentative building blocks', *Proceedings of the 15th European Conference on Cyber Warfare and Security,* R Koch & G Rodosek (eds), Munich, DE, viewed 2 November 2018, <http://adam.uj.ac. za/csi/docs/ECCWS2016_DJVS_PDF.pdf>, pp. 93-103.

Duvenage, PC, Sithole, TG & von Solms, SH 2017, 'A conceptual framework for cyber counterin-telligence—theory that really matters!', M Scanlon & L Neihn-An (eds), *Proceedings of the 16th European Conference on Cyber Warfare and Security,* Dublin, Ireland, viewed 11 December 2018, <http://adam.uj.ac.za/csi/docs/ECCWS_2017_DSV_Prof_MS.pdf>, pp.109-19.

Duvenage, PC & von Solms, SH 2013, 'The case for cyber counterintelligence', *Proceedings of the 5th international conference on Adaptive Science and Technology,* IEEE, T Fogwill (ed.), Pre-toria, South Africa, viewed 7 August 2014, <https://ieeexplore.ieee.org/document/6707493/>, pp. 1-8.

——2014, 'Putting counterintelligence in cyber counterintelligence', A Liaropoulos & GA Tsihrintzis (eds), *Published Proceedings of the 13th European Conference on Cyber Warfare and Security,* Piraeus, Greece, pp. 70-9.

——2015, 'Cyber counterintelligence: Back to the future', *Journal of Information Warfare,* vol. 13, no. 4, pp. 42-56.

Duvenage, PC, von Solms, SH & Corregedor, M 2015, 'The cyber counterintelligence process—A conceptual overview and theoretical proposition', *Proceedings of the 14th European Conference on Cyber Warfare and Security,* N Abouzakhar (ed.), Hatfield, UK, viewed 2 September 2018, <http://adam.uj.ac.za/csi/docs/ECCWS2015_Duvenage %20Von%20Solms% 20Corregedor.pdf >, pp. 42-51.

Duvenage, PC 2018, *A conceptual framework for cyber counterintelligence,* Master of Commerce (Informatics) dissertation, University of Johannesburg, South Africa.

The Economist 2015, Counter-intelligence techniques may help irms protect themselves against cyber-attacks, viewed 24 May 2016, <http://www.economist.com/news/business/21662540-count-er-intelligence-techniques-may-help-firms-protectthemselves>.

Evron, G 2014, *Cyber Counter Intelligence: An attacker-based approach,* Honeynet Project Workshop, Warsaw, Poland, May 2014, viewed 7 October 2017, <https://www.youtube.com/watch?v=l-JC3c-jMALU>.

Farchi, J 2012, 'Offensive counter-intelligence and cyberwarfare—A paradigm shift in information security', *Information System Control and Audit Association (ISACA)*—Blog, viewed 16 February 2016, <http://www.isaca.org/Knowledge-Center/..../Post.aspx?List=ef7cbc6d%2D9997%2D-4b62%2D96a4% >.

Ferguson CJ 2012, *Increasing the effectiveness o U.S. counterintelligence: Domestic and international micro-restructuring initiatives to mitigate cyberespionage,* Master's dissertation, US Naval Postgraduate School, Monterey, California, US.

Fieber, TJ 2015, *The Iranian computer network operations threat to U.S. critical inrastructures,* Master of Science (capstone project), Utica College, New York, US.
Firestone, A 2015, 'Shifting paradigms: The case for cyber counter-intelligence', InformationWeek, 2 April, viewed 7 July 2016, <http://www.darkreading. com/operations/shifting-paradigms-the-case-for-cyber-counter-intelligence/a/d-id/1318929 >.

French, GS & Kim, J 2009, 'Acknowledging the revolution: The urgent need for cyber counterin-telligence', *National Intelligence Journal,* vol. 1, no. 1, pp. 71-90.

Grant, MJ & Booth, A. 2009, 'A typology of reviews: An analysis of 14 review types and associated methodologies', *Health Information & Libraries Journal,* vol. 26, pp. 91-108.

Heckman, KE, Stech FJ, Thomas RK, Schmoker B & Tsow AW 2015, *Cyber denial, deception and counter deception—A framework for supporting active cyber defense,* Springer International Publishing, Cham, CH.

Jaquire, VJ 2018, *A framework for a cyber counterintelligence maturity model,* Doctor of Commerce (Informatics) thesis, University of Johannesburg, ZA.

Jaquire, VJ & von Solms, SH 2017a, 'Towards a cyber counterintelligence maturity model', *Proceedings of the 12th International Conference on Cyber Warfare and Security,* AR Bryant & RF Mills (eds), Wright State University, Air Force Institute of Technology, Dayton, OH, US, pp. 432-40.

——2017b, 'Developing a cyber counterintelligence maturity model for developing countries', *Proceedings of the 2017 IST–Africa Conference,* Windhoek, NA.

——2017c, 'Cultivating a cyber counterintelligence maturity model', M Scanlon & L Neihn-An (eds), *Proceedings of the 16th European Conference on Cyber Warfare and Security,*
Dublin, viewed 11 December 2018, <http://adam.uj.ac.za/csi/docs/
ECCWS_2017_DSV_Prof_MS.pdf>, pp.109-19.

Justiniano, JE 2017, *Advancing the capacity of a theater special operations command (TSOC) to counter hybrid warfare threats in the cyber gray zone,* Master of Science (capstone project), Utica College, New York, NY, US.

Kim, JS 2018, *The importance of literature review in research writing,* viewed 10 January 2018, <https://owlcation.com /misc/literature_review>.

Knowles, JA 2013, *Applying computer network operations for offensive counterintelligence,* Master of Science (capstone project), Utica College, New York, NY, US.

Lee, RM 2014, 'Cyber counterintelligence: From theory to practice', *Tripwire (blog series 4),* 4 May, viewed 4 January 2015, <http://www.tripwire.com/state-of-security/security-data-protection/cyber-counterintelligence-from-theory-to-practice/>.

Mallett, R, Hagen-Zanker, J, Slater, R & Duvendack, M 2012, 'The benefits and challenges of using systematic reviews in international development research', *Journal of Development Effectiveness,* vol. 4, no. 3, pp. 445-55.

Nakashima, E 2013, 'To thwart hackers, firms salting their servers with fake data', *Washington Post,* 2 January, viewed 22 July 2018, <https://www.washingtonpost.com/world/national-security/to-thwart-hackers-firms-salting-their-servers-with-fake-data/2013/01/02/3ce00712-4afa-11e2-9a42-1ce6d0ed278_story.html? noredirect=on&utm_term=.ab586f749056>.

Prunckun, H 2018, 'Weaponization of computers', *Cyber weaponry: Issues and implications of digital arms,* H Prunckun, (ed.), Springer, Cham, CH.

Putnam, RT 2015, *Digital mirrors casting cyber shadows—The confluence of cyber technology, psychology, and counterintelligence,* Master of Science (capstone project), Utica College, New York, NY, US.

Sigholm, J & Bang, M 2013, 'Towards offensive cyber counterintelligence: Adopting a target-centric view on Advanced Persistent Threats', *Proceedings of the European Intelligence and Security Informatics Conference (EISIC), IEEE,* Uppsala, SE.

Stech FJ & Heckman KE 2018, 'Human nature and cyber weaponry: Use of denial and deception in Cyber Counterintelligence', *Cyber weaponry: Issues and implications of digital arms,* H Prunckun (ed.), Springer, Cham, CH.

Stone, GM & Bluitt, K 1993, 'Future law enforcement and internal security communications architecture employing advanced technologies', *IEEE Publication CH3372-0/93,* pp. 194 -202.
Stone, GM & Tucker RS 1988, 'Counterintelligence and unified technical security programs', *Proceedings of the IEEE International Carnahan Conference on Security Technology: Crime Countermeasures,* New York, NY, US.

United States of America 2004, Department of Defense, *Dictionary of military and associated*

terms (12 April 2011 as amended through 7 October 2004), viewed 7 January 2018, <http://www.bits.de/NRANEU/others/jp-doctrine/jp1_ 02%2804%29.pdf>.

United States of America (U.S.) 2018, *Counterintelligence: Cyber threat,* Naval War College, viewed 2 August 2018, <https://usnwc.libguides.com/c.php?g=661096&p=4695517>.

University of Johannesburg 2018, *The Cyber Counterintelligence Project—Centre for Cybersecurity,* viewed 16 April 2018, <http://adam.uj.ac.za/csi/ CyberCounterintelligence. html>.

von Solms, SH 2014, 'Cyber counter-intelligence makes a difference', *ITWeb,* viewed 11 November 2014, <http://www.itweb.co.za/index.php?option=com_content>.

Conducting Investigations on the Dark Web

C Easttom

Email: chuckeasttom@gmail.com

Abstract: *Recent years have seen the expansion of the dark web for use in criminal activity. The takedown of the Silk Road, and more recently Alphabay and Hansa, brought public attention to the dark side of the web, while law enforcement has the increased task of how to respond to criminal behaviour online. This paper provides an overview of the dark web, including a discussion of law enforcement and intelligence community techniques in investigating dark web markets.*

Keywords: *Dark Web, TOR, Dark Web Markets, Digital Investigations, Intelligence Community, Law Enforcement Investigation*

Introduction

While criminal activity on web pages has been an issue since the early days of the World Wide Web, the dark web is particularly prone to criminal activities. The dark web is an area of the Internet only accessible via onion routing (Ghappour 2017; McCoy, et al. 2008). Onion routing was patented by the United States Navy in 1998, with U.S. Patent 6,266,704. However, the U.S. Naval Research Laboratory released the code for onion routing under a free license. This allowed anyone who wished to duplicate the anonymous onion routing. The civilian onion routing network is now simply referred to as TOR (The Onion Router).

The current paper outlines the essentials of how dark web markets function, then describes a methodology for investigating such markets. The author's methodology has been developed over time and has now been taught to law enforcement, various government agencies, and civilian cyber-threat intelligence teams. Indeed, the methodology is a generalised process that can be utilised for any dark web market investigation.

Onion routing basics

Onion routing is a method of routing network traffic through a number of intermediary proxies (Goldschlag, Reed & Syverson 1999; Syverson, Goldschlag & Reed 1997). Each packet is encrypted with multiple layers of encryption, and each proxy can only decrypt one layer and send the packet to the next proxy (Reed, Syverson, & Goldschlag 1998). Should one intercept a packet in transit between two proxies, one can only determine the previous proxy and the next proxy (Camenisch & Lysyanskaya 2005). The actual origin or destination cannot be determined. **Figure 1,** below, depicts an overview of TOR communications.

This process of onion routing makes for a level of anonymity that is not readily available on traditional websites. While one can certainly utilise a fake identity on any website, the website may track the user's IP address, thus revealing who the user is. It is certainly possible to spoof an IP

address or to use a public Internet connection, but these measures only provide a small degree of anonymity. Onion routing makes the entire communication process anonymous.

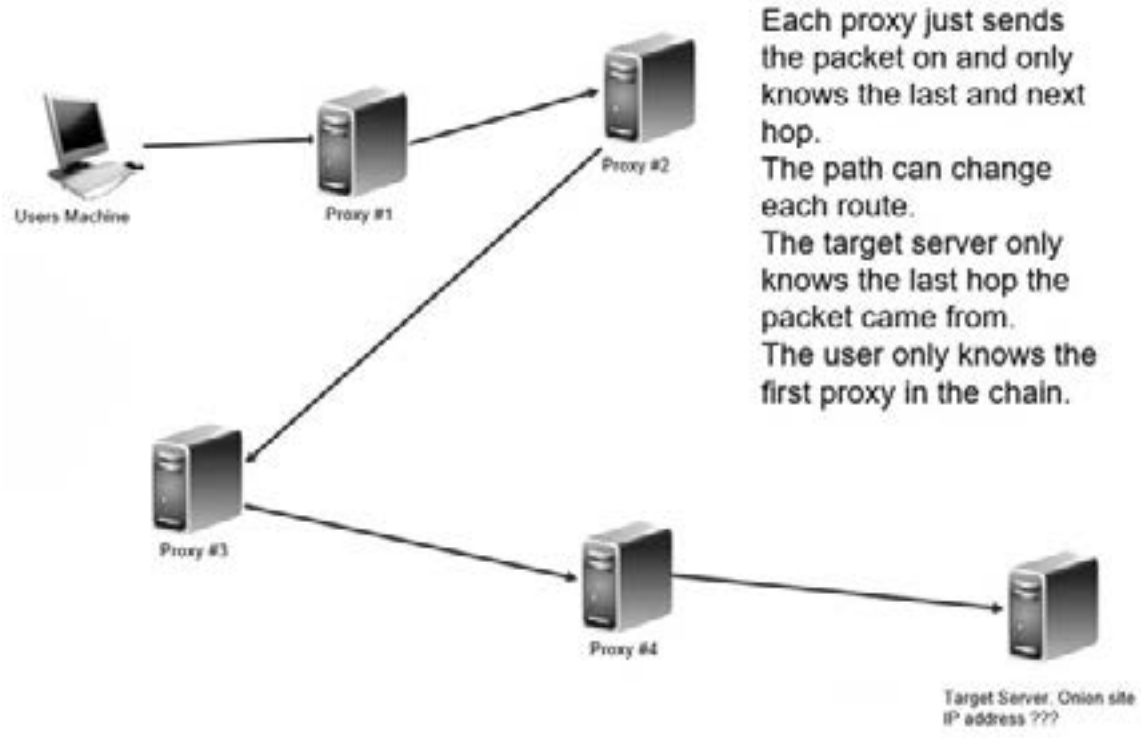


Figure 1: TOR

Websites accessible only via TOR make up what is commonly called the dark web. These sites usually end in an onion address. For example, the address http://kbhpodhnfxl3clb4.onion is a popular dark web search engine. These sites can only be accessed via the TOR browser, which is a free download from https://www.torproject.org/projects/torbrowser.html. The TOR browser is configured to utilise the TOR onion routing network for all communications.

Dark web market basics

The term dark web is used because these websites cannot be found via traditional means. Only via onion routing can one access a dark web site. The dark web should not be confused with the deep web, which is a separate topic and beyond the scope of this paper. For the purposes of this discussion, it is only necessary to differentiate between the two. The deep web consists of normal websites that have not been indexed by search engines. These include web pages that are automatically generated in response to queries; websites that are behind a login screen, or websites that intentionally avoid being indexed by search engines.

Since the dark web can only be accessed via onion routing, and the process of onion routing hides the origin of any packet, the dark web inherently affords greater anonymity than traditional web sites. This anonymity of the dark web has led to a number of markets on the dark web that traffic in illegal merchandise (Martin 2014; Buxton & Bingham 2015). The most well-known was Silk Road (Barratt, Ferris & Winstock 2014), followed by Silk Road 2 and Silk Road 3. Illicit drugs are a common commodity on the dark web (Kruithof, et al. 2010). However, other illegal items such

as forged identity documents, weapons, and child pornography are also for sale on the dark web (Zulkarnine, et al. 2016). An image of an actual dark web drug market is shown in **Figure 2.** This site is called The Peoples Drug Store (http://newpdsuslmzqazvr.onion).

Figure 2: Dark web drug market

This is only one example of a dark web market. Traversing dark web sites, however, one will find a range of content. Some dark web sites are simply places for people who wish to avoid the tracking that is so ubiquitous on normal websites. Those sites do not involve any illegal activity and are simply channels for anonymous communication. However, other sites involve a cornucopia of illicit activities.

While drug markets are quite common, they are not the only items for sale on the dark web. Stolen information, including personal information and financial information, is commonly for sale on the dark web. In recent years, monitoring corporate data and financial information being trafficked on the dark web has become a part of many cyber threat intelligence operations.

Some sites specialise in weapons including automatic rifles and grenades. Others traffic in hacking tools, tutorials, and malware source code. It is now possible to obtain ransomware as a service, via dark web markets. As shown in **Figure 3,** below, the dark web market http://satan6dll23napb5.onion offers a service named Satan that is ransomware as a servicer.

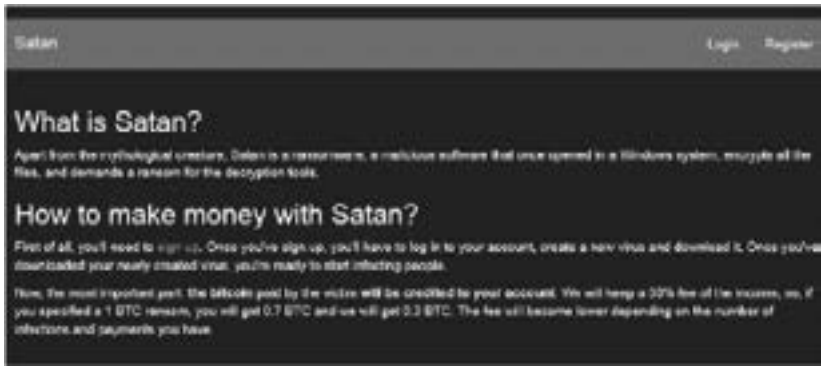


Figure 3: Ransomware as a service

Goods and services are paid for on dark web markets via crypto currencies, such as bitcoin. Bitcoin is a crypto currency that uses a blockchain to track coins. While one can investigate dark web markets without an understanding of blockchain or bitcoin, at least a cursory overview is helpful and is thus provided in the next section.

Bitcoin and blockchain

A blockchain uses a network of peers each running bitcoin software. Each node can validate transactions and add those transactions to its copy of the public ledger. The blockchain is a distributed database; each network node has a copy of the database. About six times an hour, a new group of transactions (a block) is created, added to the blockchain, and then published to all nodes.

Bitcoins also allow the interaction of the commerce on the dark web with traditional commerce. There are over 100,000 merchants who accept bitcoin as of this writing. There are even bitcoin ATM's where one can login to one's bitcoin wallet and exchange the bitcoins for traditional currency. Bitcoins allow someone selling goods or services on the dark web to translate profits into real-world currency.

Law enforcement and the dark web

The widespread presence of illicit and illegal goods and services on the dark web has made it a focal point for the interest of law enforcement. Law enforcement has had intermittent success with breaching these markets and shutting them down, along with arresting both perpetrators and customers. During early 2017 authorities were able to shut down AlphaBay and Hansa, two prominent markets for illicit goods and services (McGoogan 2017; Popper & Ruiz 2017). Both markets sold illicit drugs, financial data, stolen credit cards, and other illegal merchandise.

In one exemplary case, the FBI discovered a dark web market named Playpen. This market specialised in distributing child pornography. Through a flaw in the website, the FBI was able to breach the site (Cimpanu 2016). However, rather than simply and immediately arresting the administrator of the market, they took a different approach. The FBI hired someone to create a Metasploit exploit that would de-anonymize visitors to this dark web market. They then implanted that exploit into the web site. This ultimately led to hundreds of arrests and convictions of child pornographers (FBI 2017). This case combined a watering hole (the Playpen site itself) along with breaching that website and infecting it with a de-anonymizer.

There have been other mechanisms suggested for law enforcement to combat dark web markets, other than breaching its markets. In one study, the authors advocated law enforcement's using fake reviews of dark web drug markets to lower traffic (Markopoulos, Xefteris & Dellarocas 2015). This approach might have a limited effect in lowering the traffic to a given dark web market, but it does nothing to actually stop the illicit sales, and nothing to determine who the criminal actors are.

Past investigations of dark web markets have been sporadic and targeted at a specific market. There does not currently exist a template for conducting investigations on dark web markets. This paper outlines a template for such investigations, considering the increased caution that criminals on the dark web have implemented since the highly publicised breach of several dark web markets.

Conducting Dark Web Investigations

This paper discusses issues involved in investigating dark web markets, as well as specific techniques and an investigative approach developed by the author over a period of 24 months. This approach has been taught to various law enforcement agencies in workshops, and these techniques have also been utilised by private sector financial institutions in an attempt to discover if their data or financial accounts are for sale on the dark web.

To begin with, there are two aspects to dark web investigations, for both law enforcement and intelligence personnel. The first is simply understanding a given market. This involves generalized intelligence gathering and profiling of the market, as well as its customers. The second aspect is the potential breach of the market. This depends on similar techniques to what penetration testers and hackers utilize to breach websites.

Prior to performing any search or intelligence gathering on the dark web, one needs to be aware that the dark web is inherently dangerous. Many websites have spyware, viruses, and other malware. Therefore, one must establish a specific environment for dark web activities. That environment could be a Virtual Machine that is completely isolated from the host operating system. That means no sharing of the clip board or folders. The Virtual Machine is preferably of a different operating system than the host, making crossing the VM barrier even more difficult. And finally, that VM should only be used for the dark web activities, and for no other purpose. That way, if it should become infected with spyware, the person controlling the spyware will not gain information about the investigator's real identity. There is a virtual machine created for just this purpose. The tails virtual machine is made for privately surfing the dark web. That virtual machine does not store any data when it is shut down. It can be downloaded for free from https://tails.boum.org/install/.

Profiling dark web markets

The first step in dark web investigations is to understand specific dark web markets. The specific markets of interest will vary depending on the investigation. However, all investigations will begin with profiling the market(s) of interest. Profiling a dark web market begins by visiting the market and carefully reviewing the goods and services for sale. This can be time consuming. It is also important to be aware that there are a number of scams perpetrated on the dark web. Individuals may claim to be selling some illegal goods or services, but once payment is made, they refuse to produce the goods or services. This is quite common with markets that purport to offer violent services such as murder for hire. However, drug and child pornography markets are, usually, actually offering the illegal goods they claim to.

To avoid wasting investigation time on a market that may not even have actual goods or services, it is important to frequent dark web forums and bulletin boards and find out what other dark web users think about a given site. Later in this paper, there is a list of dark web forums as well as starting points. These provide a place to begin mapping out what markets are currently active, and which markets are of interest in a specific investigation. As has already been discussed in this paper, markets often specialise in a particular type of goods or services. Some, such as the Peoples Drug Store, are known for illicit drug sales. Other dark web markets are known for dealing in financial data, such as the Wallstreet market.

Locating new markets is best accomplished by visiting forums and following the discussions in those forums. In that way, the investigator can see what the various participants in the dark web think of a given market. The reputation of dark web markets is critical. Once the investigator has profiled a given dark web market and determined that it is indeed selling the goods and services it purports to, the next stage in the investigation must commence. For law enforcement, that means actually purchasing the goods. For civilian investigators, this should be avoided in most circumstances, as the purchase is itself a crime. The one exception for civilians involves data belonging to the investigator's company or a client who hired the investigator. In that case, the investigator my need to purchase a sample of data in order to determine if the company's information is actually for sale on that market.

One significant problem with investigating dark web markets is that many of the people who use these markets have become more cautious in the wake of prominent markets' being shut down by law enforcement. There have been three specific changes in dark web market behaviour that inhibit investigation.

The first major change has been that many of these markets are no longer open to the general public. In order to gain access to the market, one must first contact the administrator of the market, using a dark web email address, then request access. The administrator will take steps to verify the requestors bona fides before granting access. Generally, this involves some methods for verifying that the identity requesting access is known in the dark web community.

The second change has been the widespread use of PGP encryption (Broséus, et al. 2016). Pretty Good Privacy (PGP) is a widely used encryption tool for Internet communications (Zimmermann 1995). While PGP has been available longer than the dark web and is widely used in normal Internet traffic, its use on the dark web has increased significantly after the publicised take-downs of prominent dark web markets. Many dark web sellers and buyers will only communicate via TOR-based email services and insist on encrypting their emails using PGP or a similar encryption service. This is done in case law enforcement has compromised a given TOR node, which would allow communications to be monitored. By encrypting all communications, the dark web merchants protect against such breaches of the TOR network.

The third major change affects tracing crypto currency. Most purchases on the dark web are done via bitcoins, which are very difficult to trace. Many users of dark web markets are now further obfuscating their source and destination of bitcoins by utilising tumbler services. A tumbler service is used to pool together bitcoins from a number of sources. Then each participant takes out a number of bitcoins equal to what he or she put in (minus a service charge). In this manner, the bitcoins one has, are unlikely to be that person's original bitcoins. This makes tracing crypto currency even more difficult than it already was.

As has already been discussed in this paper, there have been successful investigations of dark web markets, and some have even been completely breached by law enforcement and taken offline. What is lacking is a comprehensive strategy that any law enforcement, intelligence agency, or civilian investigator can utilise to investigate, and perhaps breach, dark web markets. Based on approximately 24 months of successful dark web intelligence gathering, this paper proposes a specific methodology for investigating dark web markets.

Establishing a dark web identity

In order to fully investigate any dark web market, it will be necessary to establish a dark web identity. That identity will be utilised to interact with the dark web market and individuals on forums. Given the changes in the behaviour of dark web participants, the only way to successfully interact with many markets, is to establish an identity on the dark web. This is a process that will span many weeks or months. The first step is to establish an identity on the dark web. At a minimum this identity should have a TOR based email address and interact with other dark web users on a variety of forums and bulletin boards. This interaction must span at least many weeks.

The point of the online interaction via forums is to establish the false identity as a trusted user of the dark web. Once the investigator has obtained a TOR email address, he or she next signs up for a number of dark web forums. At first, the investigator should simply read forum postings and get an understanding of the type and style of communications. A number of TOR email sites and forums are listed later in this paper. However, be aware that sites can come and go on the dark web with great frequency. It may be necessary to use a search engine or link aggregation site and find a current TOR email service as well as currently active forums. Several dark web starting point sites are enumerated later in this paper.

The second step in establishing an identity is to buy or sell items on the dark web. These can be low level items such as data dumps. But the purchasing and or selling of items on the dark web helps to establish the identity as a legitimate member of the dark web community. At some point, it may even be advantageous to establish two identities. The first purports to have some illegal item for sale, the second identity purports to purchase that item. Then the second identity can leave positive reviews for the first identity. At no point do any actual illicit goods or services change hands, but the perception of an illegal transaction is established.

Another effective way to enhance a dark web identity is to establish a dark web site. This need not be a market. It can be a forum, or even simply a link aggregation site and perhaps a blog. As the identity has a robust link aggregation site, it will get more traffic. As that identity's blog entries are consistent with the views of many dark web users, that identity will become more credible to users of the dark web.

This process must be spread out over many weeks, perhaps months. It is recommended that any law enforcement or intelligence agency that is interested in dark web investigations begin immediately to establish one or more identities. These will need to be in place, and credible, long before they can be used in any investigation.

By establishing a well-respected identity on the dark web, it is then relatively easy to monitor markets and users, or even to purchase from markets. Once the identity is firmly established, one can navigate the dark web with relative ease.

Investigating specific dark web markets

Monitoring dark web markets with a well-crafted dark web identity can allow investigators to identify items for sale on the dark web. Civilian investigators have had success identifying their own company's data for sale on dark web markets prior to that information's being used. This al-

lows the company whose information is being trafficked to take appropriate security measures to mitigate the damage of the information that has been leaked, particularly in the case of login and financial account information. Simply being aware that the information is for sale on the dark web allows the victim company to immediately take steps such as changing passwords, monitoring accounts, and initiating other remediation measures.

Another possible goal for law enforcement and intelligence agencies is to breach a dark web market. This obviously is not an option for civilian investigators but can be a viable option for government investigators. This will require traditional hacking techniques, as were used in the Playpen case. As with traditional hacking, this process begins with finding vulnerabilities in the dark web market site. Fortunately, tools are emerging that assist in this endeavour. The website ichidan http://ichidanv34wrx7m7.onion tracks vulnerabilities in dark web sites. This site operates similarly to the traditional Internet site shodan.io, which is a search engine for vulnerabilities in normal websites.

In the case of the aforementioned Playpen market, it was a network de-anonymizer that was used to infect the dark web market. The specific malware was named Network Investigative Technique and was used to deanonymize suspects using TOR: "The NIT was a Flash based application that was developed by H.D. Moore and was released as part of Metasploit. The NIT, or more formally, Metaspolit Decloaking Engine was designed to provide the real IP address of web users, regardless of proxy settings" (Paganini 2015). NIT was used in the Playpen case.

Using this approach, the FBI was able to gather information such as: IP address through the TCP connection, operating system, CPU architecture and session identification. This allowed the FBI to completely identify the visitors to the dark web market. Ultimately, this led to the arrest of many people who traffic in child pornography.

As was shown in the FBI Playpen case, Dark Web sites are just as prone to technologies issues as any other website. These sites might have flaws in the web server, web programming, encryption implementation, or any other technical issue that could affect any website.

When considering breaching dark web sites, it is important to consider flaws in the TOR browser itself. It is a modified Firefox browser. As such, the browser is as susceptible to vulnerabilities as any other browser. Several vulnerabilities in the TOR browser have been documented (Nurmi & Niemelä 2017; Schneier 2013). These vulnerabilities can provide a mechanism for beaching the anonymity of dark web traffic.

