Conference Proceedings of

17th Australian Cyber Warfare Conference (CWAR), October 10-11th, 2018, Melbourne, Victoria, Australia.

Edited by:
Matthew Warren
ISBN 978-0-6484570-0-8.

Proceedings of

CWAR 2018

Edited by
Matthew Warren
ISBN978-0-6484570-0-8.

Organised By:

Deakin University Centre for Cyber Security Research and Innovation

Sponsor:

Australian Information Security Association.

Published by the Deakin University Centre for Cyber Security Research and Innovation, Deakin University, Geelong, Victoria, Australia.

Welcome

The 17th Australian Cyber Warfare (formally Information Warfare) conferences builds upon a number of national conferences run across Australia namely in Perth, Melbourne, Geelong and Adelaide.

This conference looks at the continued development of Cyber Security within Australia, taking into the new emerging technologies and approaches.

Papers were selected for their relevance in relation to Cyber Security and the conference theme. The aim of this conference is to further the work already achieved within Australia and bring together researchers in the field to discuss the latest issues and their implications upon Australia. All full papers were reviewed by two members of the program committee. This year papers came from Australia (ACT, NSW, VIC, SA, WA), China, Maldives, Sri Lanka and Vietnam.

We commend the author for their hard work and sharing their results, and the reviewers of the conference for producing an excellent program.

CWAR 2018 Organising Committee

Conference Chair
Matthew Warren, Centre for Cyber Security Research and Innovation, Deakin University, Australia

Program Committee
Andy Jones, University of Hertfordshire, United Kingdom.
Atif Ahmad, University of Melbourne, Australia.
Bill Hutchinson, Edith Cowan University, Australia.
Darius Štitilis, Mykolas Romeris University, Lithuania.
Debi Ashenden ,University of Portsmouth, United Kingdom.
Graeme Pye, Deakin University, Australia.
Jeff Malone, Department of Defence, Australia
Jill Slay, La Trobe University, Australia.
Johan Van Niekerk, Nelson Mandela University, South Africa.
Mathews Nkhoma, RMIT Vietnam, Vietnam.
Martti Lehto ,University of Jyväskylä, Finland.
Mike Johnstone, Edith Cowan University, Australia.
Oliver Burmeister, Charles Sturt University, Australia.
Sean Maynard, University of Melbourne, Australia.
Shona Leitch, RMIT, Australia.
Steven Furnell, University of Plymouth, United Kingdom.
Paul Haskell-Dowland, Edith Cowan University, Australia.
Raymond Choo, University of Texas at San Antonio, United States of America.
Regina Valutyte, Mykolas Romeris University, Lithuania.

Contents

Targeting Hackers:
Big Data, Self-Fulfilling Prophecies, and Inescapable Errors

Shiri Krebs

Senior Lecturer and Director of HDR, Deakin Law School, Deakin University

Stanford Center on International Security and Cooperation (CISAC), affiliate

221 Burwood Highway, Burwood, Victoria, 3125, Australia

s.krebs@deakin.edu.au

Abstract

Cybersecurity measures have been increasingly relying on algorithmic predictions and calculations, as well as on big data analytics. While incorporating valuable information into security decision-making processes, these methods also entail several inherent weaknesses, which lead to the death of innocent civilians. This contribution focuses on the heuristics and biases affecting intelligence gathering and interpretation during targeted killing decision-making processes, and explains how organizational structures and socio-psychological dynamics may lead to intelligence failures and skew risk assessment processes. To analyse these processes, it employs interdisciplinary theories of risk assessment, organizational decision-making, and international law, finding that big data analytics involve values-infused predictions and interpretations which increase the risk of error, while producing a sense of robustness and clarity. Instead, it is recommended to redefine "data" for the purposes of preventive measures, to avoid transforming inconclusive information into "facts," and to mandate further investigation, where needed, rather than completing the missing information with predetermined conceptions and untested hypotheses. Importantly, the outputs of cyber surveillance should be questioned and re-evaluated to make sure individuals are not being killed based on misrepresentations, misguided evaluations, and self-fulfilling prophecies.

Secure Smart Contract Supply Chains

Richard Greene[1] and Michael N. Johnstone[1, 2]
[1] School of Science
[2] Security Research Institute
Edith Cowan University
Western Australia

rgreene0@our.ecu.edu.au , m.johnstone@ecu.edu.au

Abstract. *The concepts of information assurance and information warfare are often framed around digital artefacts. The fabrication of these artefacts and their consequent attachment to physical products provide interesting opportunities for assuring the provenance of physical products. We examine a case study where blockchain technology offers an innovative solution for Wine Industry supply chain governance, especially in regard to the identification of counterfeit product. This paper proposes a prototype solution to directly link the advantages of blockchain technology with the requirements of the Wine Industry. The results indicate that although blockchain-based supply chains are feasible, issues around authentication, verification and protection of digital assets persist. Our contribution was to investigate the opportunities offered by blockchain and smart contracting technology to provide greater product assurance than is currently offered in the wine industry.*

Keywords: Blockchain, Information Warfare, Network Security, Communications, Ethereum.

Introduction

Wine is a significant export driver for Australia. Wine Australia (2018) report that wine exports in 2016-17 represented a $2.3B value in sales, accounting for 61% of wine production by volume. In other words, more Australian wine is exported than is consumed domestically. Anderson (2015) states that the Asian market for wine is growing. He points out that two factors account for this trend. First, a growth in incomes in Asian countries and second, discerning consumers in the region are seeking quality imported wines. According to Anderson, Asia accounted for 4-5% of Australia's wine exports in the early 2000s, but since then its share has more than doubled. Further, he notes that Australia is second only to France in supplying wines to China, making China, Australia's third biggest market.

Therefore, overseas consumers are choosing quality merchandise from Australian wine producers. Much in the same way as French wine producers have moved to protect the use of certain regional names (e.g.,

Champagne, Chardonnay), Australian producers need to consider how to protect the brand of a quality product in a growing market.

According to Glase (2017), 40% of the wine produced in Australia is exported to mainland China. Further, China is now the second largest grape growing area under cultivation in the world. Vines are grown in provinces including Shandong, Hebei, and Tianjin, as well as many other regions. With rising wages and middle-income families, China has also become the sixth leading consumer of wine in the world, just behind Germany.

This rise in the availability, production levels and market for the product has however also led to an increase in the production of counterfeit wines. Cho Lee (2017) suggests that “as much as 50% of the fine wines in China are believed to be fake”. Although this comment focused mainly on the fine wine sector, the relatively low punishment and high rewards for the practice leads to easy entry into the (counterfeit) market. Accuracy in determining the size of the problem is difficult and little published research is extant. Wine commentators used statements such as “I can’t really say how much is on the market but looking at the wines I have seen, I think a fifth of all wines being fake may be an exaggeration” (Egan, 2017). An investigation into Intellectual Property rights infringements estimated that of the 400 wines he looked at, 50% of foreign wines were fake (Boyce, 2012).

For high value wines, studies have shown that even professional tasters are unreliable for identification of wine products. Hodgson (2008) indicated that only about 10% of wine judges were able to replicate scores given in competition conditions. Samples and chemical analysis can be taken in the wine preparation process however these have inconsistencies and issues with accuracy (McCharles and Pitman, 1936).

Poor record keeping of older wines also makes identification difficult (Hellman and Frank, 2009). Modern wine producers have better record keeping and tend to implement individual protection methods, such as laser etching, bank note style labels, security seals and even DNA seals. BRL Hardy, an Australian wine company employed DNA coding to authenticate its wine (Humphries, 2001), using hand scanners to read the DNA seal, similar to the process used to verify tickets for the Sydney Olympics. These efforts are, however expensive to implement on a large scale for smaller wine producers.

Fake wine is also not limited to simply counterfeits assuming the identity of existing products. Brian Smedley from the South Australian Wine Industry Association stated "The bigger problem we are looking at these days are wines which are not trying to necessarily copy a direct label of a well-known brand but are claiming to be from Australia." (Sedghi, 2018). To address this problem, the CSIRO, in collaboration with the Australian Wine

Research Institute, are currently attempting to develop a chemical isotopic approach that can accurately fingerprint Australian wine so it can be distinguished from wine from other major wine producing countries (Blackburn and Williams, 2017).

Nonetheless, given the size of the Australian wine export market, there is a need to protect Australian wine producers. The problem here is not about information warfare via denial of service, but concerns the spread of misinformation and the consequent second order effects ion the wine industry (which would lead to a lack of trust in the legitimate product, and resulting decline in sales/profits). This paper attempts to test the hypothesis that Ethereum Smart Contract technology can provide greater protections in the supply chain than existing methods.

Use of Smart Contract and Blockchain Technology

A blockchain is essentially a distributed, tamper-proof ledger. The blocks in the ledger could be currency transactions (as in Bitcoin) or smart contract elements (as in Ethereum). Whilst implementations may differ, the core concepts of transactions, blocks and consensus are common. Apart from the data and a time stamp, each transaction encodes a hash of the previous transaction, thus integrity is preserved as it is quite difficult for an interloper to change a transaction as the hash of all prior transactions in a block would need to be re-calculated and changed as well. Further the network of devices that hold the ledger is a peer-to-peer (P2P) structure, such that consensus must be reached for a transaction to be considered verified and written to a block.

The most common application of blockchain technology (apart from virtual currency) appears to be in supply chain management, driven by a lack of visibility of consignment data as a delivery moves through the supply chain (Miller, 2018; Lu and Xu, 2017).

Orman (2018) discusses the notion of blockchains applied to the universal personal identity problem. In our case, we seek to identify a specific instance of a product, rather than a person, but the principle remains the same.

Wessling et al. (2018) point out that there are several challenges to applying blockchains to extant systems, mostly because it is difficult to define which blockchain attributes are most important in a given system and cite anonymity and immutability as potential attributes of interest. An important part of their approach is identifying the trust relation and interactions between participants, a point also made by Ayed et al. (2014) in section 4.

Tikhomirov et al. (2018) state that the requirements for code analysis tools differ across platforms and domains, which is true. In some domains, not having false positives is crucial, whilst in others, it is false negatives that matter. They claim that in smart contract programming, a low false negative rate is crucial but a relatively high false positive rate is acceptable.

Currently Australian wine producers use the Label Integrity Program (LIP) process as shown in figure 1.

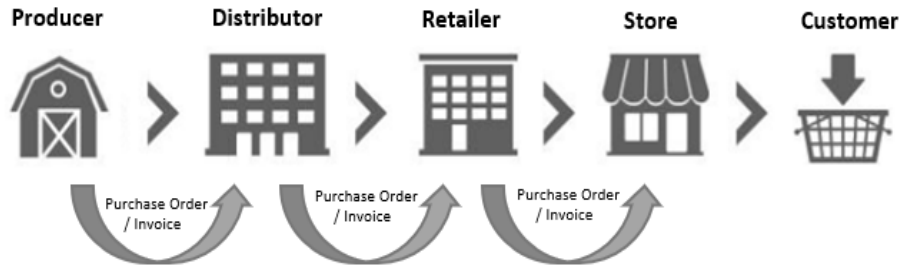

Fig 1: **Label Integrity Program Supply Chain.**

While the LIP process uses conventional data objects (e.g., invoices, purchase orders), the transactions involving these objects normally take place between two trusted trading partners, who have established some form of agreed standard of communication. An example data object in this context is a Wine Goods Supply Statement. Such a statement, however, is not necessarily provided in a standard format. While the consequences of non-compliance in Australia can range from instruction to relabel non-compliant products, to the cancellation of a person's licence to export and in extreme cases to prosecution and imprisonment (Wine Australia Act, 2013). However, the Wine Goods Supply Statement is not required outside Australia.