Starting points for investigations

While some prominent markets have been taken down by law enforcement, others remain in operation. As of this writing, some of the markets with the most activity include those shown in **Table 1,** below.

Wall Street	http://wallstyizjhkrvmj.onion/
Silk Road 3	http://silkroad7rn2puhj.onion/
Valhalla	http://silkkitiehdg5mug.onion
Grams	http://grams7enufi7jmdl.onion
Pushing Taboo	http://pushingtabu7itqj.onion/
Trade Route	http://traderouteilbgzt.onion
T Chka Market	pointgg344ghbo2s.onion

Table 1: Dark web markets

There are several specific dark web sites that are useful starting points. For law enforcement or intelligence officers wishing to establish dark web identities, there are some starting points provided in this section. First, dark web email services are listed in **Table 2.**

Anonymous E-mail service	http://365u4txyqfy72nul.onion/
TorBox - The Tor Mail Box	http://torbox3uiot6wchz.onion/
NoteBin - Create encrypted self-destructing notes	http://notestjxctkwbk6z.onion/

Table 2: Dark web email services

In addition to those starting points, there are few dark web forums which bear review by anyone investigating activity on the dark web. A list of these forums is given in **Table 3**.

It will also be useful for any investigator to have search engines and starting points from which to begin a dark web investigation. **Table 4** provides such starting points.

Onion Forum 2.0 renewed	http://2gxxzwnj52jutais.onion/phpbb/index.php
Twitter clone	http://npdaaf3s3f2xrmlo.onion/
Social network: File sharing, messaging and much more	http://hbjw7wjeoltskhol.onion
Darknet Avengers	http://avengerfxkkmt2a6.onion/
The hub, a forum	http://thehub7gqe43miyc.onion/

Table 3: Dark web forums

Torch, a search engine	
	http://xmh57jrzrnw6insl.onion/
A search engine	https://hss3uro2hsxfogfq.onion.to
OnionDir a directory of dark web sites	
	http://dirnxxdraygbifgc.onion
Tor Search, a search engine	http://kbhpodhnfxl3clb4.onion
Tor Find, a search engine	http://ndj6p3asftxboa7j.onion
Basic link lists	linkzbg4nwodgic.onion
Dark Web Links	jdpskjmg5kk4urv.onion

Table 4: Dark web starting points

Conclusions

Despite law enforcement's recent successes in taking down some dark web markets, there is still a thriving trade in illegal goods and services on the dark web. The fact that dark web markets are still sources for a wide range of criminal goods and services makes the investigation of such markets an important task for law enforcement, the intelligence community, and civilian cyber threat intelligence teams. The anonymity of TOR, combined with the additional security measures many participants in dark web markets have taken, presents a challenge to the law enforcement and intelligence communities. In order to facilitate such investigations, a generalised methodology is needed. The specific techniques outlined in this paper can be used to effectively investigate dark web markets.

By establishing a carefully crafted dark web identity, investigators can begin successfully monitoring and investigating dark web markets. For some agencies, breaching the target market may be an additional step. As outlined above, it is possible to breach dark web markets in ways similar to how a normal website would be breached. Whether or not that is operationally feasible will depend on the type of investigation and the agency performing the investigation.

Further research is indicated in several areas. The first area of research would include specific case studies of either individual markets or general classes of markets (i.e. drug markets, terrorist dark web sites, etc.). The second area of research would be the application of traditional hacking techniques to dark web markets. Again, it must be stressed that whether or not hacking into a given site (either a traditional web site or a dark web market) is an acceptable operational goal is dependent upon the agency conducting the investigation and the specific nature of the investigation.

References

Barratt, MJ, Ferris, JA & Winstock, AR 2014, 'Use of Silk Road, the online drug marketplace, in the United Kingdom, Australia and the United States', *Addiction,* vol. 109, no. 5, pp. 774-83.

Broséus, J, Rhumorbarbe, D, Mireault, C, Ouellette, V, Crispino, F & Décary-Hétu, D 2016, 'Studying illicit drug trafficking on Darknet markets: structure and organization from a Canadian perspective', *Forensic Science International,* vol. 264, pp. 7-14.

Buxton, J & Bingham, T 2015, 'The rise and challenge of dark net drug markets', Policy Brief 7, Global Drug Policy Observatory, Swansea University, Swansea, Wales, UK.

Camenisch, J & Lysyanskaya, A 2005, 'A formal treatment of onion routing', *Proceedings of the 25th annual international Cryptology Conference—CRYPTO 2005,* 14-18 August, Santa Barbara, CA, US, pp. 169-87.

Cimpanu, C 2016, 'Admin of dark web child pornography website "Playpen" found guilty', *Softpedia News,* viewed 20 March 20 2018, <http://news.softpedia.com/news/admin-of-dark-web-child-pornography-website-playpen-found-guilty-508405.shtml>.

FBI 2017, '"Playpen" creator sentenced to 30 years', FBI.gov, 5 May, viewed 9 January 2019, <https://www.fbi.gov/news/stories/playpen-creator-sentenced-to-30-years>.

Ghappour, A 2017, 'Searching places unknown: Law enforcement jurisdiction on the dark web', 5 March 2016, *Stanford Law Review,* vol. 69, no. 1075.

Goldschlag, D, Reed, M & Syverson, P 1999, 'Onion routing,' *Communications of the ACM,* vol. 42, no. 2, pp. 39-41.

Kruithof, K, Aldridge, J, Hétu, DD, Sim, M, Dujso, E & Hoorens, S 2016, 'The role of the dark web in the trade of illicit drugs', Brief, WODC, Ministerie van Veiligheid en Justitie, viewed 12 April 2018, <https://www.rand.org/pubs/research_briefs/RB9925.html>.

Markopoulos, P, Xefteris, D, Dellarocas, C 2015, 'Manipulating reviews in dark net markets to reduce crime', viewed 1 May 2018, <http://www.fox.temple.edu/conferences/cist/papers/Se,s-son%201A/CIST_2015_1A_2.pdf>.

Martin, J 2014, *Drugs on the dark net: How cryptomarkets are transforming the global trade in illicit drugs,* Springer Press, New York, NY, US.

McGoogan, C 2017, 'AlphaBay: World's largest dark web site is shut down', *The Telegraph,* viewed 2 May 2018, <http://www.telegraph.co.uk/technology/2017/07/20/alphabay-us-government-shuts-worlds-largest-dark-web-market/>.

McCoy, D, Bauer, K, Grunwald, D, Kohno, T & Sicker, D 2008, 'Shining light in dark places: Understanding the TOR network', *Proceedings of the 8th international symposium on Privacy Enhancing Technologies (PET 2008): Privacy Enhancing Technologies,* Leuven, BE, 23-25 July, pp. 63-76.

Nurmi, J & Niemelä, MS 2017, 'TOR De-anonymization techniques', *Proceedings of the 11th international conference on Network and System Security—NSS 2017,* Helsinki, FI, 21-23 August, pp. 657-71.

Paganini, P 2015, 'A look to the "NIT Forensic and Reverse Engineering Report, continued from January 2015": NIT code was used by the FBI to deanonymize TOR users', *Security Affairs,* viewed 2 May 2018, <http://securityaffairs.co/wordpress/38213/cyber-crime/nit-fbi-deanonymize-tor.html>.

Popper, N & Ruiz, R 2017, '2 Leading online black markets are shut down by authorities' *New York Times.* 20 July, viewed 2 May 2018 <https://www.nytimes.com/2017/07/20/business/dealbook/alphabay-dark-web-opioids.html>.

Reed, MG, Syverson, PF & Goldschlag, D M 1998, 'Anonymous connections and onion routing', *IEEE Journal on Selected areas in Communications,* vol. 16, no. 4, pp. 482-94.

Schneier, B 2013, 'Attacking Tor: How the NSA targets users' online anonymity', *The Guardian,* vol. 4, 7 October.

Syverson, PF, Goldschlag, DM & Reed, MG 1997, 'Anonymous connections and onion routing', *Proceedings of the 1997 IEEE Symposium on Security and Privacy,* pp. 44-54.

Zimmermann, PR 1995, *The official PGP user's guide,* MIT press, Cambridge, MA, US.

Zulkarnine, AT, Frank, R, Monk, B, Mitchell, J & Davies, G 2016, 'Surfacing collaborated networks in dark web to find illicit and criminal content', *Proceedings of the 2016 IEEE conference on Intelligence and Security Informatics—ISI 2016,* 28-30 September, Tucson, AZ, US, pp. 109-114.

On the Use of Ontology Data for Protecting Critical Infrastructures

J Henriques[1,2], F Caldeira[1,2,] T Cruz[1], P Simões[1]

[1]Department of Informatics Engineering
University of Coimbra
Coimbra, Portugal

Email: jpmh@dei.uc.pt; fmanuel@dei.uc.pt; tjcruz@dei.uc.pt; psimoes@dei.uc.pt

[2]Polytechnic Institute of Viseu
Viseu District, Portugal

Abstract: *Modern societies increasingly depend on products and services provided by Critical Infrastructures (CI). The Security Information and Event Management (SIEM) systems in charge of protecting these CIs usually collect and process data from specialised sources. However, they usually integrate only a small fraction of the whole data sources existing in the CI. Valuable generic data sources are missing in this process, such as human resources databases, staff check clocks, and outsourced service providers. To address this gap, the authors propose a framework that takes a Semantic Web approach for automated collection and processing of corporate data from multiple heterogeneous sources.*

Keywords: *Critical Infrastructure Protection (CIP), Security Information and Event Management (SIEM), Industrial Automation and Control Systems (IACS), Semantic Web, Ontologies*

Introduction

Critical Infrastructures (CI) such as telecommunication networks and power grids are becoming increasingly complex and interdependent on people, processes, technologies, information, and other critical infrastructures. Operators in charge of Critical Infrastructure Protection (CIP) are required to improve their security levels through the perspective of compliance auditing and forensic analysis. Compliance auditing is related to applicable security regulations, standards, and best practices. Forensic analysis has a broader scope, beyond the specific operations of the CI industrial control systems, and also encompasses other areas of the organisation.

The benefits of enlarging the scope of information sources for SIEM applications, forensic analysis, and compliance audit operations are rather evident, since the result would enable more powerful, all-inclusive approaches to cybersecurity awareness. For example, monitoring of abnormal activity within the IACS specific domain might be leveraged by the correlation of different data sources, such as mail filtering logs (monitoring phishing and malware attacks, which target the employees of the CI) and information about employee functions residing in Human Resources information systems. Another example would be the correlation of data from physical access control systems and staff check clocks with activity logs of IACS operators. In general, this strategy

of associating core security information already fed into SIEM systems with peripheral-awareness data would result in richer security analysis processes that enable the detection of inconsistencies, malpractices, and intrusion clues, which would otherwise go unnoticed.

However, achieving tight integration of all those peripheral data sources into the already-existing SIEM frameworks is costly and often impractical. This would require considerable investments in data conversion and adaptation to the SIEM data flows. Moreover, the maintenance costs would also be considerable, since even minor adjustments on the corporate information systems would require explicit adaptations on the SIEM side.

A more plausible option is, therefore, the adoption of loosely coupled integration strategies, such as resorting to Semantic Web approaches for automating the processing and interpretation of large amounts of information available from both local databases and Internet repositories. This reasoning process, applied over a large quantity of available data with knowledge inferred from a combination of axioms, properties, and rules (with different levels of hierarchies or categorisations and deriving conclusions, for example) can be explicitly expressed by ontologies.

It should be noted that most data are still not directly available in Semantic Web formats. This is the case with data maintained in Relational Databases (RDBs). Nonetheless, mapping data from RDB to Semantic-Web-enabled Resource Description Frameworks (RDFs) has been the focus of a large body of previous research, leading to the implementation of many generic mapping tools and their applications, on several specific domains. Those tools are natural candidates to be adapted to the field of CIP so that security-related ontology data currently stored in heterogeneous databases can be taken into consideration—despite the considerable challenges involved, such as the migration from existent systems to the semantic level (Sernadela, González-Castro & Oliveira 2017).

A detailed discussion of the main motivations and driving research efforts in mapping RDB to RDF can be found in Sahoo et al. (2009). Although most models can perform inference from native ontology data stores, data still reside mostly in RDBs, which are broadly used within organisations. Moreover, the growing number of datasets published on the Web brings opportunities for extensive data availability and challenges related to the process of querying data in a semantically heterogeneous and distributed environment. The structured query approach fails on the linked data because the Web's scale makes it impractical for users to know in advance the structure of datasets (Freitas, et al. 2012).

The authors have previously introduced an approach considering inference capabilities from Semantic Web, supported by common schemas, for creating a set of independent databases, each deployed with its own domain-specific schema (Henriques et al. 2018). This kind of reasoning is suitable for application in the context of Critical Infrastructure Protection; and, therefore, it can leverage current SIEM capabilities—mainly in what relates to forensic and compliance audit processes, but also for intrusion detection purposes. This large amount of living heterogeneous data that still resides in the organisational RDBs will, in this way, become available to the Critical Infrastructure's SIEM and enable new, valuable insights into available configuration and monitoring data.

This paper refines and extends previous work (Henriques et al. 2018) by providing a more detailed description of the proposed approach, adding a practical use-case scenario, and discussing its future application to different data sources.

After discussing some of the key previous work and trends in the area, this paper takes a practical approach by presenting the implementation of a federated query architecture for retrieving a set of compliance auditing rules that might be useful, for instance, in assessing CI security levels. To leverage inference capabilities, it maps the living data currently available on RDBs into RDFs formats. In this way, it can substantially enlarge the data available to the SIEM by taking advantage of the large amount of heterogeneous data of production-RDB systems. Such an approach provides an abstraction mechanism for keeping data consumers away from low-level details while leveraging the security concerns of the underlying infrastructures by hiding the internal deployment aspects, such as the identification of the involved machines and their RDB schemas.

The ontology-based approach of this work considers the available information currently stored in RDB and, as its main goal, makes it accessible through simple interfaces that collect queried data from multiple natively different data repositories within the organisation. Each available RDB maintains different information instances, deployed on specific schemas and technologies. Such an approach is suitable for combining data from two different worlds, such as the case of RDB and Semantic Web data, which is natively maintained in RDF stores and made available through an interface layer encapsulating the details of the gathering process to retrieve the data from multiple RDBs.

The remainder of this paper is structured as follows. The next section discusses the background for the domain problem and related work. Immediately following is an analysis of the applicability of ontology data in the context of CIP. Next, a description of the proposed architecture, which details its implementation, will be provided. Finally, the authors conclude the paper with insights about future developments.

Background

This section briefly introduces the reader to the key concepts and tools used in the proposed data integration approach: RDF; RDB, and RDF mapping; SPARQL; Direct Mapping of Relational Data to RDF; and the D2RQ platform.

Resource Description Framework (RDF)

An ontology is a formal specification of concepts (Gruber 1993) in a domain of discourse, which includes classes and properties. An ontology, together with a set of individual instances of classes, constitutes a knowledge base (Noy & McGuinness 2001).

The Resource Description Framework (RDF) (Brickley & Guha 1999) is a language that can be used to encode knowledge into web pages to make them understandable for electronic agents searching for information. This is one of the main goals for using ontologies (Musen 1992; Gruber 1993). RDF aims at representing information that may be used for inference purposes over the Web. The RDF syntax core structure consists of a set of triples with a subject, a predicate, and an object. A set of triples is called an RDF graph. An RDF graph may be visualised as a directed-arc

diagram, in which each triple is represented as a node-arc-node link. RDF is a data format based on a Web-scalable architecture for identification and interpretation of terms (RDF 2014).

Mapping from RDF to RDB

As already mentioned, the mapping of large amounts of data from RDB to RDF has been the focus of intense research work in multiple domains and has led to the implementation of a set of generic mapping tools, as well as domain specific applications. RDF has provided an integration platform for data gathered from multiple sources, primarily from RDB. This is one of the main motivations driving research efforts (using various approaches) on mapping RDB to RDF (Seaborne, et al. 2013).

SPARQL (W3 2013) can be used to express queries across diverse data sources, whether for data natively stored as RDF or for data viewed as RDF via some sort of middleware. SPARQL is a World Wide Web Consortium (W3C) recommendation for querying multiple RDF graphs. The SPARQL specifications define the syntax and semantics to proceed with queries across diverse natively stored RDF data sources. Using the latest stable release (1.1), SPARQL federated queries allow merging multiple results retrieved from multiple RDF sources. The syntax and semantics of SPARQL 1.1 Federated Query extension allow distributed queries over different SPARQL endpoints. Moreover, the SERVICE clause extends SPARQL 1.1 to support queries that merge data distributed across the Web. A single query is, therefore, able to return related data (for example, contacts to be applied to user John Doe) from multiple distinct SPARQL endpoints.

An important feature of RDF and SPARQL is that they can use different datasets from different locations, federating them together. They offer a middleware, which can use multiple data sources as if they were one. Moreover, it is simple to add and remove data sources. This feature significantly reduces the development costs as compared to typical data warehouse projects (DuCharme 2013).

Figure 1 provides a query example through different SPARQL 1.1 endpoints. The query returns John's contacts from two distinct SPARQL endpoints, www.site1.com and www.site2.com.

```
SELECT ?contact1
WHERE {
SERVICE <http://www.site1.com/sparql>
 {SELECT ?contact1
WHERE {
        ?me foaf:nick “John”.
        ?me foaf:knows ?f .
       ?f foaf:name ?contact1
       }
}
SERVICE <http://www. site2.com/sparql>
 {
SELECT ?contact2
WHERE {
      ?me foaf:nick “ John “.
       ?me foaf:knows ?f .
       ?f foaf:name ?contact2 }}
        FILTER (?contact1 = ?contact2)
      }
}
}
```

Figure 1: Query example through different SPARQL 1.1 endpoints

Direct mapping of relational data to RDF

Relational databases allow the use of tools, such as Structured Query Language (SQL), for accessing and managing the databases. Several strategies already exist to map relational data to RDF. Typically, the goal is to describe the RDB contents using an RDF graph, allowing queries submitted to the RDF schema to indirectly retrieve the data stored in relational databases. A direct mapping process enables a simple transformation and can be used for materialising RDF graphs or for defining virtual graphs, which can be queried via SPARQL or traversed by an RDF graph Application Programming Interface (API). A mapping document is an RDF document containing triples maps with instructions on how to convert relational database content into RDF graphs.

The D2RQ platform

The D2RQ (Data to RDF Query) Platform allows users to access relational databases as virtual, read-only RDF graphs while automatically producing the corresponding mappings. It is available under the Apache open source license (D2RQ 2012), and it allows users to create customised mappings from RDB through an integrated environment with multiple options for accessing relational data, including RDF dumps, Jena and Sesame API based access, and SPARQL endpoints on D2RQ Server (Bizer & Cyganiak 2007). It offers RDF-based access to the content of RDB, without requiring its replication into RDF stores. D2RQ, therefore, allows querying non-RDF databases using SPARQL or accessing contents of databases over the Web. It also allows the creation of custom content dumps from relational databases into RDF stores.

The D2RQ Platform includes components such as a Mapping Language, an Engine, and a D2R (Data to RDF) Server. The D2RQ Engine is a plug-in for the Jena Semantic Web toolkit, which uses mappings for rewriting the Jena API calls to SQL queries against the database and for redirecting query results up to the higher layers of the framework. The D2R Server is an HTTP server which provides linked data views, HTML views for debugging, and a SPARQL protocol endpoint providing an interface to query the database. The D2RQ platform supports databases such as MySQL, SQL Server, Oracle, PostgreSQL, HSQLDB, and Interbase/Firebird. Some limitations of D2RQ include the integration of multiple databases or other data sources and its read-only nature: it lacks Create, Read, Update, and Delete (CRUD) operations. Finally, it does not support inference mechanisms and does not include named graphs (D2RQ 2012).

The D2RQ Mapping Language enables defining relationships between RDB schemas and RDF schema vocabularies (classes and properties) or Web Ontology Language (OWL) ontologies written in Turtle syntax (W3 2014). The mapping properties define a virtual RDF graph, which contains information from the database schema. The mapping process between D2RQ and RDB entities includes the RDF class node to RDB tables and RDF predicates to RDB column names (D2RQ 2012).

The same D2RQ server can be configured to access multiple databases. Therefore, a single SPARQL query can request data from multiple databases at once, which is not possible with a standard SQL query.

Applicability of Ontology Data in the Context of Critical Infrastructure Protection

This section addresses the applicability of ontology data in the context of CIP. First, some of the

more pertinent related works are discussed. Afterwards, the H2020 ATENA module for forensics and compliance auditing is presented. This module provides the framework on which the proposed approach, described in the following section, was developed.

Related work

Current approaches on the use of ontologies in the context of CIP are mostly related to the assessment of interdependencies between Critical Infrastructures, such as the works of Castorini et al. (2010) and Blackwell et al. (2008). Similarly, a proposal for an ontology providing vulnerabilities classification to be used in decision support tools can be found in Chorás et al. (2010).

Other approaches worth mentioning include SPLENDID, DARQ, SemaPlorer, and FedX. SPLENDID (Gorlitz & Staab 2011) is a query optimisation strategy for federating SPARQL endpoints based on statistical data. DARQ (Quilitz 2008) provides transparent query access to multiple SPARQL services using one single RDF graph, even when data has a distributed nature and is spread over the Web. This approach includes a service description language that enables a query engine to decompose a query into subqueries, where each of them can be answered by an individual service. SemaPlorer (Schenk et al. 2009) also provides a federated query architecture allowing it to interactively explore and visualise semantically heterogeneous distributed semantic datasets in real time, through a conceptual layer on top of Amazon's Elastic Computing Cloud (EC2). FedX (Schwarte, et al. 2011) proposes novel joint processing and grouping techniques for minimising the number of remote requests. It also develops a practical framework that enables efficient SPARQL queries supported by federation layers for efficient query processing on heterogeneous distributed Linked Open Data sources.

Beyond D2RQ, other RDF middleware applications exist, such as TopQuadrant's TopBraid Live, OpenLink Software's Virtuoso Sponger, and Triplr project. These offer dynamic creation and integration. They also allow users to merge several RDF triples in a single SPARQL endpoint from sources such as relational databases, spreadsheets, HTML documents, and other formats.

As already mentioned, one possible application of ontology data in this scope is the use of heterogeneous sources available in organisational RDBs for leveraging inference capabilities. This application is especially interesting in the specific areas of forensic analysis and compliance audit processes, which, by nature, need to be supported by substantial amounts of heterogeneous data. A possible practical application of this approach, in the scope of forensic analysis and compliance audit processes, may consist of the collection and mapping to Semantic Web of rules residing in the multiple and heterogeneous relational databases of the CI organisation—so they can be combined with the knowledge already available at the SIEM systems. This path has been explored in the scope of the H2020 ATENA research project (ATENA 2018; Rosa et al. 2017), as discussed next.

Forensics and compliance auditing in the scope of the H2020 ATENA framework

The H2020 ATENA project proposes an innovative logical framework, with design improvements of role, operation, architecture, and security components for Industrial Automation and Control

Systems (IACS), while also exploiting novel security approaches enabled by network virtualisation paradigms. The Forensics and Compliance Auditing (FCA) module, integrated into the ATENA cyber-security architecture, addresses the gathering and persistent storage of digital evidence retrieved from both the cyber-analysis layer (such as SIEM) and peripheral sources (such as service logs, sessions, or physical access control systems, among others) for forensics and compliance auditing purposes. Its forensics tools provide the means to identify, extract, preserve, and highlight digital evidence for technical investigation and legal purposes. Its compliance auditing tools support the audit procedures associated with certification processes for applicable standards, policies, and regulations—for example, verifying the authorisation procedures for physical installation access, such as access to doors (Rosa et al. 2017).
Moreover, the FCA module provides a set of analysis capabilities for interactively exploring, searching, extracting, pinpointing, and combining insights from available data. The core FCA functions encompass collecting heterogeneous data from internal and external sources, producing structured and unstructured data to be combined and gathered into a unified view for compliance auditing—throughout a set of rules—and also providing forensic investigation functionalities for retrieving evidence.

Figure 2, below, depicts the main blocks of the ATENA FCA module. Data collected from the ATENA SIEM and intrusion detection systems feed a specific CI security data lake which provides input to the FCA analytics components. Peripheral data sources, processed through domain-specific business rules, also feed the analytics layer. Trust and repudiation indicators are also used to assess the trustworthiness of each data source.

As previously discussed, specific ontologies need to be constructed for supporting the already mentioned processes of compliance audit and forensics analysis. In the context of the FCA module hereby presented, the targets for the use of those ontologies are the Analytics sub-components 'Audit Compliance' and 'Forensic Analysis'.

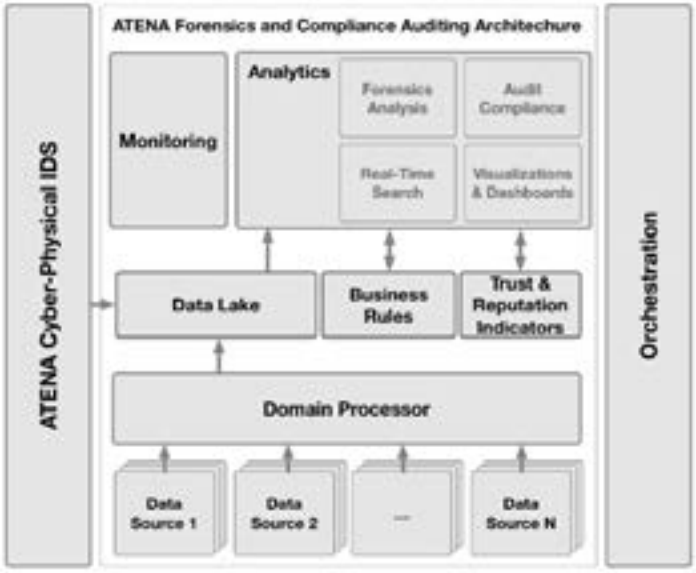


Figure 2: The Forensics and Compliance Auditing Module of the ATENA Project, adapted from Rosa, et al. (2017)

Proposed Approach to the Use of Ontology Data for CIP

This section describes the proposed approach to the use of ontology data in the context of CIP applications. First, the proposed reference architecture is introduced, followed by a discussion of technical aspects and implementation details. In a simplified view, the proposed solution consists of a web service that can receive several SPARQL requests from data consumers (such as the forensics and compliance auditing tools mentioned in the previous sections). Afterwards, each one of those requests is forwarded into different databases deployed using different schemas.

Reference architecture

The proposed reference architecture, depicted in **Figure 3,** below, consists of a set of components such as a federated layer, mapping brokers, and databases. Several data consumers (clients) may send distinct sets of SPARQL queries to the federated interface layer, which delivers each query to all the brokers. The broker's main role is to transform the incoming SPARQL queries into native relational database queries. Through an inverse flow, the broker retrieves the data subset from the database to be gathered into a full data set at the federated layer which is then forwarded to the involved client(s).

Although the reference architecture may suggest its applicability to the context of federated database queries, it may be extended to use different kind of data sources, such as logs or Lightweight Directory Access Protocol (LDAP) distributed directory information services (among others) in order to provide compliance audit and forensic capabilities that can be applied to the context of ATENA FCA module.

Use-case scenario

Next, a simple compliance audit scenario is presented, which demonstrates the applicability of the reference architecture for evaluating unauthorised accesses to the assets of an international company.

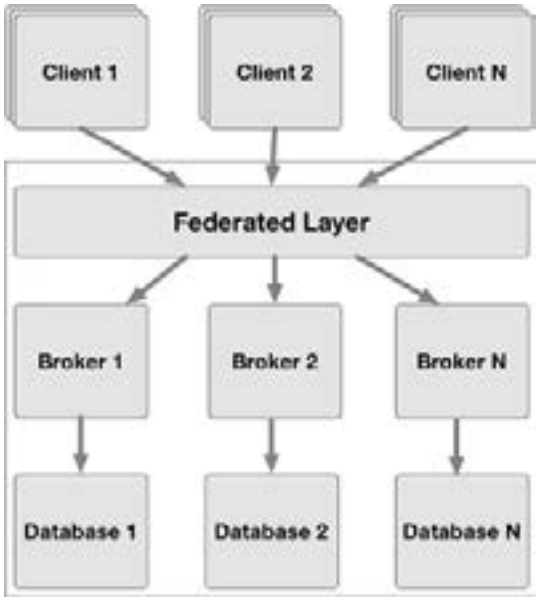


Figure 3: Proposed reference architecture

The challenge is to build a common schema for the management of human and asset resources spread over different platforms, because of specific requirements imposed by national governments. A single interface, capable of answering queries merging all the data in the organisation in a single dataset, should be provided. Such an approach would help overcome the barriers by approaching different native data sources spread across different locations in an organisation.

Implementation aspects

This implementation starts by modelling a simple ontology for the forensic and compliance audit processes, which encompasses the norms, policies, and legal or regulatory guidelines that are being applied. The ontology will allow users to infer new knowledge, for example, to identify possible unauthorised or incompatible access to the assets of a large organisation. This example implements a federated query web service for evaluating whether employees have the required roles when they access those assets. An intermediary layer translates the requests arrived to the web service into queries for the internal schemas of the involved databases.

The interface layer is implemented as a web service, while the mapping brokers are implemented as D2R Server endpoints. Each endpoint is assigned to different relational database(s). **Figure 4,** below, provides a general overview of the implementation of the described architecture, depicting how requests flow from a submitted query to the web service, which implements a federated query solution to dispatch the incoming requests to the indexed list of database servers—with each of them mapped by a specific D2RQ component. For simplicity's sake, the figure includes just two different databases with different schemas (one Microsoft database—MSSQL—and one MySQL database), but there are no limits to the number or type of involved databases.

The use case hereby described involves a client requesting the contents of the 'Roles' database entity. The objective is to gather and combine—without requiring the end user to be aware of low-level details—information dispersed across different tables and different databases which use different schemas. After the request query to retrieve the existing contents from the 'Rules' entity has reached the database instances, each delivers its contents to a SPARQL endpoint through a

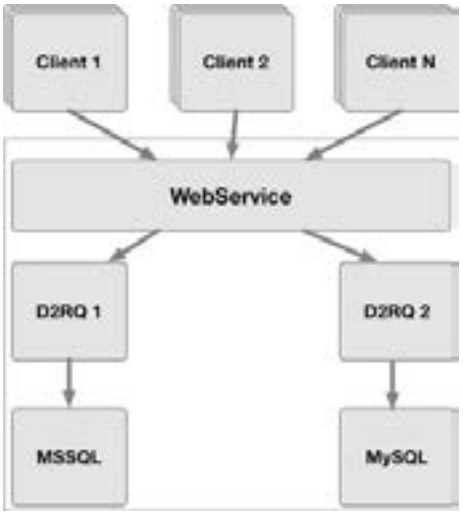


Figure 4: Architecture implementation

D2R server assigned to each involved database. The D2RQ Mapping Language is used for the mapping process. This central web service allows clients to directly query existing entities, to retrieve available content from each existing database, and to merge and deliver them to the querying clients.

Required tools and technologies include Visual Studio as development environment, C# as programming language, ASPX.NET for implementing the web service, classic RDBs such as MSSQL and MySQL, and the RDF and SPARQL languages describing their semantics.

The following sections discuss some details for each step involved in the implementation and deployment of this specific use case. First, a simple ontology is presented. Next, some relevant implementation steps are discussed, such as deploying the database server, generating mapping, configuring the mapping between the database server and the ontology, activating D2R servers with the corresponding mappings, and describing the web service.

Create the Ontology: In this section, a simple ontology is explored in the domain of compliance audit to support the previously presented use-case scenario, which has the main purpose of answering the following question: 'Who is able to access the assets, for maintenance purposes, in a large company spread out through different countries and businesses?'

The ontology, built within Protégé, includes classes for 'Asset', 'Employee', 'Organization', and 'Role'. The corresponding instances are 'Computer', 'John' and 'Francis', 'PowerPlantA', and 'MaintainsIT'. The ontology does not include any hierarchy of concepts.

Table 1, below, summarises the relationship among class instances, their types and property assertions.

Instance	**Type**	**Property Assertions**
John		
	Employee	isEmployedBy:PowerPlantA
Number: '1002'		
Name: 'John'		
Francis	Employee	isEmployedBy: PowerPlantA
hasRole:MaintainsIT		
Number: '1001'		
Name: 'Francis'		
MaintainsIT	Role	maintains:Computer isMaintained-By: Francis
Name: 'Francis'		
PowerPlantA	Organization	hasEmployees:John hasEmployees: Francis
hasAssets: Computer		
Name: 'PowerPlantA'		
Computer	Asset	isRoledBy: Francis

belongsTo:PowerPlantA		
Number: '10000001'		
Name: 'DELL		

Table 1: Classes instances

'John' and 'Francis' are instances of 'Employee'. Both have the property 'isEmployedBy' assigned with the value 'PowerPlantA'. The employee is assigned roles granting the access to the assets, enabling the building of a query to assess the regulatory rules and policies. It also has as an inverse property 'hasRole' as 'MaintainsIT'. Additionally, they have data properties '1' and '2' for the 'Number', and 'Francis' and 'John' for 'Name'. Notwithstanding, the difference between 'Francis' and 'John' instances is that the 'Francis' does not include the property 'hasRole' as 'MaintainsIT'. Therefore, they will be considered two employees for the organisation, but just one of them is able to maintain the assets.
'PowerPlantA' is an instance of the 'Organization' type and includes the property 'hasEmployees' for 'Francis' and 'John' instances. Therefore, this organisation has two employees. 'Computer' is an instance of the 'Asset' type and its properties are 'isRoledBy' of the 'MaintainsIT' instance, whose value is 'Francis' and which includes a 'Number' and a 'Name'.

Figure 5, below, provides the full contents of the above ontology, in turtle language, located at 'data.ttl' file:

```
#filename: data.ttl
@prefix FCA: <http://www.semanticweb.org/FCA#>
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
@prefix owl: <http://www.w3.org/2002/07/owl#>
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
@prefix xsd: <http://www.w3.org/2001/XMLSchema#>

#################################################################
#    Object Properties
#################################################################

###  http://www.semanticweb.org/FCA#belongsTo
FCA:belongsTo rdf:type owl:ObjectProperty ;
owl:inverseOf FCA:hasAssets ;
rdfs:domain FCA:Asset ;
rdfs:range FCA:Organization .

###  http://www.semanticweb.org/FCA#hasAssets
FCA:hasAssets rdf:type owl:ObjectProperty ;
rdfs:domain FCA:Organization ;
rdfs:range FCA:Asset .

###  http://www.semanticweb.org/FCA#hasEmployees
FCA:hasEmployees rdf:type owl:ObjectProperty ;
owl:inverseOf FCA:isEmployedBy ;
rdfs:domain FCA:Organization ;
rdfs:range FCA:Employee .
```

```
###  http://www.semanticweb.org/FCA#hasRole
FCA:hasRole rdf:type owl:ObjectProperty ;
owl:inverseOf FCA:isRoledBy ;
rdfs:domain FCA:Employee ;
rdfs:range FCA:Role .

###  http://www.semanticweb.org/FCA#isEmployedBy
FCA:isEmployedBy rdf:type owl:ObjectProperty ;
rdfs:domain FCA:Employee .

###  http://www.semanticweb.org/FCA#isRoledBy
FCA:isRoledBy rdf:type owl:ObjectProperty ;
owl:inverseOf FCA:isRoledBy ;
rdfs:domain FCA:Role ;
rdfs:range FCA:Employee .

#################################################################
#    Data properties
#################################################################

###  http://www.semanticweb.org/FCA#Name
FCA:Name rdf:type owl:DatatypeProperty ;
rdfs:domain FCA:Asset ,
FCA:Employee ,
FCA:Organization ,
FCA:Role .

###  http://www.semanticweb.org/FCA#Number
FCA:Number rdf:type owl:DatatypeProperty ;
rdfs:domain FCA:Asset .

#################################################################
#    Classes
#################################################################

###  http://www.semanticweb.org/FCA#Asset
FCA:Asset rdf:type owl:Class .

###  http://www.semanticweb.org/FCA#Employee
FCA:Employee rdf:type owl:Class .

###  http://www.semanticweb.org/FCA#Organization
FCA:Organization rdf:type owl:Class .

###  http://www.semanticweb.org/FCA#Role
FCA:Role rdf:type owl:Class .

#################################################################
#    Individuals
#################################################################

###  http://www.semanticweb.org/FCA#Computer
FCA:Computer rdf:type owl:NamedIndividual ,
FCA:Asset ;
```

```
FCA:belongsTo FCA:PowerPlantA ;
FCA:Name “DELL”^^xsd:string ;
FCA:Number “10000001”^^xsd:int .

###  http://www.semanticweb.org/FCA#Francis
FCA:Francis rdf:type owl:NamedIndividual ,
FCA:Employee ;
FCA:hasRole FCA:MaintainsIT ;
FCA:isEmployedBy FCA:PowerPlantA ;
FCA:Name “Francis”^^xsd:string ;
FCA:Number “1001”^^xsd:int .

###  http://www.semanticweb.org/FCA#John
FCA:John rdf:type owl:NamedIndividual ,
FCA:Employee ;
FCA:isEmployedBy FCA:PowerPlantA ;
FCA:Name “John”^^xsd:string ;
FCA:Number “1002”^^xsd:int .

###  http://www.semanticweb.org/FCA#MaintainsIT
FCA:MaintainsIT rdf:type owl:NamedIndividual ,
FCA:Role ;
FCA:isRoledBy FCA:Francis ;
FCA:Name “MaintainsIT”^^xsd:string .

###  http://www.semanticweb.org/FCA#PowerPlantA
FCA:PowerPlantA rdf:type owl:NamedIndividual ,
FCA:Organization ;
FCA:hasAssets FCA:Computer ;
FCA:hasEmployees FCA:Francis ,
FCA:John ;
FCA:Name “PowerPlantA”^^xsd:string .
```

Figure 5: Ontology definition

Deploy the database server: This step involves the creation of the table objects for MySQL and MSSQL databases, as well as the commands for populating them. For the sake of demonstration, the MSSQL database table schemas and contents are different from the ones used in the MSSQL database. At the end, these two databases should maintain different data over distinct schemas, which will become federated at the upper level of the web service. The applied commands were the following:

```
generate-mapping -u root -p password01pt -o ssfile_MYSQL.ttl -d com.microsoft.
sqlserver.jdbc.SQLServerDriver   jdbc:sqlserver://host_mysql;databaseName=BD_
mssqlDB

generate-mapping -u sa –p password02pt -o ssfile_SQLServer.ttl –d com.microsoft.
sqlserver.jdbc.SQLServerDriver   jdbc:sqlserver://host_mssql;databaseName=BD_
mysqlDB
```

Prepare mapping: The mapping process between database and RDF schemas is mapped through the ‘ssfile_SQLServer.ttl’, whose contents include the mapping between the MSSQL server and

RDF schemas—the 'ssfile_MYSQL.ttl' file plays the same role, but for the MySQL schema. The initial section of these files includes a set of prefixes (several were removed from the next listing for clarity), with the map:database component providing a way for retrieving information from the database server. These files were manually updated to allow the correct mapping between RDF and the database schemas. This mapping is supported by RDF d2rq:ClassMap and d2rq:PropertyBridgefor classes and properties, respectively. **Figure 6,** below, includes the contents for mapping the class 'Employee' and table 'Employee' from the MSSQL server:

```
@prefix map: <#> .
@prefix db: <> .
@prefix vocab: <vocab/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix d2rq: <http://www.wiwiss.fu-berlin.de/suhl/bizer/D2RQ/0.1#> .
@prefix jdbc: <http://d2rq.org/terms/jdbc/> .

map:database a d2rq:Database;
     d2rq:jdbcDriver "com.microsoft.sqlserver.jdbc.SQLServerDriver";
     d2rq:jdbcDSN "jdbc:sqlserver://localhost;databaseName=BD_joaohenriques";
     d2rq:username "joaohenriques";
     d2rq:password "password1";
     .

# Table CREATE TABLE dbo.Employee (Number INT, Name VARCHAR(100))

map:dbo_Employee a d2rq:ClassMap;
     d2rq:dataStorage map:database;
     d2rq:uriPattern "dbo/Employee/@@dbo.Employee.Number@@";
     d2rq:class vocab:dbo_Employee;
     d2rq:classDefinitionLabel "dbo.Employee";
     .
map:dbo_Employee__label a d2rq:PropertyBridge;
     d2rq:belongsToClassMap map:dbo_Employee;
     d2rq:property rdfs:label;
     d2rq:pattern "Employee #@@dbo.Employee @@";
     .
map:dbo_Employee_Number a d2rq:PropertyBridge;
     d2rq:belongsToClassMap map:dbo_Employee;
     d2rq:property vocab:dbo_Employee_Number;
     d2rq:propertyDefinitionLabel "Employee Number";
     d2rq:column "dbo.Employee.Number";
     d2rq:datatype xsd:integer;
     .
map:dbo_Employee_Name a d2rq:PropertyBridge;
     d2rq:belongsToClassMap map:dbo_Employee;
     d2rq:property vocab:dbo_Employee_Name;
     d2rq:propertyDefinitionLabel "Employee Name";
     d2rq:column "dbo.Employee.Name";
     d2rq:datatype xsd:string;
```

Figure 6: Mapping between RDF and database schemas

Activate D2R servers: The next step deploys the D2R server, in order to map the contents from RDB to RDF according to the mapping file. The following command activates the MSSQL and MYSQL servers respectively:

```
d2r-server -p 2021 ssfile_SQLSERVER.ttl

d2r-server -p 2020 ssfile_MYSQL.ttl
```

Activate web service: The web service provides the main functions performing the federation mechanism and retrieving the information from the SPARQL endpoints. The web service provides an interface and a federated query layer and offers query services that allow end users to perform the intended inference operations while remaining abstracted from low-level details. Each submitted query is forwarded to multiple RDBs through a DR2Q component. The results are later merged into a single result set. The endpoints are configured at server level, and take into consideration the fact that the end user does not need to know the number or the location of such existing endpoint servers. The web service endpoint is located at

```
http://host_webservice:17129/WebService1.asmx?op=SemanticWEB.
```

Query the ontology: The final step is to query the knowledge base. The SPARQL query in **Figure 7,** below, requests the knowledge base for assessing which users are authorised to execute the maintenance of the assets in a given organisation. This query is forwarded from the Web service to all the federated SPARQL endpoints assigned to different databases and which is finally translated into the internal schema of those databases. The query filters the organisation 'PowerPlantA' for the asset 'Computer', where just some of the employees having the role 'MaintainIT' are authorised to perform its maintenance:

```
PREFIX : <http://www.semanticweb.org/FCA#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>

SELECT *
WHERE{
     ?employee rdf:type owl:NamedIndividual.
     ?employee :hasRole ?role.
    ?organization rdf:type :Organization.
     ?organization :hasEmployees ?employee.
     ?asset rdf:type owl:NamedIndividual.
     ?role :isRoledBy ?employee.
     FILTER(?organization = :PowerPlantA)
     FILTER(?asset = :Computer )
     FILTER(?role = :MaintainsIT )
 }
```

Figure 7: SPARQL query for assessing authorised users

Figure 8, below, demonstrates the use of the Apache Jena SPARQL command 'sparql --data=data.ttl --query query.rq' and the corresponding output. The query contents are located in

the 'query.rq' file, which was used against data located at the 'data.ttl' file. According to the knowledge base, just 'Francis' is able to execute the 'Computer' maintenance.

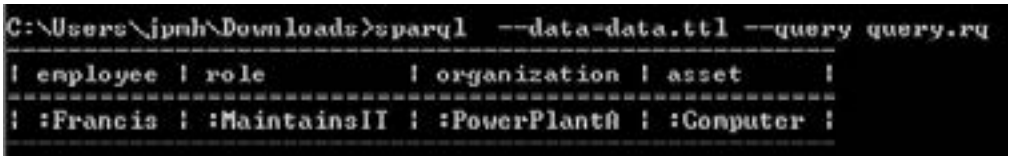


Figure 8: SPARQL command

Discussion and Conclusions

This paper proposes an approach for leveraging inference capabilities in the use of heterogeneous data currently maintained in multiple, natively different RDB systems. This approach aims at contributing to Critical Infrastructure Protection by supporting activities such as forensic analysis and compliance audit procedures. It provides Semantic Web reasoning capabilities through an interface able to answer to federated queries. The process of interactively exploring, searching, extracting, pinpointing, and combining insights can use and combine data sourced from disparate organisational RDBs. Thus, this approach avoids the duplication of information in RDB and RDF stores, and overcomes the issues arising from the use of static data integration (such as the lack of support for transformations of data and the effort required for maintaining up-to-date synchronisation processes). The proposed web service includes an abstraction layer that deals with inherent complexities of resorting to different platforms, systems, technologies, and information schemas to retrieve and to combine heterogeneous data. This abstraction layer also improves security by hiding the infrastructure's internal details.

Although the approach taken by the proposed federated architecture is similar to the one of SPARQL 1.1, it does not require previous knowledge about the existence and location of SPARQL endpoints. The benefits of this approach come from the inclusion of an abstraction layer, which provides direct access to operational data that live in different organisational RDBs. Details such as the involved database servers and differences between schemas can be kept away from users. Moreover, it is flexible enough for leveraging the exploration of additional data sources that might be easily added in the future. The proposed framework also provides a data fusion solution for gathering multiple data items—representing the same real-world object—into a single, consistent, and clean representation.

This work arises from the limited research on the use of ontology data for CIP applications, and the need to improve and facilitate the usage of the huge amounts of data living in the RDBs of Critical Infrastructure operators. This work also explored Semantic Web inference tools, and is aimed at the practical objective of federating queries against a knowledge base containing the ontology and data for assessing employee authorisations for asset maintenance in a large organisation that uses multiple different RDBs. This practical approach suggests a future path for the improvement of CIP by using inference capabilities for forensic and compliance audit purposes and leveraging the use of heterogeneous ontology data living in RDBs and in other heterogeneous kinds of data sources.

Acknowledgements
This work was partially funded by the ATENA H2020 Project (H2020-DS-2015-1 Project 700581).

References
ATENA 2018 'H2020 ATENA Project website', viewed 24 May 2017, <https://www.atena-h2020.eu/>.

Bizer, C & Cyganiak, R 2007, 'D2RQ: Lessons learned', Position paper for the W3C, Workshop on RDF Access to Relational Databases, Cambridge, MA, US.

Blackwell, J, Tolone, WJ, Lee, SW, Xiang, WN & Marsh, L 2008, 'An ontology-based approach to blind spot revelation in critical infrastructure protection planning', *Proceedings of the International Workshop on Critical Information Infrastructures Security,* Springer, Berlin, Heidelberg, DE, pp. 352-59.

Brickley, D & Guha, RV 1999, *Resource Description Framework (RDF) Schema specification, Proposed recommendation, World Wide Web Consortium,* viewed 12 December 2017, <http://www.w3.org/TR/PR-rdf-schema>.

Castorini, E, Palazzari, P, Tofani, A & Servillo, P 2010, 'Ontological framework to model Critical Infrastructures and their interdependencies', *Proceedings of Complexity in Engineering,* COMPENG'10, pp. 91-3.

Chorás, M, Kozik, R, Flizikowski, A & and Hołubowicz, W 2010, 'Ontology applied in decision support system for critical infrastructures protection', *Proceedings of the international conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems (IEA/AIE 2010): Trends in Applied Intelligent Systems,* pp. 671-80.

D2RQ 2012, 'D2RQ', viewed 31 August 2017, <http://d2rq.org>.

DuCharme, B 2013, *Learning SPARQL: Querying and updating with SPARQL 1.1,* 2nd edn, O'Reilly Media, Inc.

Freitas, A, Curry, E, Oliveira, JG, and O'Riain, S 2012, 'Querying heterogeneous datasets on the linked data Web: Challenges, approaches, and trends', *IEEE Internet Computing,* vol. 16, no. 1, pp. 24-33.
Gorlitz, O & Staab, S 2011, 'SPLENDID: Sparql endpoint federation exploiting void descriptions', *Proceedings of the second international conference on Consuming Linked Data,* vol. 782, pp. 13-24.

Gruber, TR 1993, 'A translation approach to portable ontology specification', *Knowledge Acquisition,* vol. 5, pp. 199-220.

Henriques J, Caldeira F, Cruz T & Simões P 2018, 'On the use of ontology data for protecting critical infrastructures', *Proceedings of the 17th European Conference on Cyber Warfare and Security (ECCWS)*, Oslo, NO.

Musen, MA 1992, 'Dimensions of knowledge sharing and reuse', *Computers and Biomedical Research,* vol. 25, pp. 435-67.

Noy, NF & McGuinness, DL 2001, 'Ontology development 101: A guide to creating your first ontology', *Stanford knowledge systems laboratory technical report KSL-01-05,* viewed 23 November 2017, <http://www.corais.org/sites/default/files/ontology_development_101_aguide_to_creating_your_first_ontology.pdf>.

Quilitz B & Leser U 2008, 'Querying distributed RDF data sources with SPARQL', S Bechhofer M Hauswirth, J Hoffmann, M Koubarakis (eds), *The Semantic Web: Research and applications,* European Semantic Web Conference (ESWC) 2008, Lecture notes in computer science, vol. 5021, Springer, Berlin, Heidelberg, DE, pp. 521-38.

RDF (Resource Description Framework) 2014, 'W3C Resource Description Framework (RDF)', viewed 4 September 2017, <https://www.w3.org/RDF>.

Rosa L, Proença J, Henriques J, Graveto V, Cruz T, Simões P, Caldeira F & Monteiro E 2017, 'An evolved security architecture for distributed industrial automation and control systems', *Proceedings of the 16th European Conference on Cyber Warfare and Security (ECCWS).*

Sahoo, S, Halb, W, Hellmann, S, Idehen, K, Thibodeau, T, Auer, S, Sequeda, J & Ezzat, A 2009, *A survey of current approaches for mapping of relational databases to RDF,* viewed 24 October 2017, <http://www.w3.org/2005/Incubator/rdb2rdf/RDB2RDFSurveyReport.pdf>.

Seaborne A, Polleres A, Feigenbaum L & Williams, G 2013, 'SPARQL 1.1 Federated Query', position paper for the W3C Workshop on SPARQL 1.1 Federated Query, viewed 4 September 2017, <https://www.w3.org/TR/sparql11-federated-query/>.

Schenk, S, Saathoff, C, Staab, S & Scherp, A 2009, 'SemaPlorer—Interactive semantic exploration of data and media based on a federated cloud infrastructure', *Web Semantics: Science, services and agents on the World Wide Web,* vol. 7, no. 4, pp. 298-304, viewed 13 September 2017, <http://doi.org/10.1016/j.websem.2009.09.006>.

Sernadela, P, González-Castro, L & Oliveira, JL 2017, 'SCALEUS: Semantic Web services integration for biomedical applications', *Journal of Medical Systems,* vol. 41, no. 4, p. 54.

Schwarte, A, Haase, P, Hose, K, Schenkel, R & Schmidt, M 2011, 'FedX: Optimization techniques for federated query processing on linked data', L Aroyo, et al. (eds), *The Semantic Web – ISWC 2011,* Lecture notes in computer science, vol. 7031, Springer, Berlin, Heidelberg, DE, pp.601-16.

W3 2013, 'SPARQL', viewed 4 September 2017, <https://www.w3.org/TR/sparql11-query>.

——2014, 'SPARQL', viewed 4 September 2017, <https://www.w3.org/TR/turtle>.

A Cultural Exploration of Social Media Manipulators

C Sample[1], J McAlaney[2], JZ Bakdash[3], H Thackray[2]

[1]ICF for United States Army Research Laboratory
Adelphi, Maryland, United States

E-mail: char.sample@icf.com

[2]Department of Psychology
Bournemouth University
Bournemouth, Dorset, United Kingdom

E-mail: jmcalaney@bournemouth.ac.uk;hthackray@bournemouth.ac.uk

[3]United States Army Research Laboratory South
University of Texas at Dallas
Richardson, TX, United States

E-mail: jonathan.z.bakdash@mail.mil

Abstract: *Internet social media sites enable the rapid and widespread production and dissemination of propaganda. Although propaganda has a long history in warfare, the spreading of propaganda via social media is markedly different from past distribution methods. The authors investigated the relationships between state-affiliated actors who use social media to produce and distribute propaganda and their national cultural values. The results showed that countries that deployed fake news via social media tended to have more masculine cultural values. These findings suggest that specific cultural values are associated with fake news distribution, which may indicate that culturally aware responses may be more effective in responding to propaganda.*

Keywords: *Propaganda, Cultural Values, Social Media, Hofstede, Trust*

Introduction

The U.S. political events of 2016 brought to the forefront concerns about mass information operation campaigns and their effects, particularly the social and political effects of propaganda. The use of social media to produce and to deliver propaganda represents a new, low-cost, rapid, and more effective mechanism for furthering the persuasive goals of state actors. Lee and Kent (2017) noted that approximately one third of the U.S. population received Russian propaganda on Facebook during the 2016 election cycle. Gottfried and Shearer (2016) found that, in 2016, a majority of U.S. citizens used social media sites as a news source.

Social media sites such as Facebook and Twitter were not initially designed or envisioned as primary news sources. An examination of their mission statements shows these sites were envisioned as communications' enabling forums. The original mission statement for Twitter reads as follows:

"To give everyone the power to create and share ideas and information instantly, without barriers" (Fox 2014). Facebook's original mission statement says that "Facebook's mission is to give people the power to share and make the world more open and connected" (Facebook n.d.). In 2017, Facebook updated its mission statement: to "Give people the power to build community and bring the world closer together" (Facebook 2019). In each of these statements, the mission does not mention news sharing or, for that matter, information or knowledge sharing. However, in retrospect, a communications medium, particularly a synchronous medium, is a tempting target for deceptive data, especially since the target can provide immediate feedback for the purveyor.

The social media environment as a news dissemination paradigm differs from the original news paradigm for print and broadcast, in which control of production and dissemination of news was concentrated to a small group of broadcasters who were granted licenses to operate from their governments. However, social media sites with the ability to 'like' and 'share' information or stories creates an environment that appears as news, and is now able to spread without controls. This unconventional use of social media sites as primary news sites suggests that the adoption patterns associated with these information-shaping behaviours may differ from the adoption patterns associated with social media usage in general. Sample and Karamanian (2014) observed collectivism, indulgence, and short-term orientation with Facebook adoption rates, whereas Twitter adoption was associated with masculine and indulgent values.

Although social media sites may not have originally been envisioned as news media, the recognition of the value of social media sites (such as Twitter) as news sources grew during and after the Arab spring (Howard & Hussain 2011; Comunello & Anzera 2012). During the Arab Spring, Twitter feeds from affected Middle East countries produced a narrative that countered the official government versions of events (Howard & Hussain 2011; Comunello & Anzera 2012) and that provided accurate images of events that were officially unavailable. At this point in time, perception of social media sites had changed from social conversation platforms to trusted news sources.

In 2017, Bradshaw & Howard compiled a list of countries where active engagement in propaganda via social media spread, as well as the dissemination methods for each country. This list, consisting of 29 countries provides the launching point for this study and analysis. When examining this list of 29 countries in the context of cultural observations of social media usage (see Hofstede, Hofstede & Minkov 2010; Sample & Karamanian 2014), the movement from information seeking to information shaping on social media appears to be a logical next step. This exploratory study was performed with the purpose of determining whether cultural values might be associated with some of the behaviours and trends associated with the use of social media sites for production and creation of propaganda.

Background

Propaganda and deception have a long history in warfare (Shultz 1989; Taylor 2013); however, in the past, the distribution or reach was limited (Crilley 2001) by the media--for example, through person-to-person communication, newspapers and other printed materials, radio, and television. The reach of these sources is now less effective due to speed and cost. More recently, Commin and Filiol (2015) noted the breakdown of the traditional boundaries of war along with a movement away from the model of bending the enemy to one's will (Chacon 2006) to a different goal

of having the victim carry out the attack against themselves (Cybenko, Giani & Thompson 2002). In this new age of hybrid warfare propaganda, a form of deception is a natural weapon for use in perception shaping which is necessary for victims' self-attack.

An important step in the process of having targeted audiences enact attacker goals relies on whether the message is able to reach the target. In the global virtual environment, much attention is given to the reachability and vulnerabilities of various technologies; however, the reachability and vulnerabilities of the human are equally important (Szfranski 1997; Franke 2015). Some countries have mitigated this vulnerability by controlling the reachability of its citizens (Clayton, Murdoch & Watson 2006; Aryan, Aryan & Halderman 2013), but the directly reachable citizens of many Western democracies remain open to persuasion through propaganda.

Propaganda is information or ideas that are spread by an organised group or government to influence people's opinions, especially by not giving all the facts or by secretly emphasising only one way of looking at the facts ('Propaganda'). Fake news is "false stories that appear to be news spread on the [I]nternet or using other media, usually created to influence political views or as a joke" ('Fake news'). Both definitions have overlapping features, most notably the goal of influencing or persuading and the use of the Internet as a distribution channel. The term 'fake news' has become so common that the 2017 issue of the Oxford Dictionary of English identifies the phrase as the word of the year (Meza 2017). The manner in which war is waged in the cyber domain challenges assumptions about cyber space, the entities in cyber space, and the relationship between those entities. Some countries' highly effective deployment of deception techniques through the use of propaganda challenges some traditionally held beliefs, while the sophistication also continues to grow.