Smart contracts and blockchain technology is being used in other domains. Alharby and van Moorsel (2017) conducted a meta-study of twenty-four research papers related to smart contracts. These papers focused on identifying and tackling key issues such as codifying, security, privacy and performance. There were also papers related to the development of practical applications. In the commercial environment Everledger has two blockchain projects related to supply chain integration; The De Beers Blockchain Initiative which tracks diamonds, was started in 2015 (Lewis, 2018). Chai Vault is designed to give fine wine a unique ID. As the wine changes hands, provenance information can be updated, provided the transaction is initiated by a licensed user. Brokers, retailers, auction houses, and other sales platforms can then link to the information online to show the authentication details to potential buyers. This approach appears promising, but determining the authenticity of a "licensed" user is

problematic. Finally, smart contract technology is being used in the Chinese market by Walmart to track sales of pork meat (O'Byrne, 2017).

A Prototype Solution

The methodology used in this research will be based on an evaluative and developmental research process as developed by Nunamaker et al. (1990). This will involve the construction of a new software-based process to understand the outcomes and actions needed to meet requirements.

To address the issue of cross-border authentication, blockchain technology is used as a means of assuring trust between multiple partners across a distributed network (see figure 2). Blockchain systems assume that all verification nodes are equally untrusted, their proportion in the collective decision-making/transaction verification process is solely based on their computational resources, known as the Proof-of-work algorithm (Nakamoto, 2012). The process reasons that nodes that place significant resources into the system are less likely to cheat and good actors will be rewarded.

An Ethereum "Full node" is one that coordinates and controls the blockchain database. At least one full node is required to run the network with additional nodes being added at any stage. A smart contract is computational code written in low-level bytecode and stored on the blockchain, referred to as "Ethereum virtual machine code" or "EVM code" (Ethereum, n.d.). A higher-level language is available called Solidity, which can be compiled down to bytecode. Contract data storage is based on a key/value store which persists for the long term on the blockchain. Contracts can access values on execution such as the sender address and incoming data. Contracts are executed on all nodes within the network and consensus is reached when all validating nodes agree on the resulting state of the contract.

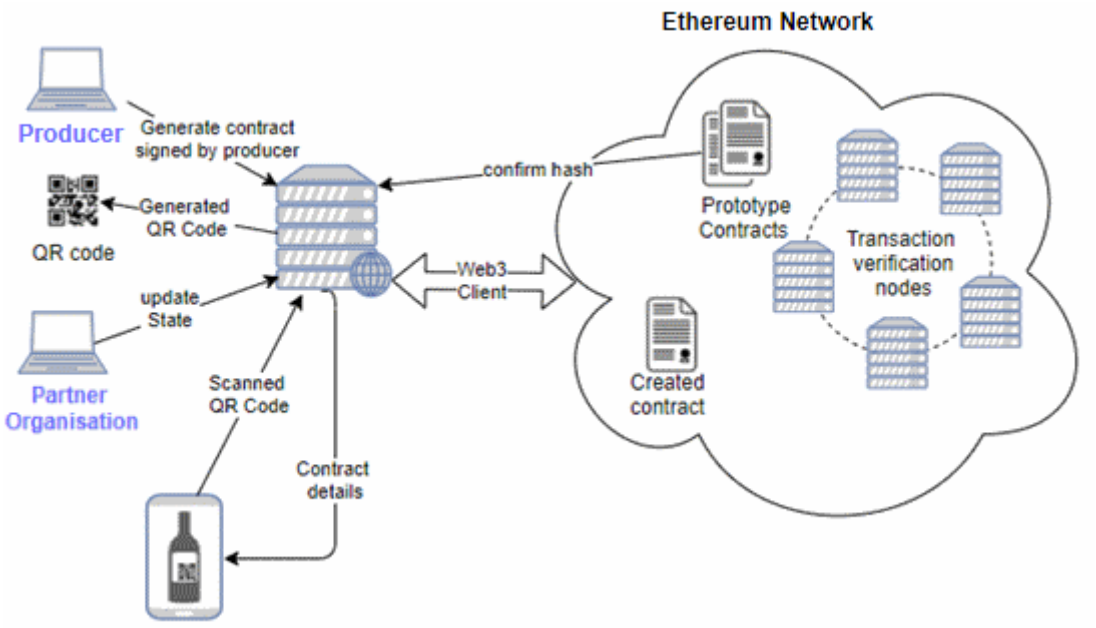

Fig 2: **Key Elements of the Design Prototype.**

In Figure 2, the process flow is as follows:

1. A Producer creates a wallet address on the blockchain, then uses this wallet address to create a new unique contract for the physical object (batch, bottle or cask of wine).
2. A contract is selected by the Web3 server from a verified list that best suits the requirements of the Producer.
3. The Producer enters all the required parameters and then cryptographically signs the request for a new contract using MetaMask and transmits this to the network.
4. When the transaction is recorded, a unique address is returned as a QR code that can be printed on the required back label.
5. The contract will accept any number updates in the form of transactions, if signed by an approved partner. All updates are recorded as data in a Merkle Patricia Tree.
6. Any individual with a connection to the network can scan the QR Code and query the contract for the verified current state, properties and all previous state transitions for the product.
7. Eventually the product can be set to a state of “Consumed”, locking the attributes and preventing any further changes to the contract.

Using the back label on a wine bottle is a cheaper alternative to more expensive tamper-proof options, such as laser etching of the wine bottle. According to the Wine Australia Act 2013 and accompanying Regulations, the Food Standards Code, the National Measurement Act, and the Consumer Act 2010, wine in Australia must contain a standard back label. When exporting however, language and cultural issues can cause

problems with back labelling, for example, China uses different codes for describing flavours, which don't necessarily match the standard Australian Wine Industry lexicon (Wine Intelligence, 2013). Using a QR code is a simple solution to this issue, as such codes are widely used across Asia (G.F, 2017).

In this model, as the product travels through the supply chain, ownership can be dynamically changed at any stage. Partner organisations, identified by their own unique wallet addresses, are permitted to update specific attributes on the contract as required (see figure 3).

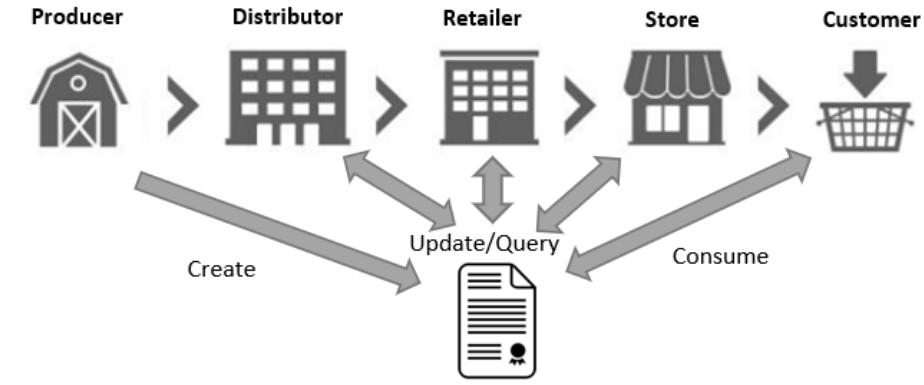

Fig 3: **Access by Various Parties to the Contract.**

The proposed system has a number of advantages over the current system.

1. Partners within the supply chain do not need to have a pre-agreed understanding with the supplying organisation. Verification is intrinsic within the blockchain infrastructure.
2. Any individual can query the attributes of a product and visually inspect the label to assure that the product has not been tampered with.
3. The unique identity of the product can be maintained and should the product QR code be duplicated, it would be clear at time of purchase that the supplier is not the one recorded on the blockchain.
4. Partners can read and update specific attributes as needed, without requiring paper records (or easily counterfeited conventional data objects).
5. Supply-chain information can be made available to all partners, providing deep insight into the movement of product.
6. The blockchain is resistant to tampering due to cryptographic signing of transactions.
7. Visual images of the product can be stored to assist with product identification.

Cyber Security Aspects of Smart Contracts

Libicki (1995) proposes seven types of information warfare. It is information economic warfare or war via the control of information trade, which is of interest here. Hutchinson and Warren (2001) point out that information is the product that has to be manipulated to the advantage of those trying to influence events. In the case of wine forgery, the product is the awareness of the value of the brand name of legitimate Australian wine and the event is the sale of inferior locally-produced wine (masquerading as high quality, imported wine), leading to a perception that Australian wine is inferior and then a second-order effect of decreased sales (exports), and possibly closure of Australian wineries, with attendant job losses in the wine and related industries.

Hutchinson and Warren (2001) note that if the target is the data, then potential information warfare actions that could be undertaken include denial of access, disruption or destruction, theft or manipulation. In our scenario, forgery is being attempted, so this could be considered theft (of the brand of the legitimate product) or manipulation (of customer's perception of the worth of their purchases).

Ayed et al. (2014) describe an implementation of an integrated authentication and authorisation framework to enable delivery of services across different organisations. As Ayed at al. point out, there is a need to securely share information with multiple, often independent parties, across organisational borders. A case in point is Figure 1, which shows business-critical information (in this case, a purchase order or invoice) traversing different organisations in a supply chain.

Ayed et al. identify the fundamental issue with such frameworks, viz. that adequate authentication and identity management systems must be used, else the parties will not trust each other. Theoretical solutions can, and have been, proposed, but implementation is a non-trivial undertaking, since the aforementioned parties use systems that belong to different security domains.

In terms of the oft-quoted principles of information security (confidentiality, integrity and availability), integrity is effectively assured by the difficulty of re-computing many hashes and having the network accept the changes, which, while not impossible, is quite difficult. This leaves attacks on confidentiality and availability as potential vectors, with the end points of the communication between parties as obvious starting points.

Denial of Service (DoS) is an attack against integrity that has proven successful in other domains (cf. the Mirai attack, which used IoT devices to perform a 1Tbit/sec distributed DoS attack). In figure 2, the transaction

verification nodes may be vulnerable to such an attack. This attack can be mitigated by having a large number of nodes, as each node is equivalent in this P2P model. Therefore, if some nodes were to be lost due to a DDoS attack, as no single node is a key point of failure, and the network is resilient, this would not unduly affect processing.

In terms of an example attack on confidentiality, Atzei et al. (2017, p173) point out that fields in contracts can be public and that simply declaring a field to be private does not guarantee that its value cannot be determined. To set the value of a field, users must send a suitable transaction which will be published on a blockchain. Since the blockchain is public (within the network), the contents of the transaction can be inspected and the value of said field inferred.

Conclusion

In this paper we investigated the opportunities offered by blockchain and smart contracting technology to provide greater product assurance than the current governance process, particularly in the Chinese market. We looked at the limited research available on the size of the issue of counterfeit wine and issues with investigating and identifying a fraudulent product. We described a prototype tool for improving the assurance processes, utilising a fully decentralised smart contract solution.