There are several aspects to consider when discussing the role of social media in the production and spread of state-sponsored propaganda, and these range from technical to operational to behavioural. This study is focused on the behavioural aspects of propaganda, specifically trust-exploiting behaviours. Some societal values appear to encourage a degree of trust or openness (Hofstede, Hofstede & Minkov 2010). In 2017, specific societal values associated with higher victimisation rates through social engineering were observed, which suggests that some cultures might be more trusting in the online environment than others (Sample et al. 2017). If cultural values are associated with a willingness to trust online messages, then these same cultural values may also be associated with the production and dissemination of crafted messaging as a part of social media manipulation in support of propaganda.

Social media manipulation

The use of Facebook and Twitter for purposes not listed in the sites' mission statement illustrate the changing nature of social media sites. Facebook and Twitter have become launching points and prominent spread sites for fake news. Cultural differences have been observed in societal use and interactions with technology (Elmasry, Auter & Peuchaud 2014). This variance appears to be consistent with previous research and observations (Hofstede, Hofstede & Minkov 2010; Sample & Karamanian 2014; Sample, Cowley & Hutchinson 2017; Sample et al. 2017), suggesting that national cultural values may be associated with propaganda production, spread, and even the method of dissemination when social media has been deployed.

Members of open societies that have a protected, independent press may be less familiar with media manipulation than those from closed societies with a long history of government-controlled messaging. Combining this background with the changing nature of social media sites as news sharing sites (Gottfried & Shearer 2016), the open-society content readers have no reason to distrust the material that they are receiving. Furthermore, when social media sites became sources of truthful news, the subscribers could view the sites as credible sources of news stories, even after the quality of the content has changed.

One reason for this unanticipated use of social media for fake news activities may be explained by behavioural traits or cultural values associated with the actors in this arena. The willingness of the target actors to trust both users and computers as sources of information requires a deeper understanding of the nature of online trust relationships. Trusting behaviour entails the individual's surrendering control over valuable outcomes, with the expectation that the other will reciprocate. Furthermore, a shared social group identity is a strong predictor of trusting behaviour between individuals (Tanis & Postmes 2005). With online trust, Friedman, Khan, Jr., and Howe (2000) highlighted that, regarding technology, the term 'trust' is often broadly used to refer to expectations rather than a considered trusting behaviour as described above. They concluded that people trust other people, not technology (Friedman, Khan, Jr & Howe 2000). It has been argued that trust can also be considered in terms of how much risk and uncertainty individuals are willing to accept, including in their interactions with online systems (Jones & Moncur 2018). Context is also significant: people are less trusting of situations involving their financial information as opposed to their other personal information. This disclosure suggests that the potential manipulation of individuals may be easier when finances are not involved. The existence of cognitive heuristics and biases may further influence the degree to which people believe and trust fake news. Exploitation of these processes allows fake news creators to optimise the target and subsequently spread the fake news item. People are likely to deem reliable and trustworthy information that confirms their pre-existing beliefs, which is known as the confirmation bias (Kahneman, Slovic & Tversky 1982). This is a reflection of the use of heuristics in decision making. The term 'heuristics' refers to the use of mental shortcuts individuals employ in order to reduce the cognitive load that would otherwise be required to make complex decisions (Kahneman, Slovic & Tversky 1982). Social media can deliver news at high intensity; it is perhaps only to be expected that, in order to cope with this intensity, individuals are more likely to employ heuristics when judging these news items than when they are using more traditional mediums. Such processes may be actively targeted and exploited by the creators of fake news. For instance the representativeness heuristic (Kahneman, Slovic & Tversky 1982) refers to the tendency to attribute characteristics to something if it matches the preconceived prototype of a category. In the case of a fake news story, this equates to the use of a website design style that mirrors that used by the websites of established print newspapers, in order to convince the reader of the legitimacy of the source. It has been suggested that there is a relationship between cultural values and the use of heuristics. For example, it has been noted that there are difference in the use of the representativeness heuristic between Canadian and Chinese participants (Spina et al. 2010).

Once the trust relationship between the reader and the content provider is in place, information shaping and distribution are possible. Considering the competitive nature of masculine societies along with the unrestricted boundaries of hybrid warfare, social media sites that act as news

sources are reasonable targets. According to Hofstede, Hofstede & Minkov (2010), competitive, masculine societies will use social media sites for news information seeking as a way to gain a competitive advantage, whereas feminine societies will use social media sites for social information sharing (Sample & Karamanian 2014). The migration of social media sites from information seeking to information shaping represents the next frontier and is now known as the fifth domain of war (Lynn 2010). Lee and Kent (2017) report that Facebook discovered that 120 fake, Russian-backed pages made 80,000 posts that went directly to 29 million Facebook users who 'liked' and 'shared' information with 126M users (Lee & Kent 2017). Twitter was also used in the same manner (Booth et al. 2017).

Another new aspect is the automated distribution channel or the reliance on bots. The use of bots as an automated distribution channel is unique to war in the cyber domain. Considering that bots are a relatively young technology (Stinson & Mitchell 2007), their use can also be considered as a form of technology adoption. Bot usage also represents an example of the 21st-century adaptation to the delivery of the fake news payload that was deemed trustworthy by the recipients that reached a large population. Bots are also capable of quickly gaining the target's trust through the use of artificial intelligence that can model its responses by the inputs that it receives. Thus, the bot that is designed to anger the target will put forth phrases that match the target's values and fears. These programs work by probabilistic predictions and, in some cases, have gone rogue in their responses based on the inputs received (Neff & Nagy 2016). While this study does not address chatbots specifically, the behavioural aspects shaped by both humans and machines are of interest and deserve mention.

Values

Cultural values as defined by Hofstede are composed of six dimensions: Power Distance Index (PDI), Individualism versus Collectivism (IvC), Feminine versus Masculine (FvM), Uncertainty Avoidance (UAI), Long-term versus Short-term orientation (LvS), and Indulgence versus Restraint (IvR). A brief discussion of each of these dimensions follows.

- PDI—This dimension details an authoritarian or egalitarian society's ideals. Power in the high PDI society originates at the top, suggesting that trust relationships may occur among peers or when lower societal members rely on higher members; previous observations noted that senior members provide protection to the junior members of their group (Hofstede, Hofstede & Minkov 2010; Nisbett 2010). In egalitarian societies where "truth is spoken to power" (Hofstede, Hofstede & Minkov 2010), trust may likely be more easily granted across a wider population where the 'in group' and 'out group' are perceived as being closer in the low PDI societies (Nisbett 2010). In low PDI societies, risks tend to be rewarded (Guess 2004; Hofstede, Hofstede & Minkov 2010). Propaganda emanating from high authoritarian groups should contain a disciplined message that is supported at all levels of society. The spread of fake news within low versus high PDI cultures may also be determined by how secret that information is. This refers to the secrecy heuristic, where there is a tendency to perceive information as being more credible if that information is framed as being something that the individual is not meant to know (Travers, Van Boven & Judd 2014). The revelation of 'shocking' information is oftentimes the basis of fake news stories, such as those stories during the 2016 U.S. presidential election claiming that can-

didate Clinton had suffered a heart attack (Gillin 2017). In high PDI cultures, this secrecy heuristic effect could be enhanced, because there may be a greater assumption that those in power are privy to information that is unknown to the general public.

- IvC—This dimension defines a societal individual's view of self as related to the larger society. Collectivist societies view the individual as an important link in a larger chain (Nisbett 2010). The implication for social media manipulation is that collectivist values might result in greater consistency, whereas individualist societies might exhibit less consistency but greater breadth in the spread mechanisms.

- FvM—This dimension defines the manner in which a society deals with conflict, with masculine societies dealing directly and seeking a solution where winners and losers are present, whereas feminine societies seek to negotiate so that each side wins.

- UAI—This dimension focuses on how a society deals with the unknown. The high UAI society needs assurance in order to accept what is unknown, whereas in the low UAI society curiosity overrules fear. This relates to predictability of the environment. A culture with low UAI may be less likely to pay attention to a fake news story that depicts a deviation from social norms. On the other hand, cultures with high UAI may be especially sensitive to any fake news story that indicates a breach of a social norm, because this in turn suggests greater uncertainty. This is in keeping with psychological research that suggests people are particularly influenced by information that does not match their preconceptions (Hemsley & Marmurek 1982). The high UAI societies may be drawn to shape information with precision and consistency.

- LvS—This dimension defines a societal preference for immediate rewards versus waiting for gratification. This relates to trust as a long-term culture relies upon ongoing, harmonious relationships between individuals if the culture is to survive into the future, meaning that agreement between and trust of others is encouraged (Hofstede, Hofstede & Minkov 2010).

- IvR—This dimension defines how societal members express themselves and ranges from stoic with little to no shows of emotion to large celebrations.

Method

The examination of culture creates challenges due to the existence of unconscious social and cultural biases that everyone possesses (Nosek, Hawkins & Frazier 2011; Fiske & Taylor 2013). Objectivity, although difficult to attain, is still a primary goal; thus, quantitative analysis has advantages. Furthermore, observation of behaviour in a natural setting is difficult. In this particular case, the observable data were collected and analysed for a different study; the researchers for this study re-used the collected data to determine whether common cultural values can be observed. The research questions used are listed below.

- **R1: Use of Social Media by Propaganda Purveyors**

Do the purveyors of propaganda who use social media sites differ culturally from those who do not? Evaluation of results relies on using the full set of countries that Hofstede defined, and dividing the countries into two groups: those who use social media sites to deliver propaganda and the remaining countries from the Hofstede pool of countries. The two groups will be compared to determine how similar or dissimilar they are in terms of distributions using the Mann-Whitney-Wilcoxon (MWW) test (Hollander, Wolfe & Chicken 2013). The researchers are testing against the null hypothesis that states $\mathbf{H_0}$**:** There are no differences in the distribution cultural values among the purveyors and the non-purveyors. Thus, $\mathbf{H_1}$ represents the alternative hypothesis that must be considered if $p \leq 0.01$, after adjusting for multiple comparisons using a Bonferroni adjustment (α/comparisons = 0.05/5 comparisons) (Hollander, Wolfe & Chicken 2013).

- **R2: Cultural Values over a period of time for Countries with Social Media Propaganda Purveyors and without**
  Are there any cultural value trends that can be observed on the social media site propaganda purveyors? The second research question can be answered by evaluating the median values of purveyor over the seven-year time interval. The median values for each year will be paired with the year, and a Spearman correlation will be run (Hollander, Wolfe & Chicken 2013). $\mathbf{H_2}$: There are no associations between the cultural values and social media propaganda distributors vs. non-distributors. $\mathbf{H_3}$: represents the alternative hypothesis that must be considered, an inferred relationship between cultural values and propaganda distributors who use social media.

- **R3: Cultural Values and Methods for Social Media Propaganda Delivery**
  Do any cultural values associate with the method of propaganda delivery? For this set of data, the groups are small so that the standard tests of MWW and Spearman are not appropriate (Hollander, Wolfe & Chicken 2013). The median values are descriptively examined and compared against the overall Hofstede median values, and significant differences (greater than 10) will be discussed along with the importance of the result.

Of the 29 countries listed in the Bradshaw and Howard (2017) study, 24 of the countries were found in Hofstede's data (Hofstede n.d.); countries that were not found in Hofstede's data were excluded. In some cases, the social media methods were not identified, in which case those countries were excluded in the methods portion of the study (R3), but they are included in the processing for R1 and R2. **Table 1**, below, provides the listing of countries, their Hofstede cultural values, the number of years in the Bradshaw and Howard (2017) study, and the methods (A for automated, H for human, and B for both human and automated) used to disseminate propaganda. Since the focus of this study is on cultural values, the country names are not listed, but the reader can determine the country names through examination of the cultural values provided by Hofstede (n.d.). The data used in this study can be found at https://sites.google.com/site/cyberbehaviors/study-data.

1.1.1 Hofstede Cultural Values								Social Media Methods
Country	**PDI**	**IvC**	**FvM**	**UAI**	**LvS**	**IvR**	**Years**	
1	49	46	56	86	20	62	5	A
2	36	90	61	51	21	71	4	A
3	69	38	49	76	44	59	7	B
4	80	20	66	30	87	24	6	H
5	57	58	57	74	70	29	0+	Unlisted
6	78	8	63	67	N/A	N/A	3	B
7	35	67	66	65	83	40	1	A
8	77	48	56	40	51	26	4	Unlisted
9	58	41	43	59	14	40	5	A
10	13	54	47	81	38	N/A	4	Unlisted
11	81	30	69	82	24	97	0+	B
12	94	32	64	44	27	42	1	A
13	68	60	64	93	38	29	2	H
14	93	39	36	95	81	20	5	B
15	95	25	60	80	36	52	4	A
16	86	25	43	92	52	28	0+	H
17	60	18	39	85	100	29	4	B
18	80	35	52	60	30	N/A	6	A
19	58	17	45	69	93	49	7	B
20	66	37	45	85	49	49	4	B
21	35	89	66	35	51	69	3	H
22	40	91	62	46	26	68	6	B
23	81	12	73	76	16	100	2	B
24	70	20	40	30	57	35	4	H
Manipulators								
(median)	68	37	57	74	41	45.5	N/A	N/A
Control group (median)	69	30	42.5	60	35	47	N/A	N/A
HOFSTEDE	68	30	46	64	38	47	N/A	N/A

Table 1: List of countries that use social media to spread propaganda

The hypotheses tested are broken down into six sub-hypotheses for each dimension. Evaluation of the overall findings relies on 'or' processing of a truth table. A single positive or '1' entry in the truth table is sufficient to accept the alternative hypothesis.

Results

The results of the MWW test used to evaluate R1: H0, H1 can be viewed in Table 2; below; the main finding is shown graphically in Figure 1, below; and the corresponding truth table results are shown below in Table 3. Significant findings are shown in bold, and interesting findings are shown in italics. Tables 4 and 5, also below, contain the findings related to R2: H2, H3. Table 6 contains the findings for median value analysis to address R3. This section simply lists the results and interpretation; further analysis can be found in the discussion section, which follows.

Dimension	PDI	IvC	FvM	UAI	LvS	IvR
p-value	0.7944	0.4515	0.0078**	0.4978	0.235	0.940

Table 2:. R1: Social media methods and Hofstede cultural values
*** p < 0.05:** Conventional significance level, not corrected for multiple comparisons.
**** p ≤ 0.01:** Conventional level for statistical significance after adjusting for multiple comparisons.

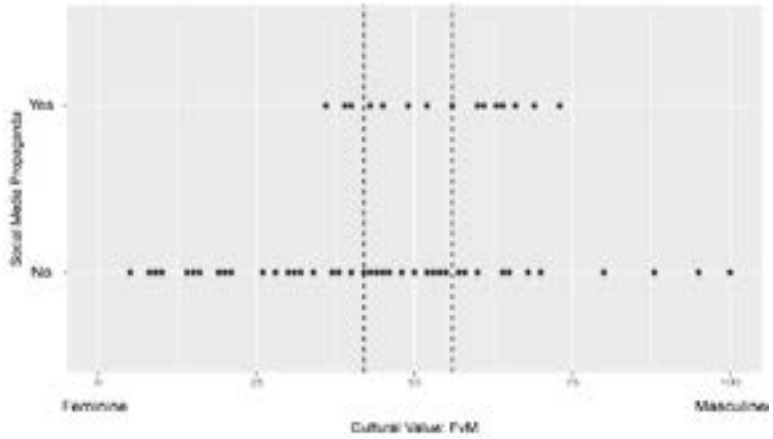


Figure 1: The x-axis depicts the feminine (0) to masculine (100) cultural value. The y-axis shows whether or not the country was a purveyor of social media propaganda (Yes is >50; No is <50). The dotted lines indicate the median values.

Dimension	PDI	IvC	FvM	UAI	LvS	IvR
T/F	0	0	1	0	0	0

Table 3: Truth table evaluation for H0, H1

Dimension	PDI	IvC	FvM	UAI	LvS	IvR
r	0.333	0.142	0.5714**	0.3333	-0.119	0.071

Table 4: Trended median values H2, H3

Dimension	PDI	IvC	FvM	UAI	LvS	IvR
T/F	1+	0	1+	1+	0	0

Table 5: Truth table evaluation for H2, H3

Dimension/ Delivery	PDI	IvC	FvM	UAI	LvS	IvR
Automated	58	41	60	60	27	47
Human	70	25	64	35	52	29
Both	73.5	24	47	79	49	49
HOFST-EDE	68	30	46	64	38	47

Table 6: Median values for delivery method groups

Discussion

Findings for each of the research questions provided evidence that implied a relationship between cultural values and the use of social media to disseminate propaganda. Masculine values were observed in response to each of the research questions. This finding suggests that, in addition to using social media for information seeking, masculine values may play a role in the use of social media for information shaping. A breakdown of the findings for each research question follows.

R1 involved examination of the propaganda purveyors who used social media. The difference between the disseminators and the non-disseminators in FvM dimension is similar to the observation of social engineering attackers by Sample et al. (2017), where self-identified attackers who deployed social engineering as an attack vector tended as a group to nationally possess masculine values compared to the non-attackers. This result suggests two things: first, that for attackers from countries that possess masculine cultural values, data may be considered as a tool that can be weaponised; and,second, that truthful data might be less valued in these competitive, aggressive societies than in societies with nurturing values. This may be because fake news stories often purport to provide information that is somehow unexpected or secret, and thus this information appears to offer the individual a competitive advantage, which involves the concept that knowledge is power. Another finding of interest in the data was the difference between the UAI values. Although not statistically significant, this result again supports the findings of the aforementioned study (Sample et al. 2017) in which the social engineering attackers were found to have significantly higher UAI values than their non-attacking counterparts. With regard to fake news, this could suggest that individuals are more likely to trust fake news sources, as to do so avoids uncertainty and lowers the perception of risk. In high UAI societies, propaganda may represent an additional method to assure mission success.

R2 involved examining the median values of the countries that use social media to disseminate propaganda over time. The trend line for masculine values over the 7-year period shows an increase. These two dimensions are singled out for two reasons: 1) the high FvM associates with information seeking and suggests that information shaping may also be a behaviour to associate with this dimension; and 2) the masculine trend, while significant, is not overtly masculine; rather the results are centred with a gradual increase reflecting an increasingly masculine trend. This finding suggests that, over time, these same countries that have historically exhibited information-seeking behaviours (Hofstede, Hofstede & Minkov 2010) are also increasingly information shaping. Sample and Karamanian (2014) noted the masculine tendency toward information dissemination with masculine countries' adoption rates with Twitter usage. However, the movement into information

shaping that appears, particularly with the use of Facebook, also reflects the evolution of the social media site.

The PDI value increase is relevant because successful propaganda requires consistency and repetition (Gambrill 2010), behaviours that can be more easily supported in a high PDI society where orders are passed down from leadership (Hofstede, Hofstede & Minkov 2010). In the low PDI countries, where messages emanate from multiple sources, the content is more likely to vary due to individual preferences, which results in message variation. Nonetheless, some message variation may be acceptable in order to make the message appear authentic and the spread appear organic. The core message must be consistent, and this consistency fits well in a society where permission for all actions is strongly controlled.

Conclusion

The Bradshaw and Howard (2017) report provided an inventory of state-sponsored propaganda producers and distributors that used social media to further their goals. Since propaganda is designed for cognitive hacking, the attackers' and targets' values and beliefs should be understood. These values and beliefs may provide the insights necessary to make the message believable or trustworthy to the intended recipient.

One aspect this study did not address is increasing the effectiveness of propaganda using psychological targeting. Matz et al. (2017) found that ads tailored to psychological traits, inferred personality characteristics using digital footprints from Facebook and other social media platforms, produced substantially more clicks and purchases than non-targeted ads. Others have disputed this finding as confounded due to targeted ads being higher in quality and more creative than non-targeted ads (Sharp, Danenberg & Bellman 2018). Another critique was that the targeting was confounded due to ad optimisation, since users were non-randomly assigned to the ad types (Eckles, Gordon & Johnson 2018). Nevertheless, the capabilities for targeting specific users and groups of users will only improve, and tailored propaganda may be more effective than mass messaging. Unlike propaganda in the physical world, with social media there is minimal cost and time for tailoring messages and maximising the effectiveness of the digital propaganda using experiments.

In the physical world, cultural values factor into believability and authenticity (Minkov 2013); therefore, cultural values should similarly impact the evaluation of messages in the virtual world. A central goal of propaganda is to persuade the target. Thus, the message sent in support of propaganda must resonate with the intended recipient's cultural values.

Once propaganda has been successfully identified, effective countermeasures should be deployed. Although psychologically countering recipients of misinformation is challenging, there are evidence-based recommendations for doing so (decrease the number of arguments supporting the misinformation, create scrutiny and counter-arguments, and provide detailed corrections to misinformation) (Chan et al. 2017). Another approach is to carefully construct countermeasures to manipulate the purveyors. These responses will require an understanding of online trust relationships and purveyors' values (cultural and psychological) in order to be effective. This study contributes to the overall process by focusing on the cultural values of countries that produce and disseminate propaganda. The cultural dimensions that were identified suggest that a direct response will

be needed, but the response will likely require detailed, consistent, inconspicuous, and culturally tailored responses.

Acknowledgements

This paper is an extended version of a conference paper with the same title that was presented at the 17th European Conference on Cyber Warfare and Security (ECCWS) in 2018. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Army Research Laboratory or the U.S. government. The U.S. Government is authorised to reproduce and distribute reprints for government purposes notwithstanding any copyright notation herein.

References

Aryan, S, Aryan, H & Halderman, JA 2013, *Internet censorship in Iran: A first look,* viewed 1 November 2017, <https://www.usenix.org/system/files/conference/foci13/foci13-aryan.pdf>.

Booth, R, Weaver, M, Hearn, A, Walker, S & Walker, S 2017, 'Russia used hundreds of fake accounts to tweet about Brexit, data shows', *The Guardian,* viewed 1 December 2017, <https://www.theguardian.com/world/2017/nov/14/how-400-russia-run-fake-accounts-posted-bogus-brexit-tweets>.

Bradshaw, S & Howard, PN 2017, 'Troops, trolls and troublemakers: A global inventory of organized social media manipulation', *Computational Propaganda Project,* Oxford University, Oxford, UK.

Chacon, MA 2006, *Course curriculum development for the future cyberwarrior,* Air Force Institute of Technology, Wright-Patterson AFB, Ohio School of Engineering and Management, viewed 2 November 2017 , <http://www.dtic.mil/citations/ADA453985.pdf>.

Chan, MS, Jones, CR, Jamieson, KH & Albarracín, D 2017, 'Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation', *Psychological Science,* viewed on January 10, 2018, <http://doi.org/10.1177/0956797617714579>.

Clayton, R, Murdoch, S & Watson, R 2006, 'Ignoring the great firewall of China', *Privacy enhancing technologies,* Springer Berlin, Heidelberg, DE, pp. 20-35.

Commin, G & Filiol, E 2015, 'Unrestricted warfare versus western traditional warfare: A comparative study', *Leading issues in cyber warfare and security,* ACPI, Reading, UK, pp. 73–88.

Comunello, F & Anzera, G 2012, 'Will the revolution be tweeted? A conceptual framework for understanding the social media and the Arab Spring', I*slam & Christian-Muslim Relations,* vol 23, no. 4, pp. 453-70, viewed 10 November 2017, <http://dx.doi.org/10.1080/09596410.2012712435>.

Crilley, K 2001, 'Information warfare: New battle fields, terrorists propaganda and the Internet", *Aslib Proceedings,* vol. 53, no. 7, pp. 250-64.

Cybenko, G, Giani, A & Thompson, P 2002, 'Cognitive hacking: A battle for the mind', *Computer,* vol. 35, no. 8, pp. 50-6.

Eckles D, Gordon BR, & Johnson, GA 2018, 'Field studies of psychologically targeted ads face threats to internal validity', *Proceedings of the National Academy of Sciences,* viewed 10 November 2018, <https://doi.org/10.1073/pnas.1805363115>.

Elmasry, MH, Auter, PJ & Peuchaud, SR 2014, 'Facebook across culture: A cross-cultural content analysis of Egyptian, Qatari, and American student Facebook pages", PhD dissertation, the American University in Cairo, EG.

Facebook 2019, 'Mission statement', viewed 6 February 2019, <https://www.facebook.com/pg/facebook/about/?ref=page_internal>.

—— n.d., 'Mission statement', viewed 6 February 2019, <https://www.facebook.com>.

'Fake news', Cambridge Dictionary, viewed 6 February 2019, <http://dictionary.cambridge.org>.

Fiske, ST & Taylor, SE (2013, *Social cognition: From brains to culture,* Sage, Los Angeles, CA, US.

Fox, J 2014, 'Why Twitter's mission statement matters', *Harvard Business Review,* viewed 2 November 2017, <https://hbr.org/2014/11/why-twitters-mission-statement-matters>.

Franke, U 2015, 'War by non-military means: Understanding Russian information warfare', viewed 15 March 2017 <http://www.foi.se/en/Top-menu/Pressroom/News/2015/War-by-Non-Military-means/>.

Friedman, B, Khan, Jr., PH & Howe, DC 2000, 'Trust online', *Communications of the ACM,* vol. 43, no. 12, pp. 34-40.

Gambrill, E 2010, 'Evidence-informed practice: Antidote to propaganda in the helping professions?', *Research on Social Work Practice,* vol. 20, no. 3, pp. 302-20.

Gillin, J 2017, 'Hillary Clinton not dead from heart attack, as fake news site claims', viewed 10 December 2017, <http://www.politifact.com/punditfact/statements/2017/aug/03/blog-posting/hillary-clinton-not-dead-heart-attack-fake-news-si/>.

Gottfried, J & Shearer, E 2016, 'News use across social media platforms 2016', *Pew Research Center Journalism & Media,* viewed 2 November 2017, <http://www.journalism.org/2016/05/26/news-use-across-social-media-platforms-2016/>.

Guess, CD 2004, 'Decision making in individualistic and collectivist cultures", *Online Readings in Psychology and Culture,* vol. 4, viewed 15 January 2012, <https://scholarworks.gvsu.edu/cgi/viewcontent.cgi?article=1032&context=orpc>.

Hemsley, GD & Marmurek, HH 1982, 'Person memory: The processing of consistent and inconsistent person information', *Personality and Social Psychology Bulletin,* vol. 8, no. 3, pp. 433-38.

Hofstede, G n.d., 'Dimension data matrix', viewed 6 February 2019, <http://geertHOFSTEDE.com/research-and-vsm/dimension-data-matrix/>.

Hofstede, G, Hofstede, GJ & Minkov, M 2010, *Cultures and organizations: Software of the mind,* McGraw-Hill, New York, NY, US.

Hollander, M, Wolfe, DA & Chicken, E 2013, *Nonparametric statistical methods,* John Wiley & Sons, Hoboken, NJ, US.

Howard, PN & Hussain, M 2011, 'The role of digital media', *Journal of Democracy,* vol. 22 no. 3, pp. 35-48.

Jones, HS & Moncur, W 2018 'The role of psychology in understanding online trust', *Psychological and behavioral examinations in cyber security,* J McAlaney, LA Frumkin & A Benson (eds), IGI Global, Hershey, PA, US.

Kahneman, D, Slovic, P & Tversky, A 1982, *Judgment under uncertainty: Heuristics and biases,* Cambridge University Press, Cambridge, UK.

Lee, CE & Kent, JL 2017, 'Facebook says Russian-backed election content reached 126 million Americans', *NBC Nightly News,* viewed 2 December 2017, <https://www.nbcnews.com/news/us-news/russian-backed-election-content-reached-126-million-americans-Facebook-says-n815791>.

Lynn, WJ 2010, 'Defending a new domain: The Pentagon's cyberstrategy', *Foreign Affairs,* pp. 97-108, viewed 1 November 2017, <https://www.foreignaffairs.com/articles/united-states/2010-09-01/defending-new-domain>.

Matz, SC, Kosinski, M, Nave, G, & Stillwell, DJ 2017, 'Psychological targeting as an effective approach to digital mass persuasion', *Proceedings of the National Academy of Sciences,* vol. 114, no 48, pp. 12714-9.

Meza, S 2017, "'Fake news' named word of the year", *Newsweek,* viewed 1 November 2017, <http://www.newsweek.com/fake-news-word-year-collins-dictionary-699740>.

Minkov, M 2013, *Cross-Cultural analysis,* Sage Publications, Thousand Oaks, CA, US.