There are some issues with the current model as described. First, the network needs to be maintained by verification nodes. Until this number is sufficient and widely dispersed, the blockchain is susceptible to a denial of service attack. Second, in the current “Proof of Work” algorithm, verification nodes are rewarded for ensuring network integrity. Alternative reward models, such as “Proof of Stake”, might better suit this application and should be explored. Finally, the Ethereum contract code and network is susceptible to a number of attacks, noted in Atzei et al. (2017).

References

Alharby M, and van Moorsel A., (2017). Blockchain-based Smart Contracts: A Systematic Mapping Study. Fourth International Conference on Computer Science and Information Technology (CSIT-2017). pp125-140.

Anderson, K. (2015). Growth and Cycles in Australia’s Wine Industry: A Statistical Compendium, 1843 to 2013. University of Adelaide Press: Adelaide.

Atzei N., Bartoletti M. and Cimoli T. (2017). A Survey of Attacks on Ethereum Smart Contracts (SoK). Proceedings of the 6th International Conference on Principles of Security and Trust - Volume 10204. pp. 164-186.

Ayed, DF., Camenisch, J., Ignatenko, T., Johnstone, M., Koster, P., Lange, B., Petkovic, M., Sommer, D., Zic, J., (2014), Authentication and authorisation in entrusted unions. Proc. 12th Australian Information Security Management Conference, 56-63, Perth, W.A.
Blackburn R. and Williams T. (2017). Fingerprinting wine to prevent fraud, CSIRO, retrieved on 15th April 2018 from https://blogs.csiro.au/ecos/wine-fingerprinting/
Boyce J, (2012). Nick Bartman's investigation of fake wine, other IPR issues in China, Grape Wall of China, retrieved on 8th April 2018 from http://www.grapewallofchina.com/2012/06/03/matured-two-years-nick-bartmans-investigation-of-fake-wine-other-ipr-issues-in-china/
Byrne D., (2017, 28 June). The Wine Detectives [Audio podcast]. Retrieved from http://www.bbc.co.uk/programmes/b08vxv32
Cho Lee, J. (2017). How To Avoid Buying Fake Wine, Forbes, retrieved on 8th April 2018 from https://www.forbes.com/sites/jeanniecholee/2017/01/22/how-to-avoid-buying-fake-wine/#5788693e49da
Egan M., (2017). Liv-ex interview with Michael Egan, counterfeit wine expert, Liv-ex, retrieved on 8th April 2018 from https://www.liv-ex.com/2017/11/liv-ex-interview-michael-egan-counterfeit-wine-expert/Hodgson, R. (2008). An Examination of Judge Reliability at a major U.S. Wine Competition. Journal of Wine Economics, 3(2), 105-113. doi:10.1017/S1931436100001152
Ethereum, (n.d.). A Next-Generation Smart Contract and Decentralized Application Platform, White Paper, retrieved on 12th April 2018 from https://github.com/ethereum/wiki/wiki/White-Paper.
G.F., (2nd November 2017), Why QR codes are on the rise, The Economist, retrieved on 17th April 2018 from https://www.economist.com/blogs/economist-explains/2017/11/economist-explains-0
Glase, C. (2017). Australia's finest wines achieve record value in 2016, Wine Australia, retrieved on 6th April 2018 from https://www.wineaustralia.com/news/media-releases/december-2016-export-report
Hellman P. and Frank M, (2009). The Crusade Against Counterfeits, Wine Spectator retrieved on 8th April 2018 from http://www.winespectator.com/magazine/show/id/41140
Humphries J. (2001). World first for wine substitution, ABC Rural News, retrieved on 5th April 2018 from http://www.abc.net.au/site-archive/rural/news/stories/s316563.htm
Hutchinson, W. and Warren, M. (2001). Principles of Information Warfare. Journal of Information Warfare, 1(1), pp. 1-6.
Kennedy S. (2015). Made in China 2025, Centre for Strategic and International Studies, retrieved on 8th April 2018 from https://www.csis.org/analysis/made-china-2025

Lewis B. (2018). De Beers turns to blockchain to guarantee diamond purity, Reuters, retrieved on 17th April 2018 from https://www.reuters.com/article/us-anglo-debeers-blockchain/de-beers-turns-to-blockchain-to-guarantee-diamond-purity-idUSKBN1F51HV
Libicki, M. (1995). What is Information Warfare? Strategic Forum Number 28, Available at: http://www.dodccrp.org/files/Libicki_What_Is.pdf
Lu, Q. and Xu, X. (2017). Adaptable Blockchain-Based Systems: A Case Study for Product Traceability. IEEE Software. 34(6), pp. 21-27.
McCharles, C.H. and Pitman, G.A. (1936). Methods of wine analysis. Industrial & Engineering Chemistry Analytical Edition, 8 (1), 55-56. DOI: 10.1021/ac50099a025
Miller, D. (2018). Blockchain and the Internet of Things in the Industrial Sector. IT Professional. 20(3), pp. 15-18.
Nakamoto S. (2012). Bitcoin: A peer-to-peer electronic cash system, Bitcoin.org, retrieved on 16th April 2018 from https://bitcoin.org/bitcoin.pdf
Nunamaker J. F., Chen M., and Purdin, T. (1990). Systems Development in Information Systems Research. Journal of Management Information Systems. 7(3), pp89-106.
O'Byrne, R., (2017). How Blockchain Can Transform the Supply Chain, Logistics Bureau, retrieved on 11th April 2018 from https://www.logisticsbureau.com/how-blockchain-can-transform-the-supply-chain/
Orman, H. (2018). Blockchain: the Emperor's New PKI? IEEE Internet Computing. 22(2), pp.23-28.
Sedghi, S., (2018). New way to identify wine will protect Australian growers from overseas fraud, scientists say, ABC, retrieved on 14th April 2018 from http://www.abc.net.au/news/2018-01-02/scientists-developing-new-techniques-to-detect-wine-fraud/9298186
Tikhomirov, S., Voskresenskaya, E., Ivanitskiy, I., Takhaviev, R., Marchenko, E. and Alexandrov, Y. (2018). SmartCheck: Static Analysis of Ethereum Smart Contracts. Proc. ACM/IEEE 1st International Workshop on Emerging Trends in Software Engineering for Blockchain (WETSEB 2018). pp. 9-16.
Wessling, F., Ehmke, C., Hesenius, M. and Gruhn, V. (2018). How Much Blockchain Do You Need? Towards a Concept for Building Hybrid DApp Architectures. Proc. ACM/IEEE 1st International Workshop on Emerging Trends in Software Engineering for Blockchain (WETSEB 2018). pp. 44-47.
Wine Australia. (2018). Australian wine: Production, sales and inventory 2016–17. Report.

Cyber-Enabled Information and Influence Warfare and Manipulation: Detection and Response

Professor Jill Slay
La Trobe Optus Cyber Security research Hub
La Trobe University, BUNDOORA, Vic
j.slay@latrobe.edu.au

Abstract
This paper examines the call for new theoretical frameworks for the detection of, and response to, cyber-enabled information/influence warfare and manipulation (IIWAM). It examines frameworks developed for more traditional IW and Deception in the early 2000s and argues that these are still valid for the increased speed and scope of IIWAM and in the development of the tools and information that underpin this activity.

Keywords: Culture, information/influence warfare and manipulation

Introduction

There have been both recent disciplinary interest and funding calls for participation in research that examines the development of techniques and approaches to information manipulation and deceptive behaviour attribution. There is also a stated need to develop technical approaches to countering this cyber-enabled information/influence warfare and manipulation (IIWAM) by preventing, detecting and responding to deception and confusion methods. We thus see an academic response to the now common phenomenon of online manipulation of information, especially mediated by social media. 'Fake news' has been weaponised.

While there might be planning in the long term to identify an enemy, or competitor's, use of IIWAM and also begin to investigate how we might produce our own 'offensive' response to this information manipulation and deception, there is very little solid disciplinary or cross-disciplinary research available to the technical or socio-technical researcher who wants to engage with this kind of research arena. The academic and public domain literature which provide a basis for a technological response is scant and spread across many overlapping disciplines.

It is claimed that modern IIWAM differs in form to traditional deception, propaganda and earlier Information Warfare techniques because of the potential scale at which it is being organized, the complexity of the targets it seeks to influence and the impact it has been having on society and even secure information targets in recent times. One can assert here that this is not necessarily a new problem and we have academic and practitioner literature that allows us to structure a theoretical approach.

It is suggested that automated development of online text is not new and that the real question is how can we be sure that the potential reader be persuaded that our 'fake' news is 'real'. Also techniques need to be developed since it may be asserted that there is a heightened awareness

of the potential of ‘fake news’ and our deceptive techniques need to be much more focused with a deeper understanding of the audience for, or target of, our deception. This means that the basic issue is worldview and culture of the target

Developing a Theoretical Basis for the Identification of Information / Influence Warfare and Manipulation

Culture and its relationship with cyber enabled ‘fake’ information

It is assumed in this paper that culture and worldview are the major factors that affect the way we think about concepts that underpin technology. Based on the well-established work of Slay (2002, 2004) it is still asserted that, with Straub (2002)

> “*it is the lack of clear concepts which makes cross-cultural research in the engineering of complex information systems difficult to conduct, and also links the effect of this lack of clarity to our inability to ‘develop and refine theories’ and to explain why there is difficulty in explaining the high degree of variance in current predictive models.*”.

The author thus asserts that IIWAM is influenced very much by culture since it is, at a basic level, a product of the engineering of complex information systems, as identified by Straub. The culture under examination might be that of a terrorist group, political party or nation state but there will be certain features of all such groups which can be identified.

Culture is a term that is perceived broadly in technological fields to play a major role in joint operations, project management, and any type of collaboration in a technological and cyber-mediated context. Culture is also very hard for the technologist to define, explain, compensate for or measure. Many attempts have been made by engineers and computing professionals to incorporate these factors into other more technically focused exploration.

As previously stated (Lin *et al* 2008; Hofstede 1980), one of the earlier modern researchers on culture and technology, proposed a term ‘national culture’ and set five dimensions for measuring national culture, and used them to evaluate cultural metrics or scores for fifty-three different countries and three multi-country regions (Hofstede 2007). Based on the explanations in Hofstede and Bond (1988), the scores are relative, and the distance between lowest-score and highest-score is about 100 points. A thorough explanation of the five national cultural dimensions is found in Hofstede (1994), and definitions of 4 useful ones are listed as follows. It should be stated here that there is often modern criticism of Hofstede and his work but I use it simply as a well-established metric to approximate worldview and culture.

1. Power Distance: According to this definition, power and inequality are two factors which can measure the cultural differences between different nations and relate to hierarchy within a specific culture. It is

thus possible to consider the extent to which a cultural expectation is one of hierarchy within 'fake' news

2. Individualism versus Collectivism: According to this definition, the meaning of individualism shows that an individual tends to behave independently. On the contrary, collectivism points out that an individual tends to seek identification within a particular group. Hofstede uses this feature as a metric for a national culture. Factors of expectation of collective thinking could be used as a metric for cultural acceptance of 'fake' news.
3. Uncertainty Avoidance: It is reasonable for people to have different level of uncertainty avoidance. Some people are willing to risk the unknown, but others are afraid of failure. Consequently, it is understandable this idea is one out of five dimensions which may be able to measure the concept of national culture and be used as a metric for response to or attribution of fake news,
4. Long-Term Orientation versus Short-Term Orientation: Some cultures have a Short-term focus in planning and belief while others have a long term perspective. 'Fake' news should thus contain elements of the cultural orientation of the target group.