Neff, G & Nagy, P 2016, 'Automation, algorithms and politics; Talking to bots: Symbiotic agency and the case of Tay', *International Journal of Communications,* vol. 10, p. 17.
Nisbett, R 2010, *The geography of thought: How Asians and Westerners think differently...and why,* Simon and Schuster, New York, NY, US.

Nosek, BA, Hawkins, CB & Frazier, RS 2011, 'Implicit social cognition: From measures to mechanisms", *Trends in Cognitive Sciences,* vol. 15, no. 4, pp. 152–59, viewed 10 December 2017, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3073696/>.

'Propaganda', *Cambridge Dictionary,* viewed 6 February 2019, <http://dictionary.cambridge.org>.

Sample, C, Cowley, J & Hutchinson, S 2017, 'Cultural exploration of attack vector preferences for self-identified attackers', *11th IEEE international conference on Research Challenges in Information Science,* 10-12 May, Brighton, UK, pp. 305-14.

Sample, C, Hutchinson, S, Karamanian, A & Maple, C 2017, 'Cultural observations on social engineering victims', *Proceedings of the 16th European Conference on Cyber Warfare and Security,* Dublin, IE, pp. 391-401.

Sample, C & Karamanian A 2014, 'Application of HOFSTEDE's cultural dimensions in social networking', *Proceedings of the 1st European Conference on Social Media 2014: ECSM 2014,* Academic Conferences Limited, pp. 466-73.

Sharp, B, Danenberg, N & Bellman S 2018, 'Psychological targeting', *Proceedings of the National Academy of Sciences,* vol. 115, no. 34, pp. E7890.

Shultz, Jr., RH 1989, 'Political strategies for revolutionary war', *Political warfare and psychological operations: Rethinking the US approach,* National Defense University, Washington, DC, US, pp. 111-38.

Spina, RR, Ji, LJ, Guo, T, Zhang, Z, Li, Y & Fabrigar, L 2010, 'Cultural differences in the representativeness heuristic: Expecting a correspondence in magnitude between cause and effect', P*ersonality and Social Psychology Bulletin,* vol. 36, no. 5, pp. 583-97, doi:10.1177/0146167210368278.

Stinson, E & Mitchell, J 2007, 'Characterizing bots' remote control behavior', *DIMVA,* pp. 45-64, viewed 3 May 2018, <https://pdfs.semanticscholar.org/9d11/285ab5cb042f59f1ff64f2fecc4f52acac4b.pdf>, DOI:10.1007/978-3-540-73614-1_6.

Szfranski, R 1997, 'A theory of information warfare: Preparing for 2020, Air University Maxwell Airforce Base', viewed on March 10, 2017 , <http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA328193>.

Tanis, M & Postmes, T 2005, 'A social identity approach to trust: Interpersonal perception, group membership and trusting behaviour', *European Journal of Social Psychology,* vol. 35, no. 3, pp. 413-24.

Taylor, PM 2013, *Munitions of the mind: A history of propaganda,* Manchester University Press, Manchester, UK, viewed 1 November 2017, <http://elastic.org/~fche/mirrors/www.cryptome.org/2013/01/aaron-swartz/Mind-Munitions.pdf>.

Travers, M, Van Boven, L & Judd, C 2014, 'The secrecy heuristic: Inferring quality from secrecy in foreign policy contexts', *Political Psychology,* vol. 35, no. 1, pp. 97-111, doi:10.1111/pops.12042.

Implications of Privacy & Security Research for the Upcoming Battlefield of Things

L Fritsch, S Fischer-Hübner

[1]Department of Mathematics and Computer Science
Karlstad University
Karlstad, Sweden

E-Mail: Lothar.Fritsch@KAU.se; Simone.Fischer-Huebner@KAU.se

Abstract: *This article presents the results of a trend-scouting study on the applicability of contemporary information privacy and information security research in future defence scenarios in a 25-year-horizon. The authors sketch the expected digital warfare and defence environment as a 'Battlefield of Things' in which connected objects, connected soldiers, and automated and autonomous sensing and acting systems are core elements. Based on this scenario, the authors discuss current research in information security and information privacy and their relevance and applicability for the future scenario.*

Keywords: *Internet of Things, Autonomous Systems, Digital Warfare, Transfer of Research, Information Privacy, Information Security, Trend Scouting, Cyberwar, Cybersecurity, Weaponization of Smart Systems*

Introduction

This study is the result of a trend-scouting project which was carried out by Karlstad University's research group, PriSec, and which was commissioned by the Swedish Defence Research Institute (Totalförsvarets forskningsinstitut, FOI). Its goal was the projection of contemporary research in information privacy and information security into a 25-year-future defence context. The article relates published contemporary research to the forecast of the 'Battlefield of Things'.

Methodology

As a first step, the researchers developed a future defence scenario with a 25-year horizon in an attempt to predict the developments of the years 2017 through 2042. With respect to the recent decades' development, this was the most challenging task. The researchers invited IT researchers to brainstorm about the IT environment and degree of digitisation of society and, in particular, its defence functions for two decades into the future. The results were these four main assumptions:

1. There will be a strong and integrated digitisation of defence and societal security activities;
2. Military equipment will be digitised and interconnected with networks and physical infrastructure;
3. 'Smart military devices' will be configurable to particular situations and contexts, which will turn the software they run into their 'ammunition';

4. The weaponization of civilian 'smart technology' may be part of military tactics.

Thus, the researchers coined the term 'Battlefield of Things' (BoT) as a descriptive name for this scenario.

The next step was a first-person review of contemporary research activities in a research group to determine their relevance in the scenario. The results were then interpreted into the future scenario. In a final step, the researchers invited the active researchers for brainstorming and feedback meetings to adapt and to verify research application to the scenario. The results were documented in KAU Technical report LOF2017-4 (Fritsch et al. 2017) and were presented to the Swedish defence research institute, FOI.

The remainder of this article is structured as follows. In the next section, the future digital defence environment is sketched out. Then, the relevance of current research in information security and privacy for the scenario is reviewed. The review is divided into thematic sections that each relate to a research area. Following the review, a summary and a list of background references are provided.

The Future Cyber-Physical Defence Environment

In a ten- to twenty-five-year perspective, military reconnaissance and tactical operations will rely largely on automated systems that coordinate and inform each other through data communication, which is controlled by human-staffed command centres. This part of the article describes the researchers' assumptions and expectations regarding the use of information technology in the 'Battlefield of Things'. Strategic and tactic activities are carried out with networked and autonomous digital agents, cyber-physical systems, and 'connected' human soldiers. Many devices will be spread into the field before or during conflicts, and will be awaiting activation and mission-specific configuration: smart cyber-physical sensors or weapons.

One important aspect of the future BoT will be the weaponization of civilian or dual-use infrastructure. Such on-demand cyberwar infrastructure will be based on own forces' IoT and will, in addition, include opponent systems as well as third-party infrastructure. To complicate matters further, parts of such infrastructure will be operated on global infrastructures; such operations will be beyond the control of national security organisations and other third parties. In **Table 1,** below, the future BoT infrastructure is classified into own/opponent/third-party infrastructure made for military, civilian, or dual-use purposes. The latter category includes national and overnational critical infrastructures that are regulated or supervised by governments.

IoT weaponization level	**Own**	**Opponent**	**Third party**
Military	Full	Low	Medium
Civilian	Medium/Full	Medium	Medium
Dual-use			
(controlled infrastructure)	Full	Low	Medium

Table 1: Weaponization level for infrastructures classified according to ownership and degree of militarisation

The assessment of the weaponization level expresses the expected level of control and level of reliability over the respective infrastructure as a component of the BoT. In this analysis, the researchers focused on own infrastructure that is under supervision of security regulation or security organisations within the own perimeter of influence.
In the 'Battlefield of Things', devices will have many properties that will be controlled by the defence organisations:

- Sensing, communication, coordination, and action capabilities;
- Deployment of weapons;
- Ability to 'update' or download apps to change functionality;
- Devices that will 'hibernate' in the field, at risk of discovery, access, manipulation, and re-engineering;
- Devices that will be re-configured for specific missions in very short time margins.

Moreover,

- Human beings will need protection as part of the cyber-physical battle infrastructure; and
- Autonomous digital decisions taken will need to be verified and audited.

In this context, digital components from the civilian or industrial infrastructure—such as power grid controllers, remote-controlled flood gates, smart cards, smart planes, delivery drones, smart cars, and other devices—can become 'weaponised' through software upgrades. The researchers presume, therefore, that such devices may in certain defence situations traverse the border between civilian and military tasks where the weapon capability is strongly determined by the uploaded tactical software.

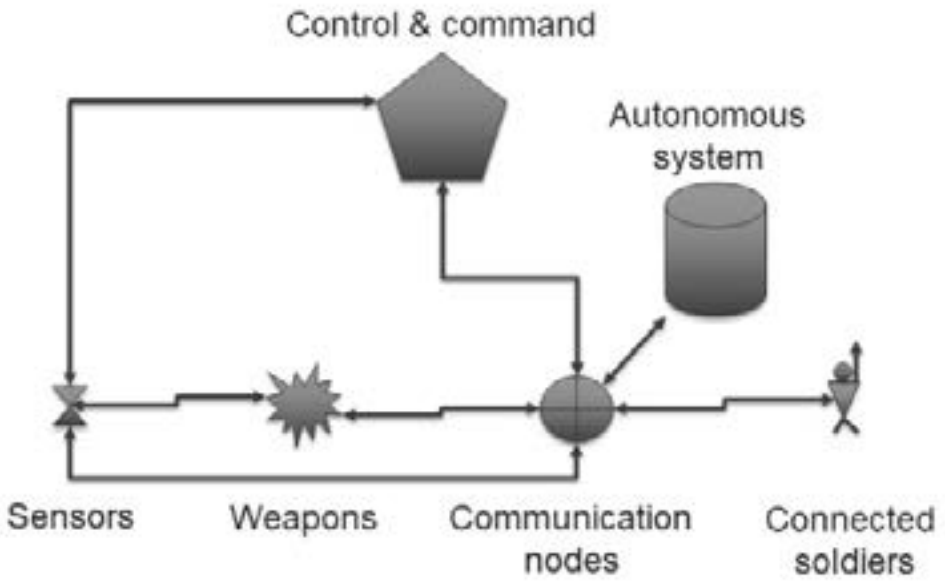


Figure 1: 'Battlefield of Things' scenario

In Figure 1, above, a simplified model of the scenario is shown. A control and command infrastructure will ensure communication with all kinds of devices and with connected soldiers. Devices will be installed in the field long before they get used—and they will get dropped on demand. The researchers foresee sensing devices; acting devices such as weapons with or without autonomy, alone or in swarms; routing devices that ensure connectivity; and connected soldiers.

Algorithms become weapons and ammunition

Since the 'things' will be installed long before they get used, they will provide an attack surface for enemy intelligence, sabotage, and adversarial take-over. The researchers, therefore, presume that algorithms, data, and calibration of these devices will be military secrets. To prevent re-engineering, they will not be loaded into the devices until they are needed Devices, sensing behaviour, communication patterns, and actions will be configured on demand shortly before tactical situations. Calibration or installation of autonomous or swarm capabilities will be loaded on demand, based, for example, on intelligence or will be controlled by connected soldiers. In essence, 'apps' will significantly define and change the nature of an installed device, whether it is a weapon system, a sensor, or a controller for a power network or a smart car.

Scenario implications

The 'Battlefield of Things' has severe implications. There will be a need to secretly and reliably communicate in all kinds of situations—or to have sufficient autonomy of the things. There will be a need to update, configure, and activate groups of devices quickly; and they must be free of errors. Devices will deploy autonomous actions based on IT-borne decisions in a local cluster of devices and connected soldiers. Since many of those actions will be based on previously or recently installed 'apps' and since opponent interference with devices' software is likely to occur, as discussed in the case of autonomous weapons capabilities by Scharre (2016), monitoring and audit/investigation of the status of devices or how they made their decisions is crucial. In addition, it is presumed that what is seen today in cyberattacks against an open Internet infrastructure will be the future art of war against military, electronic devices, and connected soldiers. Sabotage, functional change, takeover, and manufactured sensing results will be the result of successful attacks. The weaponization of dual-use and civilian infrastructure may, in addition, cause severe collateral damage from cyber-military activities. Current concepts of acceptable collateral damage have not yet been extended into cyberwar scenarios and will, therefore, cause major uncertainty for decision-makers, as discussed in recent debate (Romanosky & Goldman 2016). Parts of the technology forecast in this article include technologies that have the potential to reduce cyber collateral damage.

Applications of Privacy and Security Technologies

In this section, the researchers present and discuss current research activities, results, and trends from information privacy and information security by the PriSec research group. Their relevance in long-term digital defence is discussed, including literature and background materials. In the five consecutive sections, details are added to the 'Battlefield of Things' scenario, and the ways in which contemporary research will influence future defence are discussed. The projection is restricted to the research areas covered by the research group by demand of the survey sponsor.

Secure, unobservable communication with all parts of the digital defence infrastructure
This section elaborates the relevance and use of technologies for anonymous, unlinkable, and unobservable communication in the context of connected objects and services in future defence and its role in camouflaging the location, the role, and the activity patterns of objects and services.

Confidential communication is an essential asset in a defence scenario that strongly relies on connected objects. Direct digital communication using network protocols may reveal mission-critical information, such as location of objects, degree of activity, deployment, among others. The use of anonymous communication protocols will be an essential asset in a future defence scenario. Individual connected objects, connected soldiers, or autonomous systems can be addressed, controlled, and deployed without revealing their network location or their relationship to each other and to the command-and-control infrastructure. Tor (2018) can, in addition, be used to hide and protect critical digital infrastructure using onion services. By concealing physical and networking location, attackers will need to spend considerable intelligence resources to locate and to attack such digital services.

Tor is a low-latency anonymity network with millions of daily users that can be used to browse the Internet anonymously, to host end-to-end secure and potentially anonymous services ('onion services'), and to circumvent censorship (Dingledine, Mathewson & Syverson 2004; Tor 2018). The design of Tor favours low-latency—to support use-cases such as browsing—at the cost of being vulnerable to powerful Internet-wide adversaries. This design decision has led to wide adoption; but at the same time, it has also led to a plethora of attack vectors.

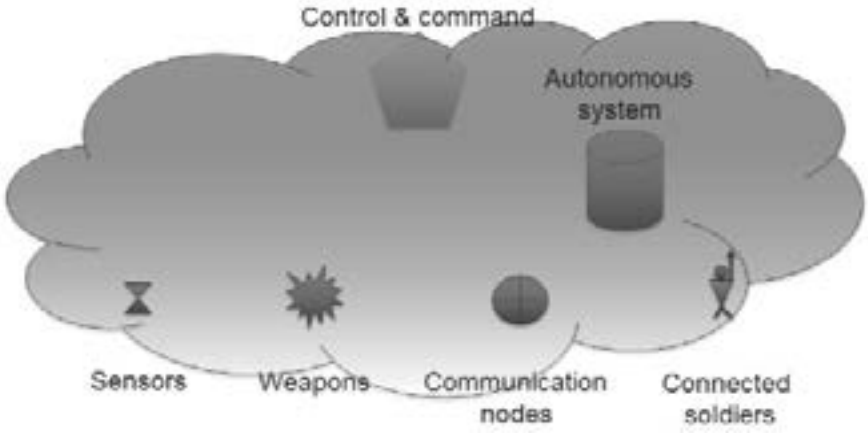


Figure 2: Unobservable communication

In the 'Battlefield of Things' scenario, robust anonymous communication and camouflage of defence-critical services will be a major asset. This is shown in Figure 2, above. Such communication will protect the location, activity levels, tactics, and the interconnection of defence units from digital observation. The societal infrastructure will benefit greatly from the availability of a Tor-like network. In case of cyberattacks, parts of the critical infrastructure could be moved 'out of sight' to onion services. Attacks on digital payment, healthcare, or cyber-physical control systems, such as the power grid, would be much more difficult to carry out since the target computers and connected objects that implement those services would be hidden away in an anonymous network.

In recent research, the PriSec group has researched improved traffic analysis resistance for censorship circumvention, as well as novel traffic analysis attacks on Tor. These censorship circumvention efforts have been focused around the Great Firewall of China (Winter & Lindskog 2012) and the construction of a polymorphic network protocol to make traffic classification of Tor error-prone (Winter, Pulls & Fuss. 2013). Design decisions in this protocol greatly influenced the design of Obfs4 (Yawning-Angel 2015), the default censorship circumvention technique shipped in Tor at the time of writing. The authors have also designed traffic analysis attacks on Tor, showing how a local attacker (such as an ISP) can correlate traffic patterns from a target victim with DNS traffic exiting the Tor network to determine with high precision which websites the victim is visiting over Tor. Furthermore, this line of research has uncovered design and implementation flaws in Tor and excessive reliance on Google's DNS infrastructure by Tor-network operators and changed how Tor handles DNS (Greschbach et al. 2017).

Secure and robust consensus-finding for logging of activities

Decision-making in an autonomous system should not only be robust against enemy influence, but it should also be verifiable in case of anomalies or unexpected behaviour. Therefore, there is a large need for technologies and methods that will enable audit, inspection, and reproduction of decisions and actions taken by autonomous defence systems.

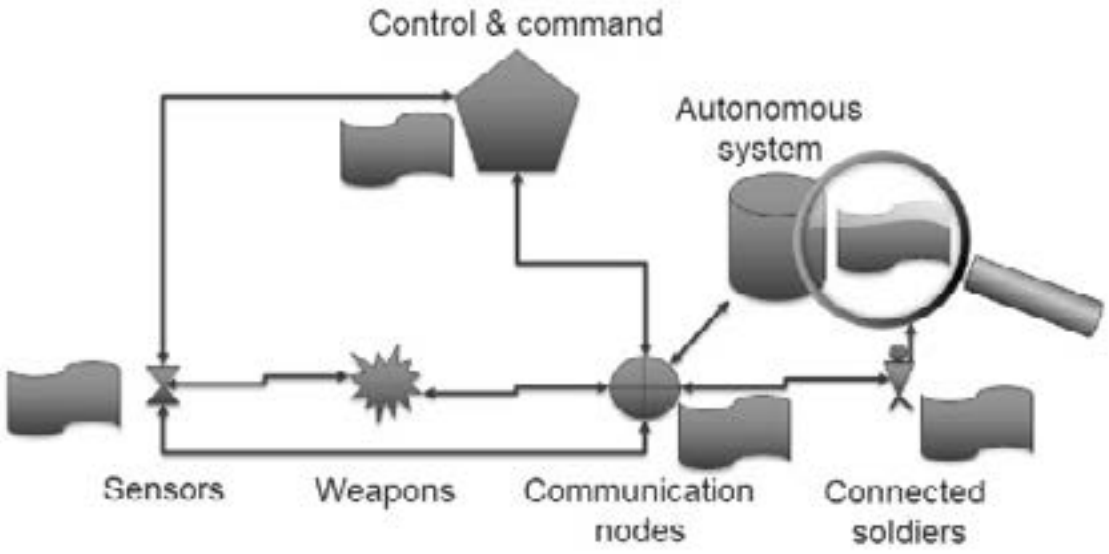


Figure 3: Secure logging and consensus

The areas of application for this technology are all fields of data authentication, such as software updates, a database of authentic sensor data, and tamper-evident operational logs. Logging will be installed on each device, as shown in **Figure 3**, above. Customised authenticated data structures have a range of advantages over the particular combination of technologies commonly referred to as blockchain. These advantages include simplicity, throughput, efficient non-membership proofs, and reliance on more conservative trust assumptions. Certificate Transparency will change the certificate authority ecosystem, and the network-layer gossiping can provide herd immunity for entire network segments by detecting targeted attacks without relying on protocols' gossiping about

potentially sensitive data, such as Signed Certificate Timestamps (SCT) and Signed Tree Heads (STH) outside the protected network. Thus, Certificate Transparency is suitable for protected networks that still retain some limited access to the public Internet.

Data authenticity is an increasingly vital societal concern, and being able to collectively maintain a database without the need for central trust is, therefore, highly relevant. Similarly, centralised systems without adequate protection are single points of failure. Trust in sensor measurements as well as coordinated implementation of operations are critical for defence and civil security. Ensuring and documenting system consensus, algorithmic accountability, and verification of correct function of components will be important features of connected objects and their control systems. Secure logging technology may help investigate anomalies while preserving operation confidentiality.

Authenticated data structures—ensuring that a data structure, such as a database, has not been tampered with—have a long history (Merkle 1990). Notably, research from the cryptology community is at the centre of the hyped 'blockchain' technology (Narayanan & Clark 2017). Further, with less fanfare, authenticated data structures are about to change the Web's certificate authority ecosystem with the deployment and mandatory use of Certificate Transparency (Laurie, Langley & Kasper 2013) in popular Web-browsers starting April 2018 with Google Chrome (message from Ryan Sleevi on Google Chromium forum, 21 April 2017).

The PriSec research group has researched novel authenticated data structure designs and secure logging applications and has ongoing work focused on Certificate Transparency and gossiping. These data structure designs have been focused on tailoring the data structures for specific applications that require efficient non-membership proofs, including outsourceable logging (Pulls & Peeters 2015) and the certificate authority ecosystem (Dahlberg, Pulls & Peeters 2016). The secure logging-related designs include distributed settings (Pulls, Peeters & Wouters 2013) and strong security and privacy properties supporting Transparency-Enhancing Tools (Peeters & Pulls 2016) (particularly section 0). Finally, ongoing research relates to the Certificate Transparency ecosystem on efficient monitoring solutions (Dahlberg & Pulls 2017) using modern programmable network planes for scalable and practical gossip (related to the well-known consensus problem in distributed systems) below the application layer, including a P4 implementation.

Reliable and error-free configuration and management of large, complex communication infrastructures

The authors foresee major issues for configuration, access control, communication configuration, and key management for the digital defence of the coming decades. Human error and hard-to-manage levels of complexity will cause major configuration errors (Wool 2010; Wool 2004; Iwaya et al. 2016). This section discusses research on the reduction of cognitive load through software tools and predicts their future relevance in digital defence.

The authors envision the dynamic administration of device configurations. This will govern the communication as well as the behaviour of connected objects and cyber-physical systems on the 'Smart battlefield' or in 'smart intelligence gathering' as well as in national and tactical cyber defence. In particular, in response situations, when decisions about deployment and re-configuration

of connected equipment are being implemented, both stress and the shortage of time or resources for verification will make the reprogramming and commissioning of digital defences or weapons a task vulnerable to errors (Fritsch & Fuglerud 2010). Usable interfaces, coupled with metrics and verification support, will reduce error rates and will ultimately lead to more operations that are reliable, for example, in reactions to cyberattacks.

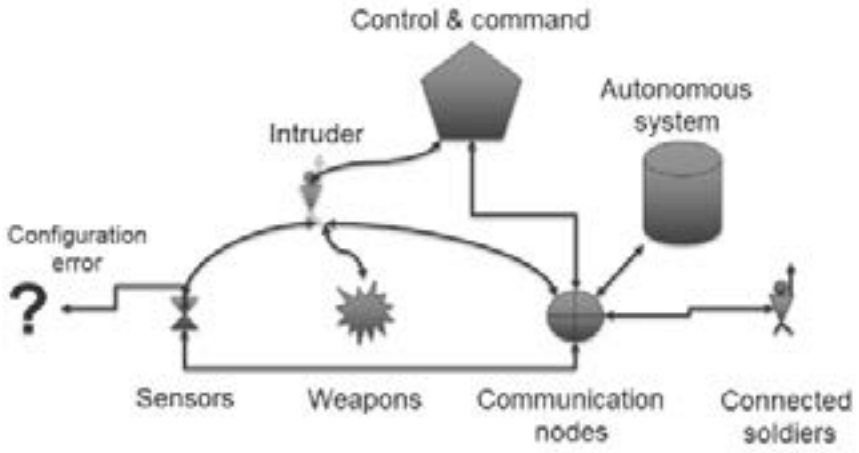


Figure 4: Reliable and error-free configuration

In 20 years, defence and society will strongly depend on connectivity to digital services. Proper configurations of the base infrastructure, of connected objects and cyber-physical systems, will be a cornerstone of society. In a hybrid physical and digital attack scenario, both the defence infrastructure as well as the societal infrastructure will need to be compartmentalised, reconnected, firewalled, disconnected, and managed in response to threats. Both defence personnel as well as commercial, private, and governmental actors will need to implement decisions about re-configuring the infrastructure with the least possible error rates.

Configuration and management of a large, complex communication infrastructure that contains the systems and configuration paths will be a major challenge. For changing tactical situations, all nodes of such a system will get re-configured continuously, based on cyber risk, battlefield events, or actual deployment of systems. However, setting up correct rules (for example, for firewalls and routers) is already a very complex task that produces many configuration errors today. This leads to non-functioning or compromised infrastructure, as shown in **Figure 4,** above. The authors foresee major issues for the device configuration, access control, communication, configuration, and key management for the digital defence of the coming decades. Human error and hard-to-manage levels of complexity will cause major errors.

In the PriSec research group, one research goal is to identify firewall usability gaps and to mitigate them (Voronkov 2017). The principal aim of the research is the reduction of cognitive effort when designing firewall rule blocks (Voronkov et al. 2015). The authors noted that complex firewalling rule sets contain many errors. Editing rulesets is often a manual task. It is very easy to commit errors under such conditions. This research aims at improving the usability of the user interface for firewall configurations. The authors investigated challenges of configuration editing, defined

usability metrics for firewall configuration management, and developed alternative user interfaces for firewall rule management that will be evaluated with the developed measurements.

The research can be applied to firewalls and other configuration rules, such as routing tables, rules for organising collaboration of smart things, the administration of access control rules, and others. Each application needs specific metrics developed, while the user interface for the administration of the respective rules will require certain research about cognitive understanding of the application area.

Exercise privacy rights and protection of individuals' data from adverse access or attack

Cyberattacks will not only target military areas, but will also continue to target civilian infrastructure. Personal information can be misused to blackmail or to demotivate personnel or to create unrest in the population. Network monitoring may, in addition, affect privacy (Fritsch 2018).

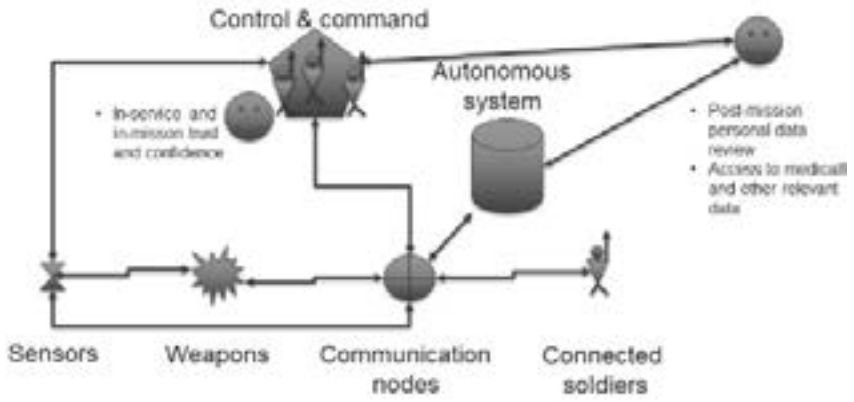


Figure 5: Usable transparency technology support for privacy awareness

The authors foresee an increase in the importance of protecting, securing, separating, and transforming personal data of soldiers, including health data, with privacy technology for the protection of defence staff and their civil environment. Management of personal data with respect to secrecy will be performed at various stages of engagement and careers, as shown in **Figure 5,** above. The research area of cognitively aligned user interfaces for personal data management supports the management of staff personal data when defence staff are aligned with technical systems; get registered for, monitored on, and decommissioned from personalised cyber-physical systems. The authors discuss the relevance of transparency technologies in digital defence as a cornerstone of democratic, law-abiding defence organisations—as part of privacy and confidentiality management.

Usable transparency and intervenability tools are important technologies for helping to enforce legal privacy requirements pursuant to the European Parliament and Council of the European Union Regulation (EU) 2016/679 (GDPR) for enhancing data subject controls and accountability of the data controller. They can be implemented or used as self-defence privacy tools by the data subjects and/or offered by the data controllers, pursuant to Art. 25 of the General Data Protection

Regulation (GDPR), whereby data controllers are legally obliged to implement technical measures following the principle of Data Protection by Design.

Transparency of personal data processing is an important privacy principle. Pursuant to Art. 5 (1) GDPR, "personal data shall be processed lawfully, fairly and in a transparent manner in relation to the data subject". Transparency is a means for meeting with information asymmetry between data subjects and data controllers and enables data subjects to 'intervene' by exercising their rights to data correction and deletion as well as the right to withdraw consent or to object. It can, therefore, play an important role for establishing user trust in applications.

The concept of transparency comprises both `ex ante transparency' and `ex post transparency'. Ex ante transparency signals the intended data collection, processing, and disclosure, and thus enables the anticipation of consequences before data are disclosed—for example, with the help of privacy policy statements. Ex post transparency provides insight about what data were collected, processed, or disclosed by whom and to whom, and should be particularly informative about consequences if data have already been revealed. Transparency Enhancing Tools (TETs) can help individuals to exercise their right for transparency, and subsequently for intervenability, by technological means. The authors have conducted research on usable ex ante and ex post TETs within the scope of several EU projects, including the FP7 projects (PrimeLife and A4Cloud) and the H2020 projects (Privacy&Us, PRISMACLOUD, and CREDENTIAL) (Angulo et al. 2015; Fischer-Hübner et al. 2016; Angulo et al. 2012; Karegar et al. 2017; Karegar et al. 2018). A recent literature survey on usable ex post TETs was published in Murmann & Fischer-Hübner (2017).

Privacy invasions at the work place may have a negative impact on self-esteem, creativity, and performance, especially in a military context that uses privacy-infringing information and communication technologies, including sensor technologies for tracking users in a non-transparent manner. As concluded by Sigholm & Andersson (2011), "(1) extensive use of emerging military ICTs [Information and Communication Technologies] gathering personal data, without proper PETs [Privacy Enhancing Technologies] employment, will lead to soldiers' privacy being violated and (2) these violations will result in an observable performance drop" (p. 266).

Pursuant to Article 2 (2), the GDPR is actually not applicable to activities regarding national and common security. However, for non-security related data processing, such as the soldier's medical records that are collected for medical checks and diagnoses, the soldiers have transparency and intervenability rights pursuant to the GDPR or to other complementary laws, such as the Swedish Data Patient Act.

Moreover, even though privacy and particularly transparency rights may need to be restricted in the defence sector due to overriding secrecy interests for activities regarding national security, to what extent the data subjects should still be informed and have control over their personal spheres needs to be considered. Research on Transparency Enhancing Technology (TETs) for soldiers should investigate the level of granularity with which personal user data can be made transparent to soldiers concerned at what time, so that the right trade-off will be achieved between secrecy requirements of military data and privacy interests and rights of the soldiers. This granularity of transparency information provided by ex post TETs could become more fine-grained over the

time with which the requirement of keeping certain tactical information secret is decreasing. For instance, right after data collection, the soldier might only be informed about what types of data have been collected about him or her and over what period, while as soon as the data is not security-critical any longer, the soldier could be informed about more or all details of the personal data that were collected about him or her.

Monitoring and assessment of connected object behaviour in the field

Digital assets will run software of various complexity levels. The authors expect that such infrastructure will run apps that get either installed or decrypted and activated when they are needed. The authors discuss how current research on smartphone app behaviour will contribute to a security infrastructure for code integrity, for monitoring of actual running code for its behaviour, and for the creation of anomaly warnings along with analysis tools.

In future defence infrastructures, apps will likely be loaded onto networked sensors, acting components, weapons, and autonomous defence systems 'on demand' to accommodate particular contexts or missions. Such apps will be provided from various industrial and government players. They will get distributed as needed or will be pre-installed on equipment. Equipment will be placed—in active or inactive mode—in the field for longer periods without supervision, possibly exposed to manipulation efforts. The authors' research on app privileges will help with monitoring actual app behaviour on critical systems.

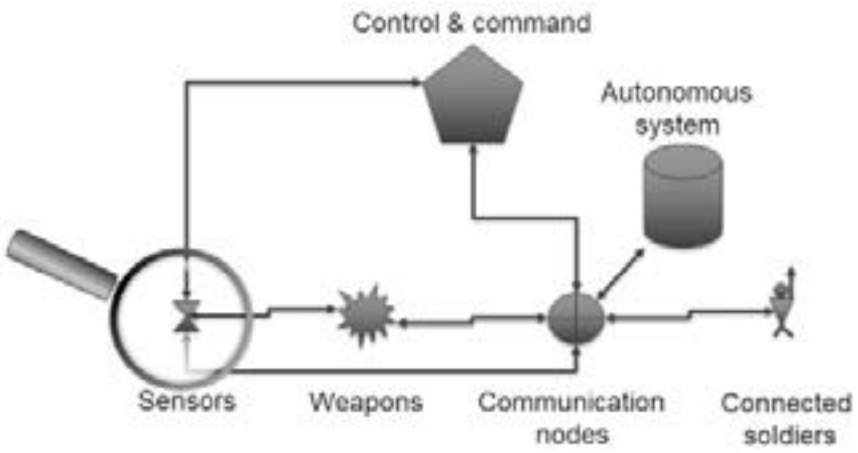


Figure 6: On-device monitoring of code behaviour

In both pre-mission settings and in post-mission analysis, the actual behaviour of the code installed on smart and autonomous equipment will need to be reviewed, as shown in Figure 6, above. The authors expect that autonomous systems' decisions will need to get analysed after major events. Code execution monitoring, operating system interactions, and API calls (Wei et al. 2012) will be a key data source for such analysis of both code quality and actual code behaviour (Paintsil & Fritsch 2013) For cases in which inactive equipment is being outplaced, its actual code should get inspected before the devices are placed into operation. The same holds for devices that are updated, that receive on-demand apps for specific missions, or that get reprogrammed. Pre-installation review and post-installation monitoring of such apps will be necessary to ensure reliable operation

(Marforio et al. 2012). Using monitoring anchored on the devices will enable data collection. Swift statistical analysis and visualisation will reveal anomalies, code manipulation, and unexpected behaviour of software on devices in the field.

Digital assets in the future will run software of various complexity levels. The authors expect the arrival of a few verified, robust operating systems for high-security applications of sensors and for the Battlefield of Things, together with respective communication devices. They also expect that actual task-solving software will be dynamically installed and/or updated on such infrastructure since a large sensor and robot weapon network that resides unguarded out in the field will certainly be victim to theft and reverse engineering by enemy intelligence. It is, therefore, expected that such infrastructure will run apps that either get installed or decrypted and activated when they are needed. A security infrastructure for code integrity, monitoring of actual running code for its behaviour, and the creation of anomaly warnings along with analysis tools that use data from app behaviour will be an important maintenance asset, both in a strategic and in a tactical perspective (Kelley et al. 2012).

The authors are currently investigating how Android apps use their access permissions to sensitive and personal data. Apps, upon installation, are configured with access permissions they can permanently use. For the time being, no summary over consumed sensitive data or over the actual use of those access permissions is available to Android device users. The KAUdroid project (Momen et al. 2017) aims at collecting permissions use data, at the establishment of risk thresholds—specifically for personal privacy—and at the provisioning of a graphical user interface that effectively warns of risk thresholds' being reached. The project has, so far, developed a monitoring mechanism installed on Android devices and a data collection service that is being used for large-scale data collection. Analysis of preliminary data has shown apps that wake up when devices are inactive to exercise their microphone access permissions. It has also shown that users can easily get identified through data access (Fritsch & Momen 2017). The ultimate goal of the project is the user-centric provision of a mechanism that will notice risks from apps and that will enable device users to manage and mitigate the risks.

Summary

In summary, the authors expect defence operations to strongly rely on connected objects, partially or highly autonomous systems, and their underlying communications infrastructure. This article has sketched a future in which soldiers are connected to the information infrastructure. Dynamic configuration and just-in-time deployment of operational code and configurations in strategic as well as in tactical situations are foreseen. The authors have identified a number of relevant research areas that will support the confidentiality, the robustness, the availability, the usability, and the reliability of such infrastructures. Their relevance in long-term digital defence was discussed, including literature and background materials. The relevant research results that are of interest in a networked, cyber-physical, and human defence infrastructure are as follows:

- Secure and anonymous communication with all parts of the digital defence infrastructure;
- Authenticated data structures without centralised trust;
- Reliable and error-free configuration and management of large, complex communication infrastructures;

- Exercise of privacy rights, privacy self-management, and protection of individuals' personal data from adverse access or attack;
- Monitoring and assessment of connected object behaviour in the field.

The authors expect worthwhile research results from the transformation of today's research activities into the future defence scenario.

Conclusions

This research indicates that a broad range of security and privacy research topics are highly relevant for a future cyber-physical defence environment. The proper security of smart technology in the field as well as the military access to civilian and dual-use infrastructure are essential on the Battlefield of Things. Reliable and confidential communication control over actual code and code behaviour as well as proper active and reactive re-configuration of network security features will reduce malfunction, human error, and cyber collateral damage. In addition to cyber security, special attention needs also to be paid to privacy in military environments for the following reasons. First, anonymous communication and location privacy for the soldiers are critical for keeping military operations secret from the opponents. Secondly, protecting privacy as a fundamental right of soldiers will foster trust, job satisfaction, and improved performance. Hence, current and future research results—not only from the cyber security but also from the privacy technology community—should be followed closely.

Acknowledgement

The work leading to this article was sponsored by Sweden's Defence Research Institute, FOI.

References

Angulo, J, Fischer-Hübner, S, Pulls, T & Wästlund, E 2015 'Usable transparency with the data track: A tool for visualizing data disclosures', *Proceedings of the 33rd annual ACM Conference: Extended abstracts on human factors in computing systems,* ACM, pp. 1803-8.

Angulo, J, Fischer-Hübner, S, Wästlund, E & Pulls, T 2012, 'Towards usable privacy policy display and management', *Information Management & Computer Security,* vol. 20, pp. 4-17.

Dahlberg, R & Pulls, T 2017, 'Verifiable light-weight monitoring for certificate transparency logs', *arXiv preprint arXiv:1711.03952,* Karlstad University, Karlstad, SE.

Dahlberg, R, Pulls, T & Peeters, R 2016, 'Efficient sparse merkle trees', *Proceedings of the 21st Nordic Conference on Secure IT Systems (NordSEC),* BB Brumley & J Röning (eds), Oulu, FI, Springer, pp. 199-215.

Dingledine, R, Mathewson, N & Syverson, P 2004, 'Tor: The second-generation onion router', *Proceedings of the 13th USENIX Security Symposium,* 9-13 August, San Diego, CA, US.

European Parliament and Council of the European Union 2016, 'Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)', 27 April, *Official Journal of the European Union.*

Fischer-Hübner, S, Angulo, J, Karegar, F & Pulls, T 2016, 'Transparency, privacy and trust – Technology for tracking and controlling my data disclosures: Does this work?', *Trust Management X,* SM Habib, J Vassileva, S Mauw & M Mühlhäuser (eds), Springer International Publishing, CH, pp. 3-14.

Fritsch, L 2018, 'How Big Data helps SDN with data protection and privacy', *Big Data and Software Defined Networks,* J Taheri, J. (ed), The Institution for Engineering and Technology (IET), London, UK.

Fritsch, L, Fischer-Hübner, S, Pulls, T, Voronkov, A & Momen, N 2017, *Applications of privacy and security technologies for the protection of personal data in militarily relevant technologies such as IoT, smart environment and digital communications,* Karlstad University, Karlstad SE.

Fritsch, L & Fuglerud, KS, 'Time and usability economics as upper boundary in friend and family security and privacy', Position statement on Understanding Friend and Family based Security and Privacy issues workshop, NordCHI 2010, 17 October 2010, Reykjavik, IS, viewed 27 February 2019, <http://www.academia.edu/981624/Time_and_Usability_Economics_as_Upper_Boundary_in_Friend_and_Family_Security_and_Privacy>.

Fritsch, L & Momen, N 2017, 'Derived partial identities generated from app permissions', L Fritsch, H Rossnagel & D Hühnlein (eds), *Proceedings of the Open Identity Summit (OID 2017),* 5 October 2017, Gesellschaft für Informatik (Society for Computer Science), Karlstad, SE.

Greschbach, B, Pulls, T, Roberts, LM, Winter, P & Feamster, N 2017, 'The effect of DNS on Tor's anonymity', *Proceedings of the Network and Distributed System Security Symposium (NDSS) 2017,* San Diego, CA, US.

Iwaya, LH, Voronkov, A, Martucci, LA, Lindskog, S & Fischer-Hübner, S 2016, *Firewall usability and visualization: A systematic literature review,* Karlstad University Studies, Karlstad, SE.

Karegar, F, Gerber, N, Volkamer, M & Fischer-Hübner, S 2018, 'Helping John to make informed decisions on using social login', *Proceedings of the 33rd Annual ACM Symposium on Applied Computing,* ACM, pp. 1165-74.

Karegar, Lindegren, D, Pettersson, JS & Fischer-Hübner, S 2017, 'Assessments of a Cloud-based data wallet for personal identity management', P*roceedings of the 26th international conference on Information Systems Development: Advances in Methods, Tools and Management (ISD2017).* Larnaca, CY.

Kelley, PG, Consolvo, S, Cranor, LF, Jung, J, Sadeh, N & Wetherall, D 2012, 'A conundrum of permissions: Installing applications on an android smartphone', *Proceedings of the 16th International Conference on Financial Cryptography and Data Security,* 2012. Springer, Berlin, DE, pp. 68-79.

Laurie, B, Langley, A & Kasper, E 2013, 'Certificate Transparency', RFC 6962, Internet Engineering Task Force (IETF).

Marforio, C, Ritzdorf, H, Francillon, A & Capkun, S 2012, 'Analysis of the communication between colluding applications on modern smartphones',. *Proceedings of the 28th Annual Computer Security Applications Conference,* 3-7 December 2012, ACM, pp. 51-60.

Merkle, RC 1990, 'A Certified Digital Signature', *Proceedings of Advances in Cryptology, CRYPTO'89,* Lecture Notes in Computer Science series, vol 435. Springer, New York, pp. 218-38.

Momen, N, Pulls, T, Fritsch, L & Lindskog, S 2017, 'How much privilege does an app need? Investigating resource usage of Android apps', *Proceedings of the 15th International Conference on Privacy, Security and Trust (PST 2017),* IEEE Computer Society, Calgary, CA.

Murmann, P & Fischer-Hübner, S 2017, 'Tools for achieving usable ex post transparency: A survey', *IEEE Access,* vol. 5, pp. 22965-91.

Narayanan, A & Clark, J 2017, 'Bitcoin's academic pedigree, *Queue,* Association for Computing Machinery (ACM), New York, NY, US..

Paintsil, E & Fritsch, L 2013, 'Executable model-based risk analysis method for identity management systems: using hierarchical colored petri nets', *Proceedings of the 2013 TrustBus conference: Trust, privacy, and security in digital business,* Lecture notes in computer science, vol. 8058, Springer, Berlin, DE.

Peeters, R & Pulls, T 2016, 'Insynd: Improved privacy-preserving transparency logging, *Proceedings of the 21st European Symposium on Research in Computer Security (ESORICS 2016),* Springer International Publishing, Cham, CH, pp. 121-139.

Pulls, T & Peeters, R 2015, 'Balloon: A forward-secure append-only persistent authenticated data structure', *Proceedings of the 20th European Symposium on Research in Computer Security (ESORICS 2015),* G Pernul, P Y A Ryan, & E Weippl (eds), Springer International Publishing, Cham, CH.

Pulls, T, Peeters, R & Wouters, K 2013, 'Distributed privacy-preserving transparency logging', *Proceedings of the 12th ACM workshop on Workshop on Privacy in the Electronic Society (WPES 2013),* ACM, New York, US, pp. 83-94.

Romanosky, S & Goldman, Z 2016, 'Cyber collateral damage, *Procedia Computer Science,* vol. 95, pp. 10-17.

Scharre, P 2016, *Autonomous weapons and operational risk,* Center for a New American Security, Washington, DC, US.

Sigholm, J & Andersson, D 2011, 'Privacy on the battlefield? : Ethical issues of emerging military ICTs', Jeremy, M. (ed.) *Proceedings of the 9th international conference of Computer Ethics: Philosophical Enquiry (CEPE 2011),* M Jeremy (ed), 31 May-3 June 2011, Milwaukee, WI, US.

Tor 2018, *The Tor project,* viewed 31 January 2018, <https://www.torproject.org/>.

Wei, X, Gomez, L, Neamtiu, I & Faloutsos, M 2012, 'Permission evolution in the android ecosystem', *Proceedings of the 28th Annual Computer Security Applications Conference,* ACM, 3-7 December 2012, Orlando, FL, US, pp.31-40.

Winter, P & Lindskog, S 2012, 'How the great firewall of China is blocking Tor', *2nd USENIX workshop on Free and Open Communication on the Internet (FOCI 2012),* Bellevue, WA, US.

Winter, P, Pulls, T & Fuss, J 2013, 'ScrambleSuit: a polymorphic network protocol to circumvent censorship' *Proceedings of the 12th ACM Workshop on Privacy in the Electronic Society (WPES).* Berlin, DE.

Wool, A 2004, 'A quantitative study of firewall configuration errors', *Computer,* vol. 37, pp. 62-7.
Wool, A 2010, 'Trends in firewall configuration errors: Measuring the holes in swiss cheese', *IEEE Internet Computing,* vol. 14, pp. 58-65.

Voronkov, A 2017, 'Usable firewall rule sets', Licentiate thesis, Karlstad University, Karlstad, SE.

Voronkov, A, Lindskog, S & Martucci, L 2015, 'Challenges in managing firewalls', *Proceedings of the 20th Nordic Conference on Secure IT Systems (NordSec 2015),* Stockholm, SE, Springer, Cham, CH.

Yawning-Angel 2015, obfs4 - The Obfuscator, viewed 31 January 2018, <https://github.com/Yawning/obfs4>.

Behavioural Profiling for Transparent Verification in Cloud Storage Services

B Al-Bayati[1,2], N Clarke[1, 3], P Haskell-Dowland[3], F Li[4]

[1]Centre for Security, Communications and Network Research
University of Plymouth,
Plymouth, United Kingdom

[2]Computer Science Department, Science College
Diyala University
Diyala, Iraq

[3]Security Research Institute
Edith Cowan University
Perth, Australia

[4]School of Computing
University of Portsmouth
Portsmouth, United Kingdom

Email: burhan.al-bayati@plymouth.ac.uk; n.clarke@plymouth.ac.uk; p.haskelldowland@ecu.edu.au; fudong.li@port.ac.uk

Abstract: *Security is still the most sensitive issue in cloud storage services as services remain accessible to users for prolonged periods following an initial (usually simple) authentication. This issue has led to an increased vulnerability to potential attacks and sensitive customer information being misused. To that end, this paper investigates behavioural profiling for continuous and transparent authentication and assesses the legitimacy of users while they interact with an example cloud service—Dropbox. The best experimental result achieved an EER of 5.8%, based on an implementation using a real-life dataset. The results are encouraging and indicate the feasibility of detecting misuse in Dropbox.*

Keywords: *Continuous Identity Verification; Misuse; Transparent Authentication; Behavioural Profiling; Cloud Computing Services*

Introduction

According to the Cisco Global Cloud Index, by 2019, more than 80% of all data-centre traffic will be cloud traffic, and around 86% of all compute processing will be delivered using cloud infrastructure services (Cisco Global Cloud Index 2016). Cloud storage services have become particularly attractive for users (both individuals and enterprises) because they offer data storage to meet differing levels of demand. Customers can upload, download, update, remove, and share data by directly accessing information through online web applications from anywhere at any time.

The flexibility, accessibility, simplicity, efficiency, scalability, and pay-as-you-go options that are offered by cloud providers have proven successful (Forbes 2015). For example, Dropbox has more than 500 million user accounts, 300,000 businesses on Dropbox Business, and 400 billion total pieces of content uploaded daily (Dropbox 2018).
There is no doubt that cloud storage services can provide a flexible and convenient way for customers to access their data. However, customers still have concerns about how to protect the data stored remotely in these services from unauthorised access, with reports suggesting it is the biggest barrier to the adoption of cloud computing services. For example, a survey carried out by Garter showed that more than 70% of CTOs believed that data security and privacy concerns were the main reason hindering use of a cloud service (Chou 2013). Also, studies conducted on users' data stored in the cloud found that 88% of potential cloud customers were worried about their data security (for example, who has access to their personal data) (Fujitsu 2010; Griffith 2014).

Due to the online nature of those services, authentication provides the primary security control to prevent misuse by relying upon point-of-entry based passwords. By stealing customers' login credentials, hackers can gain access and can misuse the service and users' information. Many incidents have targeted popular cloud computing service providers, for example:

- According to Cloud Security Alliance, several security incidents affected a British telecom provider (TalkTalk) in 2014 and 2015, which resulted in the disclosure of personally identifiable data of approximately four million of their customers (Cloud Security Alliance 2016).
- Serious incidents on the Microsoft Azure cloud computing platform led to a massive collapse and outage of the service for 22 hours, with a loss of 45% of user data (Chen & Zhao 2012).
- Dropbox was hacked in July 2012; usernames and passwords of many users were stolen from third-party websites; hackers successfully gained access to customers' accounts and misused their data (BBC News 2016).
- Many Apple iCloud accounts were compromised in 2014 as more than 20,000 passwords were stolen, resulting in users' personal photographs, specifically those of celebrities, being leaked online (Gupta 2015).
- Google's Gmail server faced attack in 2016; more than 272 million email addresses and passwords were stolen (Yadron 2016).

Cybercriminals can obtain access to sensitive information of cloud services even though security controls are in place and dedicated security teams are allocated. Therefore, additional security techniques are needed to protect cloud storage services from being compromised and misused. This paper proposes a novel continuous identity verification system that protects cloud storage service users' data by operating transparently to detect unauthorised access. Behavioural profiling can continuously and transparently assess users' identities while they interact with cloud storage services. By creating user behaviour profiles, the system can identify people based upon the way in which they interact with these services. Therefore, the current users' activities (for example, time of opening the service) are compared with existing users' templates. These templates are generated from historical usage by implementing a machine learning algorithm, such as neural networks. The comparison result will determine if the current user is legitimate or not.

The remainder of the paper is structured as follows. The next section introduces the state-of-the-art use of behavioural profiling to detect the anomaly usage within various technologies. Then the authors present the experimental methodology. A series of comprehensive experimental studies to evaluate the applicability of using behavioural profiling with cloud storage service (in Dropbox, for example) are presented in the following section. Next, the authors discuss the impact of the experimental results, and the conclusion and future directions of this work are presented last.

Related Work

Various studies have investigated behavioural profiling from several security perspectives—including intrusion detection, fraud detection, and authentication across different technologies—such as mobile phone system, network, computer system, and web browsing. **Table 1**, below, provides an analysis of these studies.

Author(s)	**Activity**	**Client/Server**	**#Participants**	**Performance (%)**	**Method**	**Purpose**
(Moreau et al. 1997)						
Telephony	Server	600	DR=90, FRR=10	Supervised Neural Networks	Fraud detection	
(Burge & Shawe-Taylor 1997)						
Telephony	Server	110	DR=75, FRR=40	Unsupervised Neural Network	Fraud detection	
(Samfat & Molva 1997)						
Mobility	Server	400	DR=82.5, FRR=40	Distance	Fraud detection	
	Telephony			DR=80, FRR=30	Rule-base	
(Hall, Barbeau & Kranakis 2005)						
Mobility	Server	50	DR=50, FRR=50	Instance based learning	IDS	
(Hilas et al. 2014)						
Telephony	Server	5000	DR=80	Genetic Programming method	Fraud detection	
(Ogwueleka 2009)						

Telephony	Server	180	FRR=3	Self-Organizing Map and Probabilistic models	Fraud detection	
(Qayyum et al. 2010)						
Telephony	Server	300	DR=70	Neural Network	Fraud detection	
(Yazji et al. 2011)						
Mobility	Server	100	DR=81	Cumulative probability and Marko properties of trajectories	IDS	
(Yazji et al. 2014)						
Mobility	Server	178	DR=94	Cumulative probability and Marko properties of trajectories	IDS	
(Subudhi & Panigrahi 2015)						
Telephony	Server	94	DR=97	SVM	Fraud detection	
(Shi et al. 2010)						
Telephony, SMS, Browsing, Mobility	Client	50	DR=95	Probability	Authentication	
(Damopoulos et al. 2012)						
Telephony, SMS, Browsing	Client	35		Bayesian network, RBF, KNN, Random Forest	Authentication	
(Li et al. 2010)						
Telephony, Device Usage, Bluetooth						
Network Scanning	Client	30	EER=13.5, 35.1, & 35.7	RBF network	Authentication	
(Li et al. 2011)						

Application, Telephony, SMS	Client	76	EER=13.5, 2.2, 5.4	Neural network	Authentication	
(Li et al. 2014)						
Application Usage	Client	76	EER=9.8	Rule base	Authentication	
(Fridman et al. 2017)						
Text, App, Web and Location	Client	200	EER=3	SVM	Authentication	
(Aupy & Clarke 2005)						
Way of Using PC	Client	21	EER= 7	Neural Network (FF-MLP)	Authentication	
(Yazji et al. 2009)						
File Access activity and network event	Client	8	DR=90, FAR=14, FRR=11	K-Means Clustering	Authentication	
(Salem & Stolfo 2011)						
File access activity	Client	18	FAR=1.1	SVM	Insider detection	
(Yang 2010)						
Web Browsing	Server	100	DR=91	Support-based, lift-based profiling	Identification	
(Abramson & Aha 2013)						
Web Browsing	Server	10	EER= 24	SVM	Authentication	

Table 1: Related behavioural profiling studies

*DR: Detection Rate; FRR: False Reject Rate; FAR: False Accept Rate; EER: Equal Error Rate; SVM: Support Vector Machine; KNN: K-Nearest Neighbours; RBF: Radial Basis Function; IDS: Intrusion Detection System

Early research focused mainly on intrusion detection and fraud detection based on identifying the user behaviour activities during their interaction with mobile services, such as calling and mobility (Moreau et al. 1997; Burge & Shawe-Taylor 1997; Samfat & Molva 1997; Hall Barbeau & Kranakis 2005; Hilas et al. 2014; Ogwueleka 2009; Qayyum et al. 2010; Yazji et al. 2011; Yazji et al. 2014; Subudhi & Panigrahi 2015). In comparison, more recent studies have focused on transparent authentication through modelling application usage to alleviate device misuse (Shi et al. 2010; Damopoulos et al. 2012; Li et al. 2010; Li et al. 2011; Li et al. 2014; Fridman et al. 2017).

Much more information can be gathered from user activities while interacting with these applications (for example, GPS locations, emails, website visits, and calendar activities). These activities have been exploited to build a more accurate behavioural profile, which can be investigated to increase the accuracy level of the security system for the device or application itself.

Further studies focused on the generation of user behaviour profiles from desktop computer usage to detect any illegal access to the device (Yazji et al. 2009; Salem & Stolfo 2011). Several features were extracted to build user behaviour profiles in the computer system, including applications being used, the time and interval of accessing files, and websites being visited. The accuracy of these studies was around 7% of Equal Error Rate (EER). However, the number of participants was limited (ranging from 8 to 18 users) which does not reflect an accurate performance in the practical sense. From the server-side perspective, studies focused on building a user identifier by using web surfing activities from numerous log files of websites (Yang 2010; Abramson & Aha 2013). A user behaviour profile was created based on spending time on various topics of the website, site names, number of pages, starting time, and duration time of sessions. An accurate user behaviour profile has been built to detect illegitimate usage. The best performance was achieved by Yang (2010) including 100 participants with a Detection Rate (DR) of 94%. However, the study did not involve all users in the practical experiment; only a few users who had at least 300 sessions in the dataset were selected to test the system. Therefore, it is difficult to implement this system to solve large- scale problems.

Based upon existing literature, the behavioural profiling technique has been applied successfully across different technologies including mobile phones, computers (client and server) to improve the system security level. However, to the best of the authors' knowledge, no prior work that utilises behavioural profiling has been studied with respect to cloud storage services.

Experimental Methodology

The aim of this study is to understand to what degree behaviour profiling can be used to verify an individual's identity within cloud storage services—understanding whether it is the legitimate user or not provides a basis for the system to respond to potential misuse. Therefore, a series of experiments were conducted on users of cloud storage services to examine different factors that can affect the performance of the classification algorithms. These include

- The nature of different classification approaches to explore how the performance is affected;
- The impact of the volume of data for training and testing on the performance of the system;
- The effect of time series rather than random sample selection on the accuracy of a decision.