The terms 'worldview' and culture are often used interchangeably. It is important to realise that, once 'culture' has been defined, then members of a particular cultural group will share a common worldview. The term 'world view' has two different connotations in English. The first has a philosophical meaning and involves a person's concepts of human existence and reality; the second is an individual's picture of the world that he or she lives in. The term 'world view' as used in anthropology refers to the 'culturally-dependent, implicit, fundamental organisations of the mind (Cobern, 1991, p.19).

Kearney's (1984) model of world view presumes a logical and structural integration of presuppositions within any individual and therefore the model is known as a logico-structural one. He then identifies seven logico-structural categories contained within a given individual's world view:

- The Other
- Classification
- Causality
- Relationship
- Self
- Time & Space

These categories serve as a framework for analysis of a world view. Kearney (1984, p.65) draws the parallel between these factors of an individual world view and the categories a doctor uses for the diagnosis of a patient's disease. In order to determine the world view of an individual, his or her understanding of the seven categories of Other, Self, Time & Space, etc., need to be identified and integrated to produce a picture of the complete world view.

Thus, an understanding of how a specific culture and worldview affects belief that cyber-mediated 'fake' news is either true or fake will to a greater or lesser extent be determined by the understanding of, and the ability to measure the effects of that specific culture and worldview. A combination of the worldview Theory of Kearney (as interpreted in Slay (2009)) and that of Hofstede gives a set of metrics which may be used and adapted for any particular group so as to understand the factors which affect that group's tendency to accept cyber-mediated 'fake' news as true or false and also in the attribution of fake news to a specific group.

These worldview approaches, differentiating by factors listed below, allow us to model tools on cultural factors which are also detailed in Hofstede's work of 1983 onwards.

- Problem solving skills
- Motivation
- Leadership style
- Individual feelings about self
- Attitude toward gender
- Time
- Ability to work in a group

Deception

I align with the model proposed by Hutchinson (2006) in his work on Information Warfare and Deception and the concept that 'for a successful deception there must be an objective (to measure your success by), a target audience (to choose the applicable means of deception), a story (as a vehicle for the deception), and a means.'

If we consider the Internet and the known issues in development of cyber-mediated information to create deception, this can be derived:

- An objective - information created TO produce response that can be determined as deceptive in desired group
- A target – information can be adjusted for a given target group
- A story – story behind the information can be easily edited for target group
- Means – the information itself, and the tools used to produce the information

Thus the author's thinking can be incorporated and the thinking of means to provide a deceptive response in the context of IIWAM Hutchinson's (2006) analysis of RAND's model of the Deception Planning process (Gerwehr, S., & Glenn, R.W. (2000)).

- Means (tactics) of deception:
- Camouflage/ concealment/ cover;
- Demonstration/feint/diversion;
- Display/decoy/dummy;
- Mimicry/spoofing;
- Dazzling/sensory saturation;

- Disinformation/ruse;
- Conditioning

Hutchinson sees deceptive defensive IW (in this paper it is extended to deceptive IIWAM) as :

> *'Presenting data to the adversary that represents the truth as you would want them to perceive it. This is achieved by presenting a tailored subset of 'real' data, and/or manipulated data, and/or depriving the foe of any data, and/or disrupting the foe's data collection'*

Hutchinson then presents the concept then as:

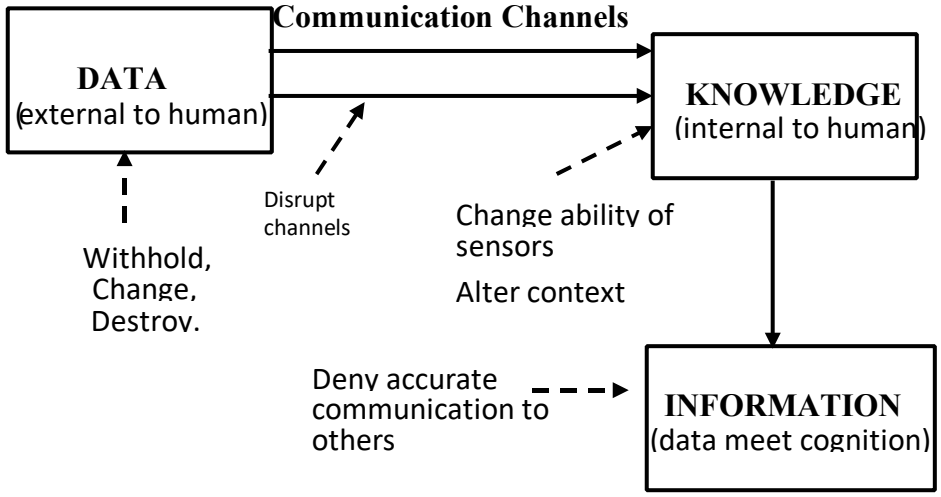

Figure 1: The relationships between data, context, knowledge, information; and the methods by which each element can be attacked to cause deception and corrupted information. (Hutchinson 2006)

Culture, Deception and IIWAM

It has been stated here that there is a need to develop a framework to understand how IIWAM can be detected, how cyber-mediated deception and 'fake' news can be attributed and, perhaps, how it can be created.

This paper has presented the following based on academic and other research in this context. This research draws out the following:

- An understanding of culture and Worldview Theory can provide a diagnostic framework, developed from the work of Kearney and Hofstede. This is useful since target groups, or victims, may share common identifiable features that can be used or manipulated to produce a belief in 'fake' news or deception. This diagnostic framework is an important offensive or defensive mechanism and can accompany attribution or targeting.

- Hutchinson's work (2006) on Information warfare and Deception is not outmoded even in an environment where speed and scale are larger than the context in which he was originally working. His analysis of RAND's model of deception planning holds well for the planning and attribution of IIWAM

Conclusion

The purpose of this work has been simply to understand if new theoretical foundations on development of techniques and approaches to information manipulation and deceptive behaviour attribution are needed. Existing literature of the field has been examined and IT IS believed that perspectives from Culture and worldview which have already been applied to IW and concepts of Deception as they relate to the IW of the earlier 2000's can be conflated to produce an overarching framework for the attribution and development of IIWAM. This topic will be covered in further work.

References

- Cobern, W.W. (1991). World view theory and science education research. NARST Monograph No. 3. Manhattan, KS: National Association for Research in Science Teaching.
- Gerwehr, S., & Glenn, R.W. (2000). Unweaving the web: Deception and adaptation in future urban operations. Santa Monica, CA: RAND.
- Hofstede, G. (1983). Dimensions Of National Cultures In Fifity Countries And Three Regions. In J.B. Deregowski, S. Dziurwiec and R.C. Annis (Eds.). Expiscations in Cross-Cultural Psychology. Lisse, Netherlands: Swets& Zeitlinger.
- Hutchinson, W, (2006) Informing Science: The International Journal of an Emerging Transdiscipline. Volume 9 . pp. 213-223. https://doi.org/10.28945/480
- Lin, YC, Slay, J & Lin, YL, 2008, 'Computer forensics and Culture, Pacific Asia Workshop on Cybercrime and Computer Forensics at ISI 2008, Taipei, Taiwan June 20 2008.
- Kearney, M. (1984). World view. Novalto, CA: Chandler & Sharp.
- Slay, J 2002, 'A Cultural Framework for the Interoperability of C2 systems', in Proceedings of 7th International ICCRTS conference ,September 11th – 13th, Quebec City, Quebec.
- Slay, J & Quirchmayr, G 2004, 'A Formal Model for the Relationship between Culture and Trust within IS Security Management, in Proceedings of the 2nd Australian Information Security Management Conference, Perth, November 2004.
- Straub, D., Loch, K., Evaristo, R., Karahanna, E. & Strite, M. 2002. Toward a Theory-Based Measurement of Culture. Journal of Global Information Management. 10(1), 13-23.

Risks of Critical Infrastructure Adoption of Cloud Computing within Government

Mansoor Al-Gharibi, Matthew Warren and William Yeoh
Deakin University Centre for Cyber Security Research and Innovation,
Deakin University, Geelong, Victoria, Australia.

malghari@deakin.edu.au, matthew.warren@deakin.edu.au,
william.yeoh@deakin.edu.au

Abstract: The purpose of this paper is to discuss the risks and reasons of adopting the cloud computing in the critical infrastructure within the government context of cloud computing adoption. The paper will also present examples of cloud computing adoption in the critical infrastructure domain. The data used in the paper was gathered from different academic, governmental and online sources. It was found that, although there are risks involved in the cloud computing adoption, governments are deploying cloud computing using different deployment models and reaching high level of deployment within the critical infrastructure. The findings of this study suggest that it is not a question of adopting or not anymore but the question of how to mitigate the risks involved after the deployment.

Keyword: Critical Infrastructure, Cloud Computing Adoption, E-Government, Cybersecurity Warfare.

Introduction:

Cloud computing is a relatively new service based information technology model that is completely or partially replacing the in-house IT delivery model. There are four adoption models which are public, private, hybrid and community cloud computing and in public sector many countries have developed what is called governmental cloud (g-cloud) for inter government engagement and engagement with citizens.

Cloud computing helps to reduce the cost of IT infrastructure and services, improve productivity and efficiency and provide on-demand services. In addition, cloud computing adoption lets the organisations focus on their core businesses and let the IT burden be on the cloud computing providers and the nature of cloud computing provide the flexibility of starting small and grow or more precise resize as per the business needs.

In private sector companies are shifting from the in-house asset based IT model to in or out house or both cloud computing service based model. In the same way, public sector is following the same movement and many governments have included cloud computing adoption in their IT strategies. Alford and Morton (2009) stated that government can achieve 50% to 67% cost saving by moving governmental applications to public or private clouds.

With all the temptation of the new technology within government, shifting the governmental critical infrastructure to the cloud is a challenge and the

cloud computing and the critical infrastructure characteristics invite vulnerabilities that can be used in cyber warfare. Despite that, different countries approached the adoption in different ways and they have different polices to mitigate the risks involved in the adoption.

In the United States they use cloud services provided by private companies (Ali 2015; Kundra 2011) while other governments developed or developing their government-cloud either internally like Oman (Information Technlogy Authority 2018) or via dedicated infrastructure developed by private companies like Australia (Microsoft). Thus, it is not a question of adopting or not anymore but the question of how to mitigate the risks involved after the deployment.

Literature review:

Critical Infrastructures:

Cyber threats started to receive attention by the government of the United States in 1996 when information and communication technology (ICT) dependency started to grow and a president commission was formed to report about threats to critical infrastructure with focus on cyber threats (Harašta 2018). In 1998, the United Nations recognized the existing and potential issues and threats of information warfare in global information and communication systems (Pye & Warren 2009). In the post 9/11 era it gained further attention and the European Union formulated a definition of an attack on critical infrastructure as, “causing extensive destruction of a Government or public facility, a transport system, an infrastructure facility, including an information system, a fixed platform located on the continental shelf, a public place or private property likely to endanger human life or result in major economic loss” (Harašta 2018).

The US government defined the critical infrastructure as, “systems and assets, whether physical of virtual, so vital to the United States that the incapacity or destruction of such systems and assets would have a debilitating impact on security, national economic security, national public health or safety, of any combination of those matters” (Cazorla, Alcaraz & Lopez 2018; Harašta 2018). The European Union define Critical Infrastructure as, “an asset, system or part thereof located in Member States which is essential for the maintenance of vital societal functions, health, safety, security, economic or social well-being of people, and the disruption or destruction of which would have a significant impact in a Member State as a result of the failure to maintain those functions” (Cazorla, Alcaraz & Lopez 2018). In Australia it is defined as, “those physical facilities, supply chains, information technologies and communication networks which, if destroyed, degraded or rendered unavailable for an extended period, would significantly impact upon the social or economic well-being of the nation or affect Australia’s ability to conduct national defence and ensure national security” (Pye & Warren 2007). Although the wordings are different, they all agree on the destruction of certain systems lead to unfavourable consequences on different aspects of social life. That is, a critical infrastructure is an infrastructure or systems when destructed for extended period of time have impact on security,

economy, health, safety and social well-being. By the definitions, the critical infrastructure spread across different sectors.

According to Dunn and Wigert (cited in Pye & Warren 2009), critical infrastructure sectors in modern societies are finance, food supply, health, government services, law and order, manufacturing, national icons, transport, water, and water waste. According to Brown, Seville and Vargo (2017), the critical infrastructure service includes water, waste water, telecommunications, energy and transportation and other services. They are also stating that there is high interdependency between the critical infrastructure systems and vulnerable to cascading failures. The critical infrastructure systems are dynamic systems and reliant and influence each other and necessary to function together in dynamic way to supply the service normally (Pye & Warren 2007). The destruction to single system has cascading effects to other systems within the critical infrastructure.

There are different sources for destruction of the critical infrastructure. It can be destroyed, damaged or disrupted by breakdowns, negligence, accidents (Pye & Warren 2007) natural disasters and extreme weather conditions (Pye & Warren 2007; Tsavdaroglou et al. 2018). In addition, the critical infrastructure can be impacted by human factors such as social engineering techniques (Ghafir et al. 2018). Since most of critical infrastructure systems are based on information and communication technologies, cyber incidents in relation to critical infrastructure can be a target for both conventional and information warfare (Cazorla, Alcaraz & Lopez 2018; Pye & Warren 2007).

Cloud Computing:

There are many definitions for cloud computing by different experts and academicians and these definitions vary in the key characteristics they identified (Madhavaiah, Bashir & Shafi 2012). The National Institute of Standards and Technology (NIST) defines cloud computing as, “a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction” (Mell & Grance 2011).

Cloud Computing Characteristics:

The cloud computing essential characteristics can differentiate it from the old information technology models and NIST has defined five essential characteristics which are: on-demand self-service, broad network access, resource pooling, rapid elasticity and measured service (Mell & Grance 2011). Hurwitz & Associates (cited in Madhavaiah, Bashir & Shafi 2012) considered elasticity and scalability, self-service provisioning, standardized application program interfaces (APIs), billing and metering of services, performance monitoring and measuring, and security as the key characteristics of cloud computing. These are:

1. On-demand self-service: is the ability for the user to be provided the requested computing capabilities without the need of human interaction;

2. Broad network access: is the accessibility of the cloud computing capabilities by different devices over a network connection;
3. Resource Pooling: is the pooling of providers' computing resources to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned according to the consumer demand;
4. Rapid Elasticity: is the flexibility of the provisioned capabilities to rapidly increase and decrease based on the demand;
5. Measured Service: is the appropriate measuring of service provided so resources used can be monitored, controlled and reported transparently for providers and consumers of the cloud services.

Cloud Computing Service Types:

There are mainly three services types of cloud computing which are infrastructure as a service (IaaS), platform as a service (PaaS) and software as a service (SaaS) (Nirenjena et al. 2017). Each of them provides a different set of services to cater for the needs of different type of users for example SaaS includes enterprise resource planning (ERP) software, PaaS includes database platforms and IaaS includes servers' usage.

IAAS

Infrastructure as a service is the base for the other two layers and provides the storage and compute capabilities for example: servers, switches and storage systems (Nirenjena et al. 2017). NIST define it as, "The capability provided to the consumer is to provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications." (Mell & Grance 2011).

PAAS

Platform as a service is providing the users with the development environment that covers the software lifecycle in which the developers can develop complete applications. In this layer the service providers have taken care of the needed infrastructure and resources needed for development (Nirenjena et al. 2017; Senarathna 2016). NIST define it as, "The capability provided to the consumer is to deploy onto the cloud infrastructure consumer-created or acquired applications created using programming languages, libraries, services, and tools supported by the provider." (Mell & Grance 2011).

SAAS

Software as a service is the upper layer of the other two services and it allows the users to install and use applications in the service providers severs and access them from anywhere with internet connection (Nirenjena et al. 2017; Senarathna 2016). NIST define it as, "The capability provided to the consumer is to use the provider's applications running on a cloud infrastructure" (Mell & Grance 2011).

Cloud Computing Deployment Models

There are four deployment models to deliver the cloud computing services which are public cloud, private cloud, hybrid cloud and community cloud

(Mell & Grance 2011; Senarathna et al. 2016). These deployment models refer to who are utilizing the cloud services infrastructure or dedicated part of it. For example in private cloud setup the resources are dedicated to one customer only while public cloud setup are shared among different unrelated customers.

Public Cloud

Public cloud models provides cloud computing services to the public and the cloud computing service provider has the control over the infrastructure used (Hsu, Ray & Li-Hsieh 2014). In this model one hardware can be shared between many organisations and the users pay for the service based on their usage or periodical subscriptions (Nirenjena et al. 2017).

Private Cloud

Private cloud models provides cloud computing services to one organisation and it has greater control over the hardware used which is available only for that particular organisation. With this model the infrastructure is either hosted on or off the organisation premises and controlled by the organisation itself or the third party cloud service provider (Hsu, Ray & Li-Hsieh 2014; Senarathna et al. 2016).

Hybrid Cloud

Hybrid cloud models are a mixed model of private and public cloud where organizations keep part of its computing in private cloud and another part in public cloud (Mell & Grance 2011; Senarathna 2016). This model is to help organizations keep higher level of control over critical computing activities in private cloud without losing the cost advantage of public cloud for other computing activities (Ali, Warren & Mathiassen 2017).

Community Cloud

Community cloud models are a cloud computing infrastructure shared and controlled by group of organizations serving common community with same concerns (Marston et al. 2011; Senarathna et al. 2016). The infrastructure management can be dedicated to one or many organizations within the community or dedicated to a third party (Senarathna 2016). An example of community cloud is government cloud (G-Cloud) which is a community or private cloud specifically designed for national government use (Zwattendorfer et al. 2013).

E-Government

History

According to Grönlund and Horan (2005) the e-government term emerged in the late 1990s but computing in government is as old as computer is. E-government is a complex change efforts to use technologies to support transforming the operation and effectiveness of government (Guo 2010) and it is gaining ground in both developed and developing countries (Bwalya & Mutula 2014). The advancement of e-government oriented technologies and services noticeably fast and the attention is shifting

toward cloud computing and use it in the e-government services (Zhang & Chen 2010).

Definition

There are many definitions for e-government by different authors. According to Guo (2010) e-government is the way for governments to use most innovative ICT, to provide citizens and businesses more convenient access to government information and services, to improve their service quality, and greater opportunities provision to participate in democratic institutions and processes. Bwalya and Mutula (2014) define it as, “the use of Information and Communications Technologies (ICTs) in public delivery frameworks” and Schnoll (2015) describe it as using ICT to support governmental functions and services, citizens to participate in the political processes. In general the definitions are agreeing on the usage of innovative ICT services to support and provide citizens and businesses access to quality and efficient governmental information and services.

Critical Infrastructure in the Cloud; Examples of the US National and Victorian State Governments

Reasons

According to Kundra (2011), cloud computing potentially will address the inefficiency and long procurement lead times of the current federal government IT environment in USA and respond faster to the constituent needs. The federal cloud computing strategy highlighting three categories of cloud computing adoption benefits to the government which are efficiency, agility and innovation. Similarly in Victoria in Australia the government want to increase its efficiency and reduce cost by adoption cloud computing (Victorian Government Solicitor's Office 2011). In addition, critical infrastructure sectors are migrating to cloud computing to realize benefits such as scalability, high availability and decreased ownership cost (Office of Cyber and Infrastructure Analysis 2017). Cloud computing can lead to faster delivery pace, continuous improvement cycles, broad services access, reduce maintenance effort and refocus that effort to improve service delivery (Digital Transformation Agency 2017).

Examples

Many countries are adopting cloud computing in for their critical infrastructure. In USA they have instituted a Cloud First policy (Kundra 2011) and the percentages of information technology systems that have adopted cloud services by industries are 55% in Financial, 60% Healthcare, 71% Retail, 86% High Tech, 78% Telecommunications, 67% Education, 59% Manufacturing and spending in cloud computing is estimated to rise from $33B in 2015 to $223B in 2025 (Office of Cyber and Infrastructure Analysis 2017). Likewise in Australia the Victorian Information and Communication Technology Advisory Committee (2014) ICT strategy has put cloud-based ICT services to be evaluated first for new and renewed systems as one of the principles which guide the ICT decision making in government and the scope of this strategy include many departments that are part of the critical infrastructure like Education and Early Childhood Development, Health, Justice, Transport, Planning and Local Infrastructure, Treasury and Finance. In addition, the scope includes many

agencies like, Ambulance Victoria, Country Fire Authority, Public Transport Victoria, Victoria Police and Victoria State Emergency Services. In the new Victorian ICT strategy Although Cloud was moved to be second to what they called “Share” which is referring to existing shared services or service that can be transitioned to shared service model, if there is no such existing service, ICT investment will be considered and Cloud is the first option (Department of Premier and Cabinet 2016).

Cybersecurity and Risks in Cyber Context

Many governments such as the Victoria State developed and maintaining cybersecurity strategies to deal with related risks. As cybersecurity strive to protect the confidentiality, availability and integrity of data and information from internal and external breaches (Victorian Government 2017), there are different possible risks in cloud services that are common with the traditional information technology model. The main difference between the two is that by default when a service is in the cloud, it is accessible by any device connected to the internet. Examples of such threats are: brute force, data leakage, denial of service, domo escalation, hyperjacking, phishing, RAM scraping (“A type of malware designed for monitoring and extracting data from a system during data processing while it is unencrypted”) and virtual machine escape (“The act of escaping a virtual machine (a virtual system or application that is running inside a physical system) and interacting directly with the virtual machine’s hosting environment.”) (Office of Cyber and Infrastructure Analysis 2017). Also there is the information risk which is the use of information power to the advantage of the attacker. Moreover, the interlinked operations of the different sectors in the critical infrastructure has the potential of cascading any damage to the other sectors risk.

Technical Risk:

By definition cloud computing involve accessing shared computing resources through network with minimal human interaction. That is implying that users can access these resource with any device connected to the internet. Although the users have to pass security measures to access these resources, there is the risk of unauthorized access either via human factor (social engineering) or systems undiscovered or untreated vulnerabilities. In addition, as the cloud computing depends on network connection to access the resources, in conventional warfare an attack to this network can damage these services. Thus there is the soft-access risk or armed-with-weapons attacks possible to the infrastructure.