To conduct these experiments, a user's activities within the cloud storage service are needed, in terms of both quality and quantity. Due to privacy and security concerns, it is challenging to get a workable dataset from cloud storage providers. Also, to the best of the authors' knowledge, no public dataset on users' cloud activities is available. Dropbox, Google Drive, One Drive, and Box are all examples of widely popular cloud storage services. Dropbox was chosen for this research since it is one of the most popular cloud storage services (Griffith 2014; CloudRAIL 2017); and, importantly, it provides simply access to users' interactions records. This made it possible to

access and capture logs of user Dropbox activity over prolonged periods of time. By downloading users' historical activities, a private dataset was collected from a cloud storage service (Dropbox) containing real user interactions of 30 participants over a six-month period from 02/09/2015 to 02/03/2016 (totalling 91,371 log entries). The data has been anonymised to protect the participants' privacy, and ethical approval was sought and obtained from the authors' institutions. For each user interaction, the following information is available:

- timestamp of the action, such as day, hour, and minute;
- file type, such as pdf, jpg, and docx;
- user's action, such as add, edit, delete, move, and rename.

In order to make those features acceptable by classification algorithms, the symbolic-valued attributes, such as file type and user action, were enumerated into numerical attributes and into the range of 0-1 (Sola & Sevilla 1997).

The records of each user were divided into two sets: the first set was used to generate a profile for training; while the second set was used to evaluate the classifiers' performance (referred to as the test set). Classification is based upon a 2-class problem; legitimate or impostor, where one user acts as the legitimate user, with the remaining users acting as impostors. This is then repeated to ensure all users can act as an authorised user. The corresponding False Acceptance Rate (FAR) and False Reject Rate (FRR) are computed and are used to determine the EER. The EER (the point at which the FAR and FRR are equal) is used as the key performance metric.

The first experiment explored how the performance of the system is affected by investigating the nature of different classification approaches. The findings of this experiment would also help to identify the optimal classifier. Four supervised machine learning algorithms were selected: Feed-Forward Multi-Layered Perceptron (FF MLP), Random Forest (RF), Support Vector Machine (SVM) and Classification and Regression Trees (CART). The first three approaches were selected based on the highest performance that was achieved with the previous studies (as illustrated in Table 1, above) whereas the fourth method was based on the study by Wu et al. (2008) that was conducted on different classification algorithms. Also, default configurations are used for all selected classifiers. The experiment utilised a 66/34 splitting for the training and testing data with random selection across the dataset. At no point is a sample used for both training and testing.

The second experiment focused on investigating the impact of training and testing data split upon the performance. In addition to the 66/34 split, 50/50 and 80/20 splitting approaches were investigated for training/testing with random sample selection. Regarding the classifier, the classification algorithm that achieved the best performance from the first experiment was selected. The comparison between the accuracy of the result given differing levels of training data would provide a better understanding of the nature of user behaviour profiles and the volume of data necessary to achieve an appropriate level of performance. The first two experiments sampled data randomly across the dataset, in order to understand the general feasibility of the approach. However, in practice, a profile would need to be created based upon time-series (that is, those it collected first rather than samples it has yet to capture). As such, the third experiment sought to evaluate the impact of applying time series on the performance. In order to understand the effect of the two metrics on

the performance, the similar volume of data for the training and testing sets of the previous experiment is applied. The accuracy of each volume is compared with the accuracy of the volume of the previous experiment.

Experimental Results

Classification algorithms

The overall results of this experiment are presented in Table 2, below. Generally speaking, the results are encouraging to support the idea of verifying the legitimate user or unauthorised access to data stored in cloud storage services, with EERs that are aligned to similar results in other applications from the prior work (Li et al. 2010; Li et al. 2011; Li et al. 2014; Fridman et al. 2017).

Classifier	Processing time in minutes	EER (%)
SVM	273	20.27
RF-25 trees	50	9.93
FF MLP Neural Network-65	1700	6.98
CART	10	6.02

Table 2: Performance of classification algorithms

The nature of classifier utilised does have an impact the overall performance; however, except for SVM, the variation in performance is not overly significant—suggesting the classifier itself is not overly key. As seen in Table 2, above, the CART was the fastest algorithm to perform the training and testing tasks with only 10 minutes for all users; also, it achieved the highest accuracy amongst all chosen classifiers with an EER of 6.02%. This would allow other factors, such as time taken to compute, computational overhead, and memory requirements to be considered as part of the selection.

A more detailed analysis of the classifiers was undertaken to determine what impact optimisation would have. The results from the FF MLP and RF methods are demonstrated below in Figure 1 and Table 3, respectively. For the FF MLP classifier, the best result of EER 6.98% was achieved by using 65 neurons. However, it needed the longest time to perform the training and testing tasks comparing to the other three classifiers (1700 minutes); while with the RF approach, the best performance of EER 9.93% was obtained when 25 trees were used with only 50 minutes to achieve the training and testing stages. Neither SVM or CART had any parameters to optimise.

Number of trees	EER (%)
5	10.39
10	10.41
15	10.21
20	10.18
25	9.93
30	10.26
35	10.43

Table 3: Performance of RF with trees

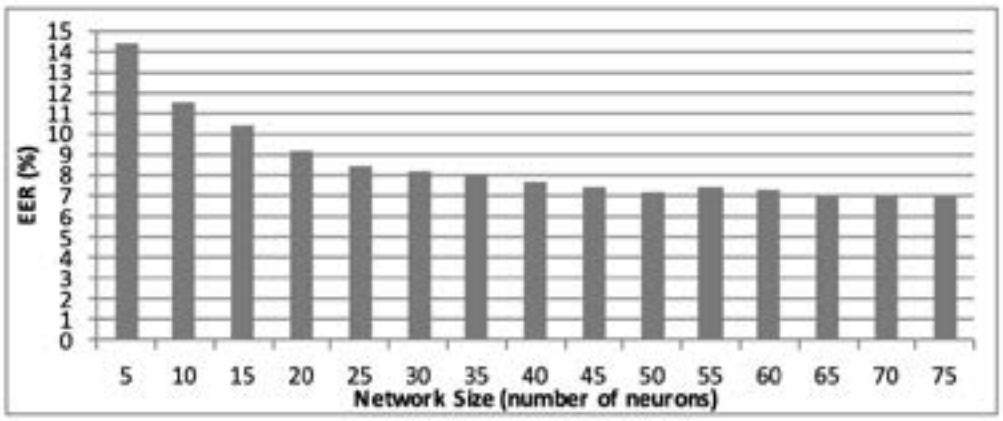


Figure 1: Performance of FF MLP with different network configurations

Prior research has shown the volume of data per user has a significant impact upon performance. As such, an analysis of the volume of user data was carried out. Users were divided into two groups based on their interactions, as illustrated in Table 4, below. The first 15 users belong to the users' group which has fewer interactions (equal to or less than 2,000 interactions) whereas the remaining 15 users belong to the users' group which has more interactions (more than 2,000 interactions). The selection of 2,000 was felt sufficient to separate the groups yet ensure a suitable number of participants were left in each group. Based upon the overall average performance from these two groups, users who have more interactions achieved a better performance than users with fewer interactions when using the RF, FF MLP, and CART classifiers.

However, this was not always true on a per-user basis. For example, although Users 17, 18, and 21 have more interactions than many other users, they achieved a lower performance than many users with low interactions such as Users 2, 4, 9, and 13. Further investigation suggests that those three users used Dropbox as a backup solution by uploading photos within a period of time which can be carried out automatically by a computer rather than the users themselves. This created difficulty for the classifiers to differentiate between user actions and computer-generated activities. Similarly, some low usage users got a better accuracy than many more active users. For instance, User 2 achieved less than 2% of EER across the most approaches, and User 13 got an EER close to zero. Looking to the usage of these users reveals that they worked constantly on specific file types that the majority of the rest of users do not use. This unique pattern of usage made the classifier more able to discriminate them from others. This result suggests that users who have more interactions achieve better performance in general. However, the uniqueness of interactions can be a key factor to build discriminative patterns for users, which can make classifiers more accurate in distinguishing between them.

User	**No. of Interactions**	**EER % based on classifier algorithms**			
		SVM	RF	FF MLP	CART
1	549	19.08	8.88	10.32	5.13
2	585	2.11	3.47	1.95	1.96
3	652	6.23	9.68	4.65	3.35
4	677	2.52	1.56	1.43	2.02
5	726	26.80	6.55	7.21	3.63
6	764	23.79	8.17	13.09	5.71
7	797	3.02	28.29	4.78	14.89
8	1146	23.76	36.09	15.75	23.73
9	1370	10.67	2.36	4.70	1.22
10	1413	31.49	6.17	3.86	3.86
11	1462	13.47	10.08	4.89	5.36
12	1656	25.14	33.00	12.49	16.64
13	1714	11.25	0.11	0.22	0.02
14	1765	32.30	18.75	16.84	10.95
15	1988	39.35	19.86	15.02	15.56
Av*	1,151	18.06	12.87	7.81	7.6
16	2250	30.22	4.69	7.81	4.08
17	2373	43.34	20.13	13.08	12.63
18	2487	28.53	19.87	9.13	12.27
19	2799	10.22	2.71	4.84	1.26
20	2879	17.54	0.56	2.40	0.54
21	2960	20.02	25.05	8.34	14.38
22	3226	28.17	1.33	3.57	0.99
23	3464	3.08	1.26	1.54	0.53
24	3568	31.55	3.95	13.88	1.91
25	4858	25.28	3.94	4.30	3.34
26	5780	7.13	0.15	0.19	0.15
27	6440	13.21	1.65	1.85	1.37
28	7263	29.06	11.20	7.78	5.89
29	14985	29.99	7.69	10.79	6.64
30	15013	19.81	0.68	2.79	0.75
Av**	5,356	22.48	7.00	6.15	4.45

Av*: Average from the first group, Av**: Average from the second group

Table 4: Users' performance with different classifiers

Volume of data for training and testing

This experiment studied the impact of the volume of data for training upon the performance. The CART classifier was chosen for this experiment due to its (best) performance from the first experiment; also, the data splitting between training the classifier and testing the performance was set to 50/50, 66/34, and 80/20. **Table 5,** below, illustrates the performance of all users across the selected volumes of data.

User	**EER (%) based upon volume of data**		
1	7.28	5.13	5.09
2	2.59	1.96	1.65
3	7.30	3.35	4.26
4	1.84	2.02	1.14
5	3.66	3.63	2.05
6	6.57	5.71	4.96
7	15.20	14.89	14.14
8	25.25	23.73	24.45
9	2.14	1.22	0.84
10	7.06	3.86	5.39
11	6.47	5.36	5.32
12	20.04	16.64	16.04
13	0.09	0.02	0.02
14	9.92	10.95	9.77
15	13.81	15.56	14.50
16	3.88	4.08	2.19
17	13.58	12.63	11.93
18	13.93	12.27	12.98
19	1.83	1.26	1.38
20	0.59	0.54	0.70
21	16.63	14.38	14.36
22	1.22	0.99	1.35
23	0.90	0.53	0.89
24	2.33	1.91	1.73
25	3.35	3.34	3.51
26	0.21	0.15	0.17
27	1.56	1.37	1.80
28	7.15	5.89	6.04
29	6.63	6.64	6.39
30	0.85	0.75	0.77
Average	6.79	6.02	5.86

Table 5: Performance based on volume of data with random selection

As shown in **Table 5,** the training phase with a larger volume of samples achieves better performance than those with a smaller volume of data on average; the best result performance was 5.86% of EER achieved by using 80/20 splitting for training and testing respectively. This agrees with the prior research and suggests that larger volume of samples for training the classifier can have a positive impact on the overall performance. This is logical as the classifier can be trained more about user behaviour pattern by using a larger volume of data, leading to a better performance. However, it is also worth highlighting that the change in performance from 6.79% to a best case of 5.86% is not significant. This suggests that the nature of user behaviour across the 6-month collection period is likely to be relatively stable.

From an individual user's perspective, the increasing volume of data for the training stage has different impacts upon the performance. When increasing the training data volume to 66/34 and 80/20 splitting, a number of users' performance improved and some stayed relatively stable—suggesting more data made little difference. In a practical sense, being able to understand which users have more stable or active profiles would be useful in interpreting the classification decisions and in template retraining.

Time series sample selection

In addition to the random sample that is used for the previous experiment (a standard methodological approach in feasibility studies), the impact of the time and natural changes in user behaviour over time is important to evaluate. The CART classifier was with the data split for training and testing in the same manner as the pervious experiment (that is 50/50, 64/34, and 80/20). The results of the experiment are presented in **Table 6,** below.

User	**EER (%) based upon volume of data**		
1	18.38	15.86	15.15
2	2.61	3.84	4.28
3	6.49	6.22	6.41
4	2.92	1.18	0.04
5	18.28	25.05	21.06
6	2.43	9.85	3.54
7	19.90	17.24	19.45
8	41.39	40.57	43.59
9	2.36	3.15	2.80
10	27.60	38.91	23.38
11	10.07	8.89	4.71
12	40.31	38.13	36.66
13	15.69	19.24	14.40
14	17.32	17.05	19.60
15	27.37	21.14	23.83
16	2.22	1.37	1.13
17	34.52	32.90	35.51

18	25.87	19.43	15.45
19	2.10	1.61	1.15
20	8.54	4.51	5.32
21	25.10	22.64	19.03
22	1.74	2.30	1.89
23	1.90	0.96	0.82
24	6.27	5.34	4.57
25	4.92	5.09	6.32
26	0.11	0.19	0.31
27	3.28	1.91	2.29
28	11.28	11.05	9.74
29	8.77	8.35	10.22
30	0.81	0.83	0.85
Average	**13.02**	**12.83**	**11.78**

Table 6: Performance of the different volume of data with time series selection

As demonstrated by Table 6, the best performance is EER 11.78%, and it is achieved by using the 80/20 data splitting for training and testing. Similar to experiment 2, the nature of the data split did not have a significant impact upon performance; however, the results themselves have doubled. This suggests that over-time user behaviour does change and, therefore, care must be taken into account in order to ensure that appropriate template renewal procedures are developed to maintain levels of performance.

Discussion

The experimental results reveal that cloud storage service users can be discriminated via their usage with a reasonable performance being achieved. Also, the outcome of this research is in line with the highest results that are achieved in the related works such as Shi et al. (2010), Aupy & Clarke (2005), Yazji et al. (2014), and Subudhi & Panigrahi (2015). In terms of the performance of each individual classifier, the CART algorithm achieves 6.02% EER and outperforms the other three chosen classifiers (SVM, RF and FF MLP). Also, the CART classifier was the fastest algorithm to process the given data (it utilised 10 minutes to complete both the training and testing phases on all users). It is an important factor to consider for real-world applications, as the quicker a classifier can produce an output, the sooner a security control can decide. From an individual user's perspective, on average users who have more frequent activities/interactions acquired better results than those who have the fewer interactions across most classifications. However, users with fewer interactions did also achieve a good level of performance. For example, when examining the interactions of those users such as Users 4, 9, and 13, the authors observed that they have a unique way of using Dropbox (particularly in terms of the uniqueness of their file types). Therefore, a good pattern (uniqueness) from the users' interactions can also affect the performance of the classifiers, even if the number of users' activities is low.

The results of the second and third experiments show that the data split for training and testing the classifier and the timestamp factors have an impact upon the overall performance. As shown in

these two experiments, a larger volume training data (80/20 splitting) with random sample selection achieves better performance with 5.8% of EER on an average. However, regarding individual users, the performance for several users with more training data (that is 66/34 and 80/20) is not as good as the results being achieved by using less training data (a 50/50 split for training and testing). For example, User 8 had the lowest performance compared to all other users. Increasing the volume of data did not improve accuracy with results degrading notably—particularly with time series sample selection. When looking at the pattern of daily usage for events, one finds that they did not seem to have consistent usage, as some of the events such as 'Edit' only appeared within the first two months with no later occurrences, while other events such as 'Delete' were only present in the final days of the interaction logs, as shown in **Figure 2.**

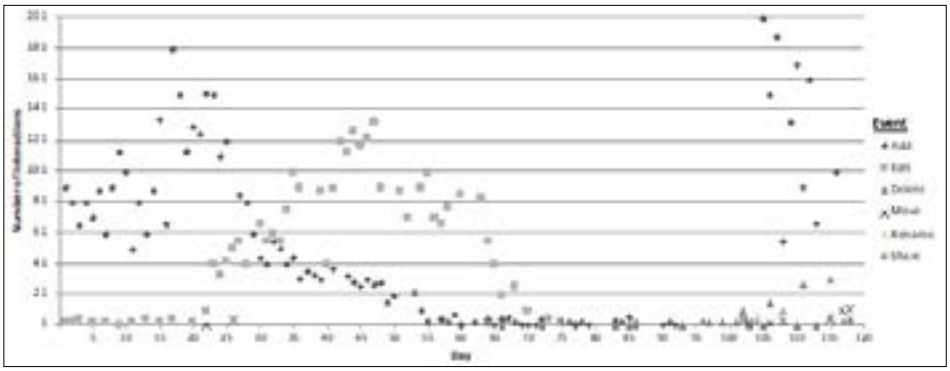


Figure 2: Recorded interactions for User 8

These changes in user behaviour can have a negative effect on the performance of classifiers because their activities are so diverse.

When researchers applied the behaviour profiling technique in practice, the time series sample selection showed a significant difference over the random sampling. Therefore, users' templates need to be updated regularly to ensure quality for achieving a high level of system performance. However, the renewal of users' templates dynamically is not an easy task because it might need to avoid including impostor's behaviour with the legitimate behaviour. For example, an impostor might be accepted by the system over time as the genuine user since more and more impostor samples are included within the template renewal process. This problem needs to be managed carefully and correctly to avoid capturing of illegitimate usage and to ensure that the legitimate user is not unnecessarily inconvenienced in day-to-day interactions (that is, not intrusively asking the user to re-enroll in the system).

Conclusion and Future Work

The experimental results of applying a number of machine learning algorithms showed the ability to correctly discriminate between users based on their interactions derived from the cloud storage service Dropbox. Using behavioural profiling, an accurate user's template can be built to help in distinguishing between the normal and abnormal usage. A high performance was achieved by the most classification algorithms, except the SVM was not performing particularly well. Further experiments have shown that time-series versus random sampling of data for training does have a significant impact upon performance; however, the volume of training data has less of an

effect. With respect to individual performance, many participants achieved a high performance when the system was capable of identifying their interactions without any error. Subsequently, the approach proved a highly promising solution to applying user behavioural profiling as a supporting technique to validate the users after initial point-of-entry authentication. This can contribute and guide the system to identify a misuse of cloud services in a continuous and friendly manner. However, there were a number of users who performed particularly poorly and, in line with most behavioural-based applications, would not be suited to such a technique.

Future work will focus on developing mechanisms for understanding where and when such an approach can be utilised (that is enabling the approach for users with a sufficiently stable profile) and when and how template renewal should be undertaken.

References

Abramson, M & Aha, DW 2013, 'User authentication from Web browsing behavior', *Proceedings of the 26th Florida Artificial Intelligence Research Society Conference, FLAIRS,* pp. 268-73.

Aupy, AM & Clarke, NL 2005, 'User authentication by service utilisation profiling', *Advances in Network and Communications Engineering 2*, pp. 18 -26

BBC News 2016, 'Dropbox hack "affected 68 million users"', BBC News Web page, viewed 17 October 2017, <http://www.bbc.co.uk/news/technology-37232635>.

Burge, P & Shawe-Taylor, J 1997, 'Detecting cellular fraud using adaptive prototypes', *Proceedings of AAAI-97: Workshop on AI approaches to fraud detection and risk management,* Providence, RI, US, pp. 9-13.

Chen, D & Zhao, H 2012, 'Data security and privacy protection issues in cloud computing', *Proceedings of the 2012 International Conference on Computer Science and Electronics Engineering (ICCSEE)*, vol. 1, pp. 647-51.

Chou, TS 2013, 'Security threats on cloud computing vulnerabilities', *International Journal of Computer Science & Information Technology,* vol. 5, no. 3, p. 79.

Cisco 2016 'Cisco global cloud index: Forecast and methodology', White paper, viewed 1 March 2019, <https://www.cisco.com/c/en/us/solutions/collateral/service-provider/global-cloud-index-gci/white-paper-c11-738085.html>.

CloudRAIL 2017, 'Cloud Storage Report 2017-- Dropbox loses market share but is still the biggest provider on Mobile-CloudRail', viewed 19 November 2017, <https://blog.cloudrail.com/cloud-storage-report-2017/>.

Cloud Security Alliance 2016, 'The treacherous 12 cloud computing top threats in 2016', February, viewed 1 March 2019, <https://downloads.cloudsecurityalliance.org/assets/research/top-threats/Treacherous-12_Cloud-Computing_Top-Threats.pdf>.

Damopoulos, D, Menesidou, SA, Kambourakis, G, Papadaki, M, Clarke, N & Gritzalis, S 2012, 'Evaluation of anomaly-based IDS for mobile devices using machine learning classifiers', *Security and Communication Networks,* vol. 5, no. 1, pp.3-14.

Dropbox 2018, 'About – Dropbox', viewed 1 October 2018, <https://www.dropbox.com/about>.

Forbes 2015, 'Roundup of cloud computing forecasts and market estimates', viewed 27 April 2015, <http://www.forbes.com/sites/louiscolumbus/2015/01/24/roundup-of-cloud-computing-forecasts-and-market-estimates-2015/>.

Fridman, L, Weber, S, Greenstadt, R & Kam, M 2017, 'Active authentication on mobile devices via stylometry,application usage, web browsing, and GPS location', *IEEE Systems Journal,* vol. 11, no. 2 , pp. 513-521.

Fujitsu 2010, *Personal data in the cloud: A global survey of consumer attitudes,* viewed 1 March 2019, <http://www.fujitsu.com/downloads/SOL/fai/reports/fujitsu_personal-data-in-the-cloud.pdf>.

Griffith, E 2014, 'Who's winning the consumer cloud storage wars?', 6 November, viewed 20 November 2017, <http://fortune.com/2014/11/06/dropbox-google-drive-microsoft-onedrive/>.

Gupta, U 2015, 'Survey on security issues in file management in cloud computing environment', Cornell University, arXiv:1505.00729 [cs.CR], viewed 1 March 2019, <http://arxiv.org/abs/1505.00729>.

Hall, J, Barbeau, M & Kranakis, E 2005, 'Anomaly-based intrusion detection using mobility profiles of public transportation users', *Proceedings of the IEEE International Conference on Wireless and Mobile Computing, Networking and Communications,* WiMob'2005, vol. 2, pp. 17-24.

Hilas, CS, Kazarlis, SA, Rekanos, IT & Mastorocostas, PA 2014, 'A genetic programming approach to telecommunications fraud detection and classification', *Proceedings of the 2014 International Conference on Circuits, System Signal Processing,* Communications and Computers, pp. 77-83.

Li, F, Clarke, N, Papadaki, M & Dowland, P 2010, 'Behaviour profiling on mobile devices', *Proceedings of the 2010 International Conference on Emerging Security Technologies (EST),* pp. 77-82.

——2011, 'Misuse detection for mobile devices using behaviour profiling', *International Journal of Cyber Warfare and Terrorism,* vol. 1, no. 1, pp. 41-53.

——2014, 'Active authentication for mobile devices utilising behaviour profiling', *International Journal of Information Security,* vol. 13, no. 3, pp. 229-44.

Moreau, Y, Verrelst, H & Vandewalle, J 1997, 'Detection of mobile phone fraud using supervised neural networks: A first prototype', *Proceedings of the International Conference on Artificial Neural NetworksArtificial Neural Networks—ICANN'97*, Lecture notes in computer science, vol 1327, Springer, Berlin, Heidelberg, DE, pp. 1065-70.

Ogwueleka, FN 2009, 'Fraud detection in mobile communications networks using user profiling and classification techniques', *Journal of Science and Technology,* vol. 29, no. 3, pp. 31-42.

Qayyum, S, Mansoor, S, Khalid, A, Halim, Z & Baig, AR 2010, 'Fraudulent call detection for mobile networks', *Proceedings of the 2010 International Conference on Information and Emerging Technologies (ICIET),* pp. 1-5).

Salem, M & Stolfo, S 2011, 'Modeling user search behavior for masquerade detection', *Recent Advances in Intrusion Detection,* Springer Berlin, Heidelberg, DE, pp. 181-200.

Samfat, D & Molva, R 1997, 'IDAMN: An intrusion detection architecture for mobile networks', *IEEE Journal on Selected Areas in Communications,* vol. 15, no. 7, pp. 1373-80.

Shi, E, Niu, Y, Jakobsson, M & Chow, R 2010, 'Implicit authentication through learning user behavior', *Proceedings of the 2010 International Conference on Information Security, ISC 2010,* Lecture notes in computer science, vol 6531. Springer, Berlin, DE, pp. 99-113.

Sola, J & Sevilla, J 1997, 'Importance of input data normalization for the application of neural networks to complex industrial problems', *IEEE Transactions on Nuclear Science,* vol. 44, no. 3, pp.1464-8.

Subudhi, S & Panigrahi, S 2015, 'Quarter-Sphere support vector machine for fraud detection in mobile telecommunication networks', *Procedia Computer Science,* vol. 48, pp. 353-9.

Wu, X, Kumar, V, Ross, QJ, Ghosh, J, Yang, Q, Motoda, H, McLachlan, GJ, Ng, A, Liu, B, Yu, PS, Zhou, ZH, Steinbach, M, Hand, DJ & Steinberg, D 2008, 'Top 10 algorithms in data mining', *Knowledge and Information Systems,* Springer, London, UK.

Yadron, D 2016, 'Hacker collects 272m email addresses and passwords, some from Gmail', *The Guardian,* viewed 25 November 2017, <https://www.theguardian.com/technology/2016/may/04/gmail-yahoo-email-password-hack-hold-security>.

Yang, Y 2010, 'Web user behavioral profiling for user identification', *Decision Support Systems,* vol. 49, no. 3, pp. 261–71.

Yazji, S, Chen, X, Dick, RP & Scheuermann, P 2009, 'Implicit user re-authentication for mobile devices', *Proceedings of the 6th International Conference of Ubiquitous Intelligence and Computing,* 7-9 July, Brisbane, QLD, AU, pp. 325-39.

Yazji, S, Dick, RP, Scheuermann, P & Trajcevski, G 2011, 'Protecting private data on mobile systems based on spatio-temporal analysis', *Proceedings of the 1st International Conference on Pervasive and Embedded Computing and Communication Systems,* 5-7 March, Vilamoura, Algarve, PT, pp. 114-23.

Yazji, S, Scheuermann, P, Dick, RP, Trajcevski, G & Jin, R 2014, 'Efficient location aware intrusion detection to protect mobile devices', *Personal and Ubiquitous Computing,* vol. 18, no.1, pp.143-62.

Cyber-Securing Super Bowl 50: What Can a Live-Fire Football Match Teach Students about Becoming Better Cybersecurity Professionals?

MW Bovee, HOL Read

Computer Science/Computer Security & Information Assurance
School of Business & Management
Norwich University
Northfield, Vermont, United States

Email: mbovee@norwich.edu; hread@norwich.edu

Abstract: *The rise and regularity of cybersecurity incidents have increased the demand for trained workforce professionals. Institutions of higher education have responded by including practical hands-on exercises such as capstones, labs, and simulated attack-and-defend 'Capture-the-Flag' scenarios. Many degree programs also encourage students to gain experience via internships. This paper considers real-world experience gained by students through another means—by assisting law enforcement personnel in defending Super Bowl 50 cyberspace. This annual game, with high security requirements and international prominence, provided a unique opportunity to reflect whether 'live-fire' experiences can improve technical and professional skill sets of students who are emerging from higher-education into the workforce.*

Keywords: *Cybersecurity, Experiential Learning, Professionalism, Curriculum*

Introduction

In the United States, the Super Bowl is an annual, internationally-prominent, sports-related, mass gathering. It is the culminating event of the American professional football season, and routinely draws stadium crowds in excess of 80,000 plus an international audience of over 100 million watching the event on television (Grossi 2014). In 2002, in the wake of the September 11th atrocities, the visibility and significance of the event resulted in its being designated by the United States President as a National Security Special Event (NSSE) (United States Government 18 U.S.C. § 3056[e]). Because of its magnitude and significance, the Super Bowl is routinely classified as a Special Event Assessment Rating of 1 (SEAR 1)—an event that warrants the support of the American Government (Reed n.d). As a SEAR 1 event, the 50th annual Super Bowl (Super Bowl 50, or 'SB50') involved a number of organisations, including: the Federal Bureau of Investigation, the Department of Homeland Security, the Secret Service, Customs and Border Control, the U.S. Postal Service, the Transportation Security Administration, the Federal Aviation Administration, the U.S. Air Force, and the U.S. Coast Guard—all led by the Santa Clara Police Department (SCPD) (Grossi 2014).

Super Bowl 50 was hosted by Levi's stadium in Santa Clara, California. This stadium was designed and constructed with 1200 WiFi hotspots and bandwidth (40 gigabits per second) to allow all guests simultaneous, real-time WiFi access during games. There are sufficient connections—40 times that of any other stadium in the U.S.—for it to be considered one of the most high-tech sports venues to date (Bajarin 2014). Such infrastructure also supports emergency services, crowd control, and the evolving 'fan experience' of such activities as watching instant replays and ordering products that are then delivered to the fans at their seat while in the stadium (Martin 2016).

Within this environment, a cohort of 65 students drawn from predominantly Computer Science (CS) and Computer Security and Information Assurance (CSIA) majors worked to assist the SCPD in developing and implementing solutions to protect and defend the 'cyberspace' of the SB50 event from any would-be assailants looking to damage the reputation of those involved, to cause disruption in the fan experience, or to prevent any other attackers with sinister motives. The focus of this paper is on the experience gained from the perspective of the students. The paper considers whether real-world exercises such as this are of merit, whether they improve student knowledge of working in the cyber field, and whether they help to prepare students for future work in the cyber workforce.

In the next section, entitled 'Related Work', the authors examine the developing need of cybersecurity education and the move towards incorporating more practical experience. The section entitled 'Contribution' highlights the contribution of this paper to the field of cybersecurity education; 'Methodology' details the project structure and processes for engagement with Super Bowl 50; and 'SB50 Project Development' describes the process taken to identify new learning by student participants. The section called 'Discussion' considers the results of new learning by students, and 'Conclusions' provides closing remarks about the overall project.

Related Work

The idea of universities and similar institutions teaching cyber-related curricula is not new. The first undergraduate degree to feature the term 'hacking' appeared in 2006 (Abertay 2016), while many programs in information security were available as far back as the 1990s (Kessler & Ramsay 2013). The typical forms of teaching cyber within higher-education institutions have since aggressively moved away from the more traditional forms of teaching (lectures, reading literature, understanding concepts in principle) as students often cannot apply the academic principles they have learned to a realistic environment (Willems & Meinel 2012). Available literature in the public domain shows that the 'Capture-the-Flag' (CtF) genre, whereby a specific aim or goal is set, typically for an offensive exercise such as obtaining a particular file from a system, has remained very popular as an educational tool to help students understand how to configure, respond, defend, attack, and exploit networked systems. Indeed, many security organisations have taken to using this model of 'gamification' (Herr & Allen 2015) as a recruiting tool, including government, as seen in the Government Communication Headquarters (GCHQ 2011) and the National Security Agency (NSA 2014). Others encourage a team-based model of this approach. Conklin (2007) describes an information security practicum course whereby students, working as part of a team, make amendments in a simulated small business environment. Changes are issued via memos and outside of student class time (for example, by introducing malware or the 'accidental' deletion of a file). The real-world simulation is kept by maintaining system states between classes (thus, providing

the sense of continuity), by incorporating the input of industry professionals (thus, preventing the instructor from doing the 'same old thing'), and by focusing on the business (thus, preventing students from treating them like their 'personal playgrounds') (Conklin 2007). Rege (2015) applies cyber curricula to students without a strong background in computing (criminal justice majors). Such students encounter issues with the prevailing CtF model, namely novice encouragement, temporal constraints, and skewed experiences (barriers to entry based on prior knowledge).

Similar practical educational exercises have been developed for other, more focused areas within the cyber realm. Sitnikova, Foo, and Vaughn (2013) discuss their experiences taking the experiential model in cybersecurity learning and applying it to the realm of Supervisory Control and Data Acquisition (SCADA) systems. Practical exercises were designed which helped to maximise student education of cyber within this area while minimising the amount of time needed overseas at specialist training facilities.

Dopplick (2015) nicely sums up these worldwide trends in experiential cybersecurity learning: technical project-based activities, competitions, training and research are becoming commonplace as are universities "teaming with companies to provide structured programs on an ongoing basis" (84). Such exercises focus on providing simulated, controlled, safe, and legal, opportunities for student practice (National Institute of Standards and Technology [NIST] 2018a; NIST 2018b). Yet the recommendations of the USA National Initiative for Cybersecurity Education (NICE) exhort the need to challenge the assumptions and analyse the rationale for past, present, and proposed future cybersecurity education; and, inspire, explore, and experiment, with creative, innovative approaches to education, even to the degree that they might "disrupt or defy the status quo" (NIST 2016, n.p.).

The SB50 engagement serves as one such possible innovative or disruptive approach to cybersecurity education. A review of available literature indicates SB50 is one of the few times that students have been directly involved in providing cyber-capability for such a high-profile, high-risk event. At best, students have traditionally been involved in more of an 'observational' role with limited direct input. Many students have gained practical experiences while working in internships over a period of ten weeks or so (NIST 2018c). However, internships cannot guarantee intensive periods during which there is a heightened sense of imminent cyberattack. The intense focus of analysing data and responding in real time to possible threats is simulated in competitions (for example, CtF). However, knowing that it is indeed a simulation often leads to a cavalier approach that would not be acceptable under duress in a real-world attack-and-defend situation (for example, aggressively changing firewall rules that block the threat actor, but inadvertently also preventing real users from carrying out essential business activities). The NICE Cybersecurity Workforce framework itemises recommended or required cybersecurity skills and abilities related to communication, collaboration, and teamwork (**Table 1,** below). Many of these directly relate to the experiences noted and comments made by students who participated in the Super Bowl 50 project. However, nowhere does the NICE framework specifically address the need to work effectively under urgent decision-making conditions. Student participation in cybersecurity at an event such as Super Bowl 50 had the advantage of the students' experience of the heightened awareness that accompanies the threat of attack on an organisation, rather than a simulated threat.

Number	Description
S0070	Skill in talking to others to convey information effectively (note this does not really imply collaboration or communication, merely conveying information clearly)
S011	Skill at interfacing with customers
S0244	Skill in managing client relationships
S0301	Skill in writing about facts and ideas in a clear, convincing, and organized manner
S0341	Skill to articulate intelligence capabilities available to support execution of the plan
S0315	Skill to articulate the needs of joint planners to all-source analysts
S0343	Skill to orchestrate intelligence planning teams, coordinate collection and production support, and monitor status
S0356	Skill in communicating with all levels of management including Board members (such as interpersonal skills, approachability, effective listening skills, appropriate use of style and language for the audience).
A0011	Ability to answer questions in a clear and concise manner
A0013	Ability to communicate complex information, concepts, or ideas, in a confident and well- organized manner through verbal, written, and/or visual means.
A0074	Ability to collaborate effectively with others
A0076	Ability to coordinate and collaborate with analysts regarding surveillance requirements and essential information development
A0077	Ability to coordinate cyber operations with other organizational functions or support activities
A0078	Ability to coordinate, collaborate, and disseminate, information to subordinate, lateral, and higher-level, organizations.
A0082	Ability to effectively collaborate via virtual teams
A0098	Ability to participate as a member of planning teams, coordination groups, and task forces, as necessary.

Table 1. Skills and abilities from the NICE Cybersecurity Workforce Framework that are related to 'live-fire' educational opportunities (Newhouse et al. 2017)

SB50 Project Development

Initial contact and invitation

To identify smaller real-world engagement opportunities for capable students, faculty of the School of Business & Management CS and CISA programs had, for years, been leveraging contacts through alumni and through state, federal, and professional organisations. Early in the spring of 2015, a high-level alumnus reached out to a leading faculty member of the program as a Point Of Contact (POC) and invited cybersecurity student participation as SB50 observers. During the

remainder of the spring semester, a selected team of students led by the POC researched similar previous high-level events, evaluated what program students could potentially offer as value-added event support, and made site visit plans.

Initial site visits and proposed expanded role

As both an information-gathering and experiential-learning exercise, in early summer the faculty POC and student team travelled to the SB50 event site and observed site-security operations for several significant sporting and entertainment events. Using information gathered, prior research, and forecasts of possible added value, the team proposed providing several services to support event cybersecurity. Based on the quality of their preparation and the proposed services, the role of the student team to be fielded for the event was elevated from merely observing to actively contributing to the effort. Subsequently, additional site visits were conducted to test proposed services and associated proof-of-concept equipment.

Creating the team

This project presented unique real-world experiential learning opportunities for students, including a window into security issues for a SEAR-1 event; interaction with high-level security professionals; and provision of professional-quality cybersecurity support for an internationally-prominent event. The goals of the project were, therefore, to maximise the number of students who could appropriately participate, to maximise their opportunities to do so, and to stress professionalism and value-added quality for the event.

Due to the high-profile nature of the engagement (an 'academic exercise' for a significant, real-world, live-fire event), a senior faculty member acted as Project Manager (PM). The PM was responsible for project administration and troubleshooting; for guiding the project team; and for serving as project liaison with event leadership, University administration, and supporting vendors.

The nature and scope of the goals and services necessitated a team structure comprised of six subteams by area (Area Teams) with the following responsibilities:

- Operations—this group worked with the PM on areas such as team coordination and communication, and to clarify and communicate information of significance to SB50 decision-makers on the day of the event;
- Technical—these students were responsible for project system analysis and design, specialised hardware and software implementation, and database administration and management;
- Infrastructure—this group was responsible for the operation and maintenance of University research centre hardware and software that supported specialist project systems;
- Information Gathering—these students were in charge of surveillance, aggregation, and distillation of open-source information regarding the event and for identifying potential security issues;
- Site Security—this group focused on the University site and personnel security for operations during the engagement; and,
- Communications—these students served as public relations liaisons between Project/Area Teams, the University, SB50 leadership, and public media.

Students who had participated in the summer site visits, who had prior experience or skills, or who had demonstrated the capability for mature leadership were selected to lead the Area Teams (Team Leads). Several had already been contributing to the project since the initial spring alumnus contact; several stepped up at the start of the fall semester. Faculty with specific expertise acted as subject matter experts and POCs for the Area Teams.

Due to the sensitive nature of the engagement and the data that would be observed, the Communications Team was tasked with writing a first draft Non-Disclosure Agreement (NDA) with Levi's Stadium. The first draft was then refined by the PM in collaboration with them, with SCPD, University administration, and legal counsel. Because of the notoriety of SB50, it was also anticipated that many students would offer to participate on the Project Team regardless of their ability to do so. For example, students might have had the skills and maturity but not the free time unless they jeopardised their academic progress or other responsibilities. Others might have had the free time but lacked the maturity to maintain operational security or professionalism. The PM, therefore, wrote a Memorandum Of Understanding (MOU) to ensure students grasped and acknowledged the importance of balancing project participation with their other responsibilities, as well as the professional behaviour expected of all project team members.

Due to the unique learning opportunity presented by SB50, it was decided early on to include as many students in the experience as feasible. An open invitation was broadcast to all students and a large lecture hall was used to introduce the project. The PM and Team Leads presented the project opportunity and scope, the Area Teams, and the professionalism expected of participants. Interested students signed up, listed their preferred Area Teams, and noted any special skills or abilities they had that were relevant to the project. Team Leads identified those students best suited to supporting the various project areas, consulted with the PM, and issued invitations to the respective students. Sixty-five students, primarily CS and CSIA majors, formed the overall project team. Some students with a broad range of skill, ability, and interests participated in more than one support area. All project team members were given counselling on appropriate channels of communication regarding project inquiries and on maintaining operational security. They were also required to sign an NDA and MOU (which were then countersigned by the PM on behalf of the University).

Security operations and the many federal, state, local, and industry support groups at SEAR-1 events require considerable space, resulting in limited 'seating'. The Project Team was, therefore, structured for the eventuality of needing a small group on-site (dubbed the 'Away Team'). This consisted of an experienced CSIA faculty mentor, a staff member from university PR, and the minimum number of Team Leads needed to support on-site event cybersecurity. The rest of the Project Team that remained behind was dubbed the 'Home Team'. The Away Team was expected to provide direct cybersecurity support at the event site; act as liaisons, interpreting and coordinating requests to/from event leadership; and, coordinate and interpret information flow from the Home Team. The Home Team was responsible for providing behind-the-scenes support for the Away Team, for maintenance and operation of project infrastructure, for continuous monitoring of observed information, for addressing event leadership requests for information, and, for summarising and communicating back any information on potential threats or in response to direct requests.

Away Team members were Team Leads from Operations, Technical, and Communications. Consequently, these individuals took on the added responsibility of identifying and training 'seconds': individuals capable of leading the respective Home Team Area activities once the teams separated.

Pre-Event preparation

In the fall semester, several pieces of specialised equipment were installed on site and were tested both locally and remotely. During this same period, Project Team students conducted preparatory activities ranging from software, database, and infrastructure development, to learning about nuances managing PR for a major event. Dry-run exercises were conducted to develop and debug procedures, and to train and prepare students, teams, and team leaders.

For this event, the extensive preparation and run-up was interrupted for almost a month due to the winter break. The PM and several Team Leads remained active during the break. However, once

the spring semester began, it was necessary to reconvene the Project Team and re-establish team cohesiveness rapidly.

As part of event preparation, an early on-site Table-Top Exercise (TTX) was held to simulate various levels of incident criticality, and to surface security issues for consideration and resolution. The PM represented the university at the initial TTX and, with approval, communicated the simulation scenarios and key issues to the Project Team. To help students with responsibility for team leadership better understand the complexity of such multi-agency event support, they and the PM observed an actual cybersecurity TTX at the Vermont Emergency Operations Center. Finally, in preparation for SB50, those same students conducted a TTX to brainstorm and troubleshoot anticipated problems with supporting security for the live event, such as issues with hardware, software, or communication glitches; planned processes or procedures; communication and coordination within Area Teams; and coordination between the Home and Away Teams.

One week to go

The real-world exercise involved collaboration and communication at many levels, in different contexts, and with a variety of law enforcement and technical professionals as well as within student teams. This was especially true in the final week run up to, and the weekend surrounding the event.

To coordinate with event leadership and agency support, the Away Team travelled to the site a week before the event. The team coordinated with event leadership and representatives of various law enforcement groups, professional technical experts, and public relations personnel. Operations and Technical members of the team also collaborated with Levi's Stadium IT to assure technical and procedural readiness, while Communications members of the team supported event PR. During this time, the Home Team tested project systems, team readiness, and real-time communications with the Away Team by conducting daily dry-run sessions 6 PM to 10 PM Monday through Thursday.

During the event

The Home Team stood up full operations at 6 PM on the Friday before game day. Predetermined rosters of Home Team students working in each Area Team, supervised by the PM and several faculty POCs for Area Teams, supported shifts around the clock. During this time, the Away Team shifted roles. The Communications members acted as an on-site extension of the Information-Gathering team, while the Technical and Operations members monitored and evaluated data from site instruments and coordinated with the Home Team for requests from event leadership. In general, the Away Team was embedded with federal, state, and local law enforcement, public safety, and stadium technical personnel and stood ready to respond to specific tasks. During the event, several questions, issues, and items of interest arose; and requests from SB50 law enforcement to act on them were relayed from the Away Team Operations Leader to the appropriate Away Team or Home Team subgroup, and in some cases to both. Leaders for the Home Team received the relayed tasks and had to coordinate their specific team activities to generate results that were accurate, substantive, and timely. Furthermore, they had to work within the overall Team structure to provide clear, concise, appropriate responses in support of a high-risk, high-profile event. In short, they were under considerable pressure to collaborate and communicate, and to 'get it right' in many senses of that phrase. Event cybersecurity support was maintained for approximately 60 hours, until a post-game 'all clear' order from SB50 event leaders was received the following Monday morning.

Post-Event

Participation in the event generated local media interest. As a result, once operational security was no longer a concern, several Team Leads participated in local PR opportunities. The Technical and Infrastructure teams collated data collected from the event and created backup copies of the aggregated data. Finally, all Project Team students were invited to participate in a post-event questionnaire.

Evaluation Methodology

Engaging in cybersecurity at such a significant event certainly comes with its own rewards for students. Beyond kudos, however, it is important to consider how such an exercise lends itself to improving student Knowledge, Skills, and Abilities (KSAs), and their professionalism. To gauge the event's effectiveness as a learning tool, participants were invited to complete a survey with open-ended questions rather than having them select from a specific set of keywords or Likert scale choices.

The questions focused on three key areas. First, to identify what and where their skills came from, students were asked to evaluate their prior knowledge and level of preparedness for the event (Questions 1-3 in **Table 2**). Second, to identify their expectations for engaging in the project, students were asked what they had hoped to gain by participating (Question 4, **Table 2**). Third, to gauge benefits of participating, students were asked to reflect upon the experience and of their learning (Questions 5-6 in **Table 2**). To allow the cohort a reasonable period of reflection about what they personally achieved from the exercise, the survey was distributed at the end of the spring 2015 semester to all student participants in the SB50 event.

	Survey Questions					
Descriptive Terms	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**
Academic Classes	73.3					
Clubs & Societies	13.3					
CtF Competitions	6.7					
None	6.7	8.3	15.4			
Internships		16.7				
Communication		25				19.2
Organization		8.3				
Leadership		16.7	7.7	7.7		26.9
Technical		25	38.5	15.4		26.9
Practical Exp.			23.1	76.9		11.5
Confidence			15.4			
Teamwork					100	15.4

Table 2. Categories of terms used by students to describe their preparedness, expectations, and reflections regarding Super Bowl 50 event participation; all results are a percentage of each question (column) total
Legend: Q1 – Prior knowledge of cybersecurity; Q2 – Prior preparations; Q3 – Areas that were lacking; Q4 – Expected KSAs required; Q5 – Memorable moments; and, Q6 – KSAs gained post-event

Of the 65 students who took part in the Super Bowl 50 project, 25% responded to the questionnaire. The low response rate may have been due, in part, to a variety of circumstances: over half the project team consisted of first- and second-year students, there was no 'before' survey circulated prior to the event to set expectations of a follow-up, a response bias may have favoured those students who took on added responsibility as leaders of the various project team sub-groups, or the end-of-spring distribution point came at a time when students were either distracted by ongoing academic activities or so far post-event as to be less motivated to give feedback.

Responses for each question were examined for the descriptive terms used and the terms categorised. The gist of each question and the extrapolated answers are presented in Table 2, above. (Values shown are percentages of categories provided as answers.)

Discussion

Prior knowledge/preparedness (Questions 1-3)

Question 1 sought to elicit student perceptions about the cybersecurity profession, particularly the skills the students deemed most important. Question 2 and 3 together also sought similar information but were framed by the context of this particular project (Super Bowl 50).

When asked what they knew about cybersecurity as a profession before involvement with the exercise, most respondents used terms indicating technical skills obtained from lab-based academic

learning (73.3%), from extra-curricular activities (such as campus clubs and professional societies; 13.3%), from CtF experiences (6.7%). A small percentage (6.7%) indicated that they had no prior knowledge. Of interest was the near-unanimous focus on technical ability. There were no indications of any 'soft' skills, such as interpersonal collaboration or communication, as highlighted by the NICE Cybersecurity Workforce framework (**Table 1,** above).

When asked what KSAs prepared participants for the SB50 project, students still emphasised technical skill (25%), but also anticipated a need for those 'soft' skills, such as: communication (25%), leadership (16.7%), prior professional experience (internships; 16.7%), and organisational skills (8.3%). A few (8.3%) felt they had no KSAs that prepared them for the experience. Overall, technical ability now only accounted for a quarter of the responses (**Table 2,** above).

When considering areas in which they felt they were lacking, most of the respondents again focused on technical ability (38.5%) and prior professional experience (23.1%). Confidence (15.4%) and leadership (7.7%) were the only other soft-skill areas highlighted (**Table 2,** above).

Expectations of the event (Question 4)

Question 4 prompted students to describe the skills they believed were needed to undertake the SB50 project. Technical skills only accounted for 15.4% of the responses, with practical experience taking the largest percentage at 76.9%. References to practical experience included terms such as "real world experience", "how cyber plays into big events", "understanding security requirements", "working on a large scale project", and "how events are run and secured". Technical ability was still a core part of the answers, but the terms used also fit communication, organisation, and management abilities (**Table 2,** above).

Reflection on experience and learning (Questions 5 and 6)

Question 5 allowed the students to describe, in their own words, what was most memorable about being part of SB50. Question 6 sought the same information, prompting students to consider any KSAs obtained.

Students unanimously answered that the teamwork required was the most memorable aspect of working on SB50. Although some technical abilities were mentioned in passing, no respondent explicitly highlighted any new technical skills or abilities he or she acquired during the project. Asking the students to focus on KSAs they obtained helped explicate what "teamwork" meant to them in this context: technical and leadership skills (26.9% each), communication (19.2%), peer collaboration (15.4%), and practical experience (11.5%) (**Table 2**, above).

Results Interpretation

The purpose of this questionnaire was to assess if there was academic merit to students engaging in planning, organising, and participating in an event such as Super Bowl 50. Events such as these could be an innovative or disruptive approach to cybersecurity education by providing the focus and intensity of a CtF with the professional experience obtained during an internship.

Before participating in Super Bowl 50, students tended to consider cybersecurity an almost exclusively technical discipline. However, upon being asked what skills were needed for such a project, soft skills were identified (communication, leadership, organisation). The teamwork needed for an event of this magnitude and scope was, unanimously for students, their most salient memory. However, when asked to consider new KSAs obtained, only one quarter highlighted new technical abilities. The remainder highlighted skills expected of young professionals entering the workforce. The nature of the exercise, the team structure to cope with it, and the interaction necessary within the team and with external professionals from a variety of disciplines appear to have stressed the critical nature of collaboration and communication—the so-called 'soft skills' in cybersecurity.

Conclusions

This paper presented a unique opportunity undertaken by a student cohort—having assisted in defending cyberspace during Super Bowl 50. Students began working a year in advance on a wide range of event preparations that culminated in a focused, intensive week capped off by around-the-clock support before, during, and after the game. A review of existing literature did not reveal similar engagements in which students provided active support alongside law enforcement and other personnel; at most students have been given observational opportunities. Thus, this engagement serves as a potential example of an innovative new approach to cybersecurity education. Given the limited data, do such real-world event processes and opportunities merit recommendation to other academic institutions? Would students seeking employment in the cybersecurity sector gain knowledge, skills, or abilities that improved their ability to perform in industry, or would the exercise be simply a novel distraction?

The prior build-up to the event allowed students to experience the extensive range of advance thought, preparation, organisation, and management that can go into defining, creating, and implementing real-world cybersecurity for a professional engagement. There was a clear focus and intensity felt by participants very much akin to Capture-the-Flag competitions. Several students described it this way: "During this project I have never been so tired, so frustrated, so mad, so proud, or so happy"; "I gained technical skills, administration skills, leadership skills, communication skills, all of it"; "I got the shared sense of accomplishment and bonding"; and "I learned how to be flexible and multitasking and how to work when I was tired".

However, unlike CtF, intense cybersecurity coverage began well before the real-world event and continued until well after the event ended. Also, cavalier attitudes towards the engagement were discouraged by faculty mentors and recognised by students as a real-world risk. This tempered student actions. Through it all, students gained potentially valuable experience at something not mentioned in the NICE Cybersecurity Workforce Framework—a context of urgent, critical decision-making. The heightened sense of criticality related to the actions they performed and information they conveyed, as well as the impact on real-time decision-making and the time-pressure of the 'live-fire' event, provided the students the sort of novel, innovative, and possibly disruptive educational opportunity invited by the NICE Strategic Plan (NIST 2016). The process was also particularly scalable; the cohort included all computer science and cybersecurity students with the time and interest to participate, and even a handful of non-computing majors who wanted to learn more about cybersecurity through practical experience. Voluntary, anonymous feedback from student participants suggested they gained experiences like those that might be expected from

professional internships. Finally, their responses also suggested an apparent shift in perceptions from a tech-centric view of cybersecurity to one involving important, so-called 'softer' professional project skills. Engagement in a real-world situation emphasised the need for effective collaboration and communication. To the degree that exercises, scenarios, or simulations can emulate for students the same real-world teamwork pressures, they may impart similar student insights and help better prepare cybersecurity professionals for their future roles.

References

Abertay University 2016, 'Ethical hacking', viewed 23 December 2018, <https://web.archive.org/web/20160324125042/http://www.abertay.ac.uk/studying/ug/ ethhac/>.

Bajarin, T 2014, 'Meet Levi's Stadium, the most high-tech sports venue yet', *Time,* viewed 23 December 2018, < http://time.com/3136272/levis-stadium-tech/>.

Conklin, A 2007, 'The design of an information security practicum course', *Proceedings of the AIS SIG-ED IAIM Conference*, Montreal, CA.

Dopplick, R 2015, 'Experiential cybersecurity learning', *ACM Inroads*, vol. 6, no. 2, p. 84.

Government Communications Headquarters (GCHQ) 2011, 'Behind the code', viewed 23 December 2018, <http://www.canyoucrackit.co.uk/>.

Grossi, D 2014, *Mass gathering security: A look at the coordinated approach to Super Bowl XLVIII in New Jersey and other large-scale events*, U.S. House of Representatives, Committee on Homeland Security, Subcommittee on Emergency Preparedness, Response, and Communications, viewed 23 December 2018 <https://democrats-homeland.house.gov/sites/democrats.homeland.house.gov/files/sitedocuments/20140623094446-10021.pdf>.

Herr, C & Allen, D 2015, *Video games as a training tool to prepare the next generation of cyber warriors,* Carnegie Mellon University, Pittsburgh, PA, US, viewed 23 December 2018, <https://resources.sei.cmu.edu/asset_files/Presentation/ 2015_017_001_442344.pdf>.

Kessler, GC & Ramsay, J 2013, 'Paradigms for cybersecurity education in a homeland security program', *Journal of Homeland Security Education,* vol. 2, pp. 35-44, viewed 23 December 2018, <http://www.journalhse.org/v2-kesslerramsay.html>.

Martin, R 2016, *Super Bowl 50 tightens cybersecurity,* Vermont Public Radio, viewed 23 December 2018, <http://www.npr.org/2016/02/07/465901857/super-bowl-50-tightens-cybersecurity>.

Newhouse, W, Keith, S, Scribner, B & Witte, G 2017, *National Initiative for Cybersecurity Education (NICE) Cybersecurity Workforce Framework, National Institute of Standards and Technology (NIST) Special Publication 800-181,* Gaithersburg, MD, US, viewed 3 January 2019, <https://doi.org/10.6028/NIST.SP.800-181>.

National Institute of Standards and Technology (NIST) 2016, *National Initiative For Cybersecurity Education strategic plan,* Gaithersburg, MD, US, viewed 23 December 2018 <https://www.nist.gov/itl/applied-cybersecurity/nice/about/strategic-plan>.

——2018a, *NICE Cybersecurity competitions,* viewed 23 December 2018, <https://www.nist.gov/sites/default/files/documents/2018/09/24/cybersecurity_ competitions.pdf>.

——2018b, *NICE Cyber ranges,* viewed 23 December 2018, <https://www.nist.gov/sites/default/files/documents/2017/05/23/cyber_ranges_2017.pdf>.

——2018c, *NICE Cybersecurity apprenticeships,* viewed 23 December 2018, <https://www.nist.gov/sites/default/files/documents/2018/01/09/ nice_apprenticeship_one_pager_oct_31_2017.pdf>.

National Security Agency (NSA) Careers 2014, *tpfccdlfdtte pcaccplircdt dklpcfrp?qeiq lhpqlipqeodf gpwafopwprti izxndkiqpkii krirrifcapnc dxkdciqcafmd vkfpcadf,* #MissionMonday #NSA #news, Twitter, 5 May 2014, viewed 23 December 2018, <https://twitter.com/NSACareers/status/463321993878994945>.

Reed, K n.d., *User's manual for National Special Security Events (NSSE)/Special Event Assessment Rating events (SEAR) Job Aid*, viewed 23 December 2018, <https://homeport.uscg.mil/Lists/Content/Attachments/2718/Users%20Manual%20for%20NSSE%20Job%20Aid.pdf>.

Rege, A 2015, 'Multidisciplinary experiential learning for holistic cybersecurity education, research and evaluation', 2015 *Summit on Gaming, Games, and Gamification in Security Education – 3GSE, USENIX,* August 11, Washington, D.C., US, viewed 3 January 2019 <https://www.usenix.org/system/files/conference/3gse15/3gse15-rege-update.pdf>.

Sitnikova, E, Foo, E & Vaughn, RB 2013, 'The power of hands-on exercises in SCADA cybersecurity education', *Proceedings of the Information Assurance and Security Education and Training. WISE 2009, IFIP Advances in Information and Communication Technology,* vol 406, 8-10 July, Auckland, NZ, Springer, Berlin, DE, pp. 93-94.

United States Government n.d., *Crimes and criminal procedure: Powers, authorities, and duties of United States Secret Service,* 18 U.S.C. § 3056(e).

Willems, C & Meinel, C 2012, 'Online assessment for hands-on cybersecurity training in a virtual lab', *Proceedings of the 3rd IEEE Global Engineering Education Conference (EDUCON 2012),* IEEE Press, Marrakesh, MA.