Information Risk

With e-government using information and communication technologies to provide citizens and businesses more convenient access to information and services to improve service quality and greater opportunities to participate in democratic institutions and processes, there is a great risk of external interferences to influence these processes and shape their results according to the intruder intentions. In addition, as critical information become available in the cloud, they can be used in conventional and cyber warfare for the advantage of the attacker.

Cascading Effects Risk
By the nature of the critical infrastructure of being interconnected systems, any success of the attacker to destruct or disrupt the cloud services, there are possibilities of negative cascading consequences affecting the different sectors of the critical infrastructure.

Government Cloud Computing Adoption Approaches

Despite all the risks involved in cloud computing adoption in government and shifting the critical infrastructure to the cloud, the financial crisis and the continues strive to efficiency and the benefits overlook the risks involved and many government resolved to different models to adapt cloud computing in its operations.

In Australia the Federal government is using Microsoft cloud services named Azure from data centres located in Canberra, Sydney and Melbourne (Microsoft) and enforce a policy that data has to be kept within the Australian borders. France was in favour of developing nation-wide government cloud and it is called “Andromeda” which was set up by two companies; Orange and Thales (Zwattendorfer et al. 2013). In USA government entities are procuring commercial cloud services as per their needs and selection processes. UK has set up their own governmental cloud and has their CloudStore which offers infrastructure, software, platform and specialized services(Zwattendorfer et al. 2013). In Oman, the government set up their own infrastructure and called it G-Cloud and it offers IaaS, PaaS, SaaS and business processes as a service (Information Technlogy Authority 2018).

Conclusion and future research

In conclusion, cloud computing is service based information technology model that is replacing in-house asset based information technology model and can be deployed in four different ways. That is, public, private, hybrid and community cloud. Cloud services are three types of service which are infrastructure as service, platform as a service and software as a service. There are many advantages to cloud computing which include reducing cost and improving efficiency.

Adopting cloud computing in government involves shifting critical infrastructures to the cloud. The nature of cloud computing and the critical infrastructure make them vulnerable to cyber warfare activities and attacks but the risks involved in this shift are not stopping governments from adopting cloud computing in its operation and they are approaching the deployment differently. Some governments resolved to buy the cloud services from private companies like the United States while some others are deploying it through building cloud specific to government and in some cases called government-cloud like Australia, UK, France and Oman. The g-cloud cloud resources dedicated solely to the government and are deployed by government or private companies.

It is clear that cloud computing is been adopted by many governments and by highlighting the characteristics of the cloud computing and the critical infrastructure and possible risks involved in addition to common technology associated risks, this paper point out how government-cloud is a potential cyber warfare battlefield. Also it highlighted that it is the question of how to mitigate the risks rather than to adopt or not. Thus, researchers are encouraged to research not only common risks (in peace era) but as well on different risk mitigations strategies to reduce the loss in case of such cyber warfare that target cloud computing services used by governments.

References

Alford, T & Morton, G 2009, 'The Economics of cloud computing', *Booz Allen Hamilton*.

Ali, A, Warren, D & Mathiassen, L 2017, 'Cloud-based business services innovation: A risk management model', *International Journal of Information Management*, vol. 37, pp. 639-49.

Ali, K 2015, 'Factors determining the implementation of cloud computing in Michigan local government agencies', D.B.A. thesis, 3707561 thesis, Baker College (Michigan).

Brown, C, Seville, E & Vargo, J 2017, 'Measuring the organizational resilience of critical infrastructure providers: A New Zealand case study', *International Journal of Critical Infrastructure Protection*, vol. 18, pp. 37-49.

Bwalya, KJa & Mutula, SM 2014, *E-government : implementation, adoption and synthesis in developing countries*, Global studies in libraries and information: Volume 1, Berlin ; Boston : De Gruyter/Saur, [2014].

Cazorla, L, Alcaraz, C & Lopez, J 2018, 'Cyber Stealth Attacks in Critical Information Infrastructures', *IEEE Systems Journal, Systems Journal, IEEE*, no. 2, p. 1778.

Department of Premier and Cabinet 2016, *Information Technology Strategy Victorian Government 2016–2020*, by Department of Premier and Cabinet.

Digital Transformation Agency 2017, *Secure Cloud Strategy*, Australian Government Digital Transformation Agency, Australia.

Ghafir, I, Jaf, S, Prenosil, V, Saleem, J, Hammoudeh, M, Faour, H, Jabbar, S & Baker, T 2018, 'Security threats to critical infrastructure: the human factor', *Journal of Supercomputing*, pp. 1-17.

Grönlund, Å & Horan, TA 2005, 'Introducing e-gov: history, definitions, and issues', *Communications of the association for information systems*, vol. 15, no. 1, p. 39.

Guo, Y 2010, 'E-Government: Definition, Goals, Benefits and Risks', in *2010 International Conference on Management and Service Science*, pp. 1-4.

Harašta, J 2018, 'Legally critical: Defining critical infrastructure in an interconnected world', *International Journal of Critical Infrastructure Protection*, vol. 21, pp. 47-56.

Hsu, P-F, Ray, S & Li-Hsieh, Y-Y 2014, 'Examining cloud computing adoption intention, pricing mechanism, and deployment model', *International Journal of Information Management*, vol. 34, no. 4, pp. 474-88.

Information Technlogy Authority 2018, *G-Cloud*, Information Technlogy Authority, retrieved 05 February 2018, <https://www.ita.gov.om/G-Cloud/G-Cloud.aspx>.

Kundra, V 2011, *Federal cloud computing strategy*, The White House, Washington.

Madhavaiah, C, Bashir, I & Shafi, SI 2012, 'Defining Cloud Computing in Business Perspective: A Review of Research', *Vision (09722629)*, vol. 16, no. 3, pp. 163-73.

Marston, S, Li, Z, Bandyopadhyay, S, Zhang, J & Ghalsasi, A 2011, 'Cloud computing - The business perspective', in *The 44th Hawaii International Conference on System Sciences* vol. 51, pp. 176-89.

Mell, P & Grance, T 2011, 'The NIST definition of cloud computing'.

Microsoft, *Microsoft First to Offer Azure Cloud Regions for Australian Government - Microsoft News Centre Australia*2018, <https://news.microsoft.com/en-au/features/microsoft-first-offer-azure-cloud-regions-australian-government/>.

Nirenjena, S, Divya, A, Aswini, R, Jayalakshmi, S & Saradhambal, G 2017, 'A cloud computing revolution in business perspective', *Advances in Natural and Applied Sciences*, vol. 11, p. 558+.

Office of Cyber and Infrastructure Analysis 2017, *Risks to Critical Infrastructure that Use Cloud Services*, US Department of Homeland Security.

Pye, G & Warren, M 2007, 'Modelling critical infrastructure systems', *Journal of information warfare*, vol. 6, no. 1, pp. 41-53.

Pye, G & Warren, M 2009, *An emergent security risk : critical infrastructures and information warfare*, Mindsystems Pty. Ltd.

Schnoll, HJ 2015, *E-Government : Information, Technology, and Transformation*, Hoboken : Taylor and Francis, 2015.

Senarathna, I, Yeoh, W, Warren, M & Salzman, S 2016, 'Security and privacy concerns for Australian SMEs cloud adoption: Empirical study of metropolitan vs regional SMEs', *Australasian Journal of Information Systems*, vol. 20, p. 20p.

Senarathna, R 2016, 'Cloud computing adoption by SMEs in Australia', Thesis thesis, Deakin University, Faculty of Business and Law, Department of Information Systems and Business Analytics.

Tsavdaroglou, M, Al-Jibouri, SHS, Bles, T & Halman, JIM 2018, 'Proposed methodology for risk analysis of interdependent critical infrastructures to extreme weather events', *International Journal of Critical Infrastructure Protection*, vol. 21, pp. 57-71.

Department of Premier and Cabinet 2017, *Cyber Security Strategy*, by Victorian Government, The Victorian Government.

Victorian Government Solicitor's Office 2011, *Cloud Computing in a Government Context*, Victorian Government Solicitor's Office, retrieved 27 August 2018, <https://www.vgso.vic.gov.au/news-events/news/cloud-computing-government-context-speakers-notes>.

2014, *Victorian Government ICT Strategy 2014 To 2015*, by Victorian Information and Communication Technology Advisory Committee.

Zhang, W & Chen, Q 2010, 'From E-government to C-government via cloud computing', *Proceedings of the International Conference on E-Business and E-Government, ICEE 2010*.

Zwattendorfer, B, Stranacher, K, Tauber, A & Reichstädter, P 2013, 'Cloud Computing in E-Government across Europe', in Berlin, Heidelberg, pp. 181-95.

Is Your Office Environment Conducive to Good Information-Security Behaviour?

Malcolm Pattinson[1], Beau Ciccarello[1], Marcus Butavicius[2], Kathryn Parsons[2], Agata McCormac[2], Dragana Calic[2].

[1] Adelaide Business School, University of Adelaide, South Australia
{malcolm.pattinson, beau.ciccarello}@adelaide.edu.au

[2] Defence Science and Technology Group, Edinburgh, South Australia
{marcus.butavicius, kathryn.parsons, agata.mccormac, dragana.calic}@dst.defence.gov.au

Abstract

This paper describes a research proposal to investigate the extent to which the information-security (InfoSec) behaviour of employees is affected by workplace distractions. This concept paper introduces a research design that may highlight the types of office environments, in terms of layout or configuration, that are less likely to result in the types of distraction that are associated with naïve, accidental and unintentional InfoSec behaviour. The proposed research consists of two stages. Stage 1 will use the Repertory Grid Technique (RGT) to interview a selection of approximately 20 employees from different organisations to identify the types of distraction that are perceived to cause them to unintentionally behave in a risk-inclined manner. This Stage will also establish the type of office they predominantly work in. Stage 2 will deploy an online survey distributed via email to each of the Stage-1 participants to determine their InfoSec awareness (ISA) score. The results of Stages 1 and 2 will then be compared to investigate any associations between distractions, office environments and ISA.

Keywords: Information Security (InfoSec), InfoSec Awareness (ISA), Office Layouts, Environmental Distractions.

1. Introduction

The need for adequate security of information systems and the data that they store, process and transmit, has never been greater for organisations and individuals. Factors that contribute to this need include:

- the increased use of, and dependence on, the internet within public and private sectors
- the emergence and increased use of technologies such as wireless, mobile commerce and social networking
- the expectation by customers and business partners that the security of an organisation's information systems is adequate

- The exponential increase in the number of data breaches that have caused huge monetary and other losses to all types of businesses and industries
- The vast majority of InfoSec incidents are caused by human error.

This ever-increasing need has triggered the realisation of Board Directors, Information System Managers and Chief InfoSec Officers (CISOs) that the most effective means of mitigating the risk to the information systems within an organisation is to address the InfoSec behaviour of digital-device users in parallel with, and not instead of, implementing hardware and software controls. This human behavioural approach to managing InfoSec supports the Schneier (2004) claim that "...the biggest security vulnerability is still that link between keyboard and chair" (p. 1).

1.1. Aim of this paper

The main aim of this paper is to describe a research proposal that will answer the following research questions:

- How does the office environment of an employee (e.g. open-plan or enclosed offices) influence their InfoSec behaviour when they are using a digital device at work?
- What types of distractions have the most impact on employee naïve, accidental and unintentional InfoSec behaviour?

The proposed research described in this paper examines the office environments of digital-device users and the respective distractions associated with these environments. In addition, this research assesses the InfoSec awareness (ISA) of individual employees as an indicator of how well they behave when using a digital device at work. It is anticipated that the findings of this research will have practical implications for management relating to the configuration of offices that promote risk-averse InfoSec behaviour (compared to risk-inclined InfoSec behaviour). This may be more cost-effective than implementing a range of controls that involve hardware, software and policies and procedures.

2. Theoretical Background

2.1. InfoSec Behaviours by Digital-Device Users

For the purposes of this research and this paper, the term 'InfoSec behaviours by digital-device users' refers to the full spectrum of InfoSec behaviours displayed by employees who use digital devices as part of their job. Table 1 below, which was developed by the authors, shows these behaviours range from deliberate risk-averse behaviours, to accidental neutral behaviours to deliberate risk-inclined behaviours.

Risk-averse behaviour (deliberate)	Neutral behaviour (accidental)	Risk-inclined behaviour (deliberate)
Always log-off when computer unattended	Leaving a digital device unattended	Installing/using unauthorised software
Disallow email attachments from unknown sources	Opening unsolicited email attachments	Creating and sending SPAM email
Installing anti-virus software and updating regularly	Not installing anti-virus software	Writing and disseminating malicious code
Changing passwords regularly	Sharing passwords	Hacking into other people's accounts
Vigilance in recognising and approaching unauthorised personnel	Not being vigilant re unauthorised personnel	Giving unauthorised personnel access to authorised precincts
Backing up work regularly	Not backing up work often enough	Theft or destruction of hardware or software
Always reporting security incidents	Not reporting security incidents	Conducting fraudulent activities
Installing firewall software	Accessing dubious web sites	Executing games on company digital devices

Table 1**: A Categorisation of InfoSec Behaviours by Digital-device Users (Adapted from (Pattinson & Anderson 2007)**

InfoSec behaviours by digital-device users have also been categorised by Stanton et al. (2005) who refer to the deliberate risk-averse behaviours (above) as 'Aware Assurance' or 'Basic Hygiene'; the accidental neutral behaviours (above) as 'Dangerous Tinkering' or 'Naïve Mistakes'; and the deliberate risk-inclined behaviours (above) as 'Intentional Destruction' or 'Detrimental Misuse'. The research described in this paper focuses on the accidental neutral behaviours shown in the middle column of Table 1 above.

2.2. Distractions

When people are distracted, they lose concentration on the task at hand (ISO_3382-3 2012) and are more likely to make mistakes. This premise was supported by Waroquier et al. (2009) who provided evidence that conscious thought is beneficial in decision-making. In terms of InfoSec, mistakes take the form of naïve, accidental, unintentional, risk-taking behaviour. The risks associated with this type of behaviour include unauthorised access to private and sensitive information, the inability to

access computer systems and making decisions based on incorrect information. The consequences of these risks include loss of life, financial losses, reputational damage and going out of business.

2.2.1. Causes of distraction

Noise is a predominant cause of distraction in offices, particularly intrusive noise such as speech from general conversation (Clements-Croome 2006), telephones ringing and telephone conversations. In fact, Hongisto (2005) maintains that distraction increases as speech intelligibility increases. In summary, distraction from intrusive speech not only depends on the distance between the speaker and the distracted employee (ISO_3382-3 2012), but also depends on office-type variables such as room-geometry, furniture (i.e. office layout or configuration), occupancy and continuous steady noise (e.g. from mechanical services).

2.2.2. Relationship between distractions and office layouts

The employees of an organisation may be required to sustain concentration when performing certain tasks while using digital devices. For organisations that manage information of a highly sensitive nature (e.g. defence organisations and banking institutions), the work environments are typically offices with varying degrees of privacy. Kaarlela-Tuomaala et al. (2009) looked at office-workers' perceptions of their work environments and how they felt about being relocated between private offices (one person per room) and open-plan offices occupied by numerous employees. Office-type determines the distraction that the employees of an organisation experience (Lee & Brand 2005).

2.3. InfoSec Awareness (ISA)

In this study, an individual's InfoSec behaviour will be reflected by his or her ISA score. This current research uses Parsons et al. (2014) definition of ISA. This definition is made up of the following three components:

- What a person 'knows' about behaving in a safe manner (Knowledge);
- How a person 'feels' about behaving in a safe manner (Attitude) and
- What a person actually 'does' when using a digital device (Self-reported behaviour).

Parsons et al. (2017) demonstrated that an individual's ISA, as measured by 63 knowledge, attitude and self-reported behaviour (KAB) survey items, is a valid and strong indicator of how securely the respondent behaves when using a digital device. Furthermore, Parsons et al. (2017) showed that a respondent's InfoSec behaviour is also highly predictive by using only knowledge scores, that is, without attitude and self-reported behaviour scores. Therefore, this proposed research intends to use only the 21 knowledge items via an online survey, to evaluate an employee's ISA score. These questions are shown in Table 2 in Section 4.3.

2.3.1. Relationship between ISA, distractions and office layouts

It could be argued that people who work in open-plan offices (compared to enclosed offices) might be less likely to sustain concentration due to aural and visual distractions and therefore more likely to engage in poor InfoSec behaviour. For example, someone who is distracted by a noisy telephone conversation nearby might be less likely to notice a suspicious email, and could be more susceptible to a phishing attack. This suggests that open-plan office environments might be less conducive to good ISA. On the other hand, these same individuals who work in open-plan offices might be less likely to engage in poor InfoSec behaviour. For example, they probably would not leave sensitive documents lying around, would not write down their password, and would not access dubious websites. This suggests that open-plan environments might be more conducive to good ISA.

3. Research Design

3.1. Stage 1

This stage is a qualitative, ground-up approach with the aim of collecting qualitative data to provide a basis for developing the quantitative instrument to be used in Stage 2. More specifically, the main objective of Stage 1 is to identify the different types of distractions that are perceived by employees that cause them to unintentionally behave in an unsafe manner. For this stage, the Repertory Grid Technique (RGT) will be used to interview approximately 20 employees from different organisations. This is a cognitive technique that was developed by, and is grounded in, George Kelly's Personal Construct Theory (Kelly 1955). It is a method of interviewing in which interviewees divulge their attitudes, thoughts and views about a situation, object or event. In this study, the domain of investigation is distractions in an office environment. i-polar constructs that relate to grid elements will be developed using the techniques of triading, laddering and pyramiding. These grid elements, will be devised by the researchers, and will comprise approximately 10 office distractions that can be addressed by changing office layout, such as:

- Intrusive speech by others nearby
- Sudden non-speech sounds e.g. doors that squeak and bang when closed.
- People stopping by their workplace
- Visual distraction outside (through the window)
- Visual distraction inside.

Also in this stage, interviewees would score each grid element (i.e. each supplied distraction) on a scale between 1 and 5 for each developed bi-polar construct.

In addition, this Stage 1 will establish the extent of distractions in each participant's office by getting them to respond to the following Lee and Brand (2005) semantic differential items:

1. I find it difficult to concentrate on my work.
2. I experience auditory distractions in my work area.

3. I have adequate privacy in my primary, individual work area.
4. I experience visual distractions in my work area
5. My work environment is too noisy.

These questions will be answered on a 7-point scale, ranging from ‘Yes, all the time’ to ‘No never’, except for the third question which ranges from ‘Yes, most definitely’ to ‘No, definitely not’.

Finally, participants will be asked to describe their office environment by selecting from six office-type options as defined by Kim and De Dear (2013) and being either Enclosed or Open-plan as follows:

- Enclosed private
- Enclosed shared (with 2-3 people)
- Open-plan cubicles with high partitions
- Open-plan cubicles with low partitions
- Open-plan with no partitions.

These office-type options are representative of the typical range of office configurations based on the literature, such as Kim and De Dear (2013) and Lee and Brand (2005).

3.2. Stage 2

This stage will comprise the design, development and distribution of a web-based survey questionnaire distributed via email to the 20 or so participants who were interviewed in Stage 1. The purpose of this survey would be to collect some demographic data about each participant and to ask the 21 knowledge questions, shown in Table 2 below, to enable the ISA score to be calculated.

No.	Knowledge Item	Focus Area
1	A mixture of letters, numbers and symbols is necessary for work passwords.	PM
2	It's acceptable to use my social media passwords on my work accounts.	PM
3	I am allowed to share my work passwords with colleagues.	PM
4	I don't need to be cautious when clicking on links in emails from people I know.	EU
5	I should be careful when clicking on links in emails from unknown senders.	EU
6	I don't have to be careful when opening email attachments from unknown senders.	EU
7	I am allowed to enter information on any website if it helps me do my job.	IU
8	I am allowed to download any files onto my work computer if they help me to do my job.	IU
9	While I am at work, I shouldn't access certain websites.	IU
10	When working in a public place, I have to keep my laptop with me at all times.	MC
11	When working on a sensitive document, I must ensure that strangers can't see my laptop screen.	MC
12	I am allowed to send sensitive work files via a public Wi-Fi network.	MC
13	I should periodically review the privacy settings on my social media accounts.	SMU
14	I can't be fired for something I post on social media.	SMU
15	I can post what I want about work on social media.	SMU
16	Sensitive print-outs can be disposed of in the same way as non-sensitive ones.	IH
17	If I find a USB stick in a public place, I shouldn't plug it into my work computer.	IH
18	I am allowed to leave print-outs containing sensitive information on my desk overnight.	IH
19	If I see someone acting suspiciously in my workplace, I should report it.	IR
20	I must not ignore poor security behaviour by my colleagues.	IR
21	It's optional to report security incidents.	IR

Table 2**: Knowledge questions to predict InfoSec Awareness**

Each question is answered on a 5-point scale ranging from 'Strongly Disagree to 'Strongly Agree'. Three knowledge statements were presented

for each of the seven InfoSec focus areas, namely, password management (PM), email use (EU), internet use (IU), social media use (SMU), mobile computing (MC), information handling (IH) and incident reporting (IR). Approximately half of the statements are negatively worded, and statements across the seven InfoSec focus areas are randomly ordered. Negatively-worded statements will be taken into consideration prior to data analysis. Therefore, a respondent's ISA score is the sum of the 21 scores between 1 and 5. The higher this 'knowledge score' is, the better behaved the individual is likely to be. For groups of individuals, the mean of the scores is used.

4. Implications and Summary

The focus of this research proposal is the InfoSec behaviour of employees who use digital devices to do their work. This implies that the research approach described herein is predominantly focussed on the behaviour of human beings as it relates to the security of organisational information systems. This proposed research attempts to identify the environmental distractions that influence this type of InfoSec behaviour.

It is anticipated that the research put forward in this paper will provide management with answers to the following questions:

- How does the office environment of an employee (e.g. open-plan or enclosed office) influence their InfoSec behaviour when they are using a digital device at work?
- What types of distraction have the most impact on employee naïve, accidental and unintentional InfoSec behaviour?

The development of offices ranging from "conventional private (or cellular) spatial configuration to modern open-plan" (Kim & De Dear 2013, p. 18) should be proportional to the information at stake for the organisation in question. For example, for the areas in defence organisations, legal firms and banking institutions that are associated with highly-sensitive information, is it prudent to deploy private offices to prevent unauthorised access to information? Our proposed research seeks to examine whether, in considering the choice of office accommodation, we should also be aware of the effect of distractions in these environments and how they may negatively impact the information security behaviours of workers in them.

5. References

Clements-Croome, D 2006, *Creating the productive workplace*, Taylor & Francis.
Hongisto, V 2005, 'A model predicting the effect of speech of varying intelligibility on work performance', *Indoor air*, vol. 15, pp. 458-468.
ISO_3382-3 2012, *Acoustics – Measurement of room acoustic parameters – Part 3: Open Plan Offices*, International Standards Organization.
Kaarlela-Tuomaala, A, Helenius, R, Keskinen, E & Hongisto, V 2009, 'Effects of acoustic environment on work in private office rooms and open-

plan offices–longitudinal study during relocation', *Ergonomics*, vol. 52, no. 11, pp. 1423-1444.
Kelly, G 1955, *The Psychology of Personal Constructs* vol. 1 & 2, Norton, New York.
Kim, J & De Dear, R 2013, 'Workspace satisfaction: The privacy-communication trade-off in open-plan offices', *Journal of Environmental Psychology*, vol. 36, pp. 18-26.
Lee, SY & Brand, JL 2005, 'Effects of control over office workspace on perceptions of the work environment and work outcomes', *Journal of Environmental Psychology*, vol. 25, no. 3, pp. 323-333.
Parsons, K, Calic, D, Pattinson, M, Butavicius, M, McCormac, A & Zwaans, T 2017, 'The human aspects of information security questionnaire (HAIS-Q): two further validation studies', *Computers & Security*, vol. 66, pp. 40-51.
Parsons, K, McCormac, A, Butavicius, M, Pattinson, M & Jerram, C 2014, 'Determining Employee Awareness Using the Human Aspects of Information Security Questionnaire (HAIS-Q)', *Computers & Security*, vol. 42, pp. 165-176.
Pattinson, M & Anderson, G 2007, 'How well are information risks being communicated to your computer end-users?', *Information Management & Computer Security*, vol. 15, no. 5, pp. 362-371.
Schneier, B 2004, *The People Paradigm*, viewed June 23 2011, <http://www.csoonline.com/article/219787/bruce-schneier-the-people-paradigm>.
Stanton, J, Stam, K, Mastrangelo, P & Jolton, J 2005, 'Analysis of end user security behaviors', *Computers & Security*, vol. 24, no. 2, pp. 124-133.
Waroquier, L, Marchiori, D, Klein, O & Cleeremans, A 2009, 'Methodological pitfalls of the unconscious thought paradigm', *Judgment and Decision Making*, vol. 4, no. 7, pp. 601-610.

Social Media and Cyber Security Awareness Program: Is Security Communication A Social Activity?

Hiep Pham, Mathews Nkhoma, Irfan Ulhaq, *RMIT University Vietnam.*

With the dramatic development of communication technologies, virtual communities on social media platforms have emerged as a new way of group knowledge sharing, which allow people to share information and experience without meeting face-to-face (Chang et al., 2015). Rather than searching for information passively, social media offers more cooperative and open communication, where a large number of people are free to share any of their thoughts, experiences, opinions, feedbacks and perspectives (Kaplan & Haenlein, 2010). The increased mobility of social media due to of the popularity of smartphone possession naturally leads to the interaction between people and social media tools become a daily activity (Kwahk & Park, 2016). Furthermore, social media platforms offer a better way to acquire new knowledge from peers, networks and through live engagements (Wasko & Faraj, 2000). Social media tools are considered as a knowledge management system, as they allows a flexible mode of dialogue, blogging, sharing of information in several format as well as live streaming with peers (Kwahk & Park, 2016; Oostervink et al., 2016). Social media enables dissemination of knowledge and information in multiple forms such as videos, photos and audios, which increases the effectiveness of sharing knowledge by offering clearer images and visions for users (Kwahk & Park, 2016). Although, the role of social media as knowledge sharing platform is well recognized, research on use of social media as security awareness and communication platform is scarce (Aloul, 2012; Gupta & Brooks, 2013; Hajli & Lin, 2016).

Our paper presents initial findings regarding the use of social media to influence security awareness and behaviour of employees. 25 participants from five financial organisations in Vietnam shared their experience and opinions towards the current use of social media at work for stock information sharing which can also be used to disseminate urgent security notices and share personal security incidents. Implications of social media as a timely and interactive cyber information sharing channel to both organisations and individuals are also provided in the paper.

References

Aloul, F. A. (2012). The Need for Effective Information Security Awareness. *Journal of Advances in Information Technology,* Vol 3 No. 3, pp. 176-183.

Chang, C. M., Hsu, M. H., & Lee, Y. J. (2015). Factors Influencing Knowledge-Sharing Behavior in Virtual Communities: A Longitudinal Investigation. *Information Systems Management,* Vol 32 No. 4, pp. 331-340. doi:10.1080/10580530.2015.1080002

Gupta, R., & Brooks, H. (2013). *Using Social Media for Global Security*: John Wiley & Sons.

Hajli, N., & Lin, X. (2016). Exploring the Security of Information Sharing on Social Networking Sites: The Role of Perceived Control of

Information. *Journal of Business Ethics,* Vol 133 No. 1, pp. 111-123.
Kaplan, A. M., & Haenlein, M. (2010). Users of the World, Unite! The Challenges and Opportunities of Social Media. *Business Horizons,* Vol 53 No. 1, pp. 59-68.
Kwahk, K.-Y., & Park, D.-H. (2016). The Effects of Network Sharing on Knowledge-Sharing Activities and Job Performance in Enterprise Social Media Environments. *Computers in Human Behavior,* Vol 55 No., pp. 826-839.
Oostervink, N., Agterberg, M., & Huysman, M. (2016). Knowledge Sharing on Enterprise Social Media: Practices to Cope with Institutional Complexity. *Journal of Computer-Mediated Communication,* Vol 21 No. 2, pp. 156-176.
Wasko, M. M., & Faraj, S. (2000). “It Is What One Does”: Why People Participate and Help Others in Electronic Communities of Practice. *The Journal of Strategic Information Systems,* Vol 9 No. 2, pp. 155-173.

A Study of Cybersecurity Awareness in Sri Lanka

R. T. S. Nagahawatta[1], Matthew Warren[2], William Yeoh[3]

[1]University of Kelaniya, Sri Lanka.
[2]Deakin University Centre for Cyber Security Research and Innovation, Deakin University, Geelong, Victoria, Australia.
[3]Department of Information Systems and Business Analytics, Deakin University, Australia.

[1]ruwann@kln.ac.lk; [2]matthew.warren@deakin.edu.au; [3]william.yeoh@deakin.edu.au

Abstract
Cybersecurity is focused on helping the community to make knowledgeable decisions on its adaptation and mitigation. This survey evaluated the level of cybersecurity awareness and discernment amongst university students in Sri Lanka. The study was based on primary data collected through a questionnaire on awareness and perception of cybersecurity from respondents in different degree programs among universities in Sri Lanka. The results indicated that experience and the level of cybersecurity awareness among university students in Sri Lanka are not significantly low, but there are some knowledge gaps with new threats. Further, the results showed that university students in Sri Lanka were able to identify cybercrime as a threat. These findings necessitate building awareness and developing capacity to improve student's knowledge on the cybersecurity subject especially if universities are to be used as a key focal point in cybersecurity awareness campaigns in Sri Lanka.

Keywords: Cybersecurity, Awareness, Perception, Knowledge.

1. Introduction

The Internet has become a part of the life of many people around the world (Kritzinger, 2010). "There is no argument whatsoever that the proliferation of devices and information are empowering. Technology is today far more democratically available than it was yesterday and less than it will be tomorrow" (Geer, 2015). The Internet evolvement in Sri Lanka is remarkable and most of the internet related latest technologies were introduced to Sri Lanka sometimes even before the other countries in the region (Abeysekara et al., 2012). Both the government and the corporate sectors of Sri Lanka have also incorporated the cyberspace into their operations. Thus, operations of the government and private sector institutions, heavily rely on computers and the internet. However, there are many threats and risks incorporate with the internet (Riem, 2001). Furthermore, the internet has exposed to criminal activities due to private information on it (Joode, 2011). Hence, there is a risk of misusing and compromising personal data on the internet. Stone (2013) illustrated cyber risk on students into three groups as shown Table 1.

Individuals' intention to harm the learner	Learners' exposure to harmful online interactions	Leaner places her / himself in a harmful situation
Cyberbullying: trolling, flaming excluding, masquerading, mobbing, denigrating, outing, harassing, cyber grooming, impersonation, blackmail, cyber snooping, identity theft, social engineering, online predators.	The inappropriate content/material, digital reputation ruin, social platforms, and chat rooms, viruses, malware, cookies.	Illegal file sharing, plagiarism, inappropriate posting online, free downloads, non-ethical postings of others' materials, sexting.

Table 1: **Cyber Risk Category (Stone, 2013)**

Lack of awareness and knowledge cause inability to protecting their personal data (Thomson, 2006). Also, the lack of awareness about cybersecurity among parents has a negative impact on the role of protecting their children from cybercrimes (Lange, 2011; Atkinson, 2009). Hence, students should equip with the necessary knowledge for the cyber threats and risks that they have to face on every day (Kritzinger, 2017). According to Lange and Solms (2012), most adults do not have enough knowledge about online threats in order to protect their children from unsecured internet access. Also, hackers aim to lose points that made by lack of knowledgeable online users (Kritzinger, 2012). Christensen (2003) argued that providing awareness about cybersecurity would facilitate secure online behavior. In addition, promoting the education could contribute to minimize the risk of users' insecure online behaviors (Kritzinger, 2013). Therefore, the aim of this study is to examine the knowledge and perception of cybersecurity among university students in Sri Lanka.

1.1 Background of Sri Lanka

Sri Lanka is an island located in the Indian ocean and has a population over 22 million. Out of the total population in Sri Lanka, 32% are using the Internet (CIA, DN). In Sri Lanka cybersecurity has become a key issue due to numerous reasons. For instance, the official website of the Sri Lankan President Maithripala Sirisena, (www.president.gov.lk) suffered by Sri Lankan youths in 2016 (Senarathna and Warren 2017). One of the main reasons is the lack of implementation of already enacted policies and regulations which is being put up to regulate the illegal activities in cyberspace and mitigate the misconducts. Nonetheless, in Sri Lanka, most of them are being ignored by the law enforcement authorities causing criminals to act without any consent.

The Sri Lankan government is aware of and concerned about cybercrime as a development issue (Palliyaguru, 2015). In this regard, the government has established several authorities. Sri Lanka Computer Emergency Readiness Team (SLCERT) which is fully affiliated to Information and

Communication Technology Agency (ICTA) of Sri Lanka. This is a national organization which acts as the main policy body for Information Technology of the nation. The Figure 1; (Palliyaguru, 2015), shows that a dramatic increase of cyber incidents afterward in 2010.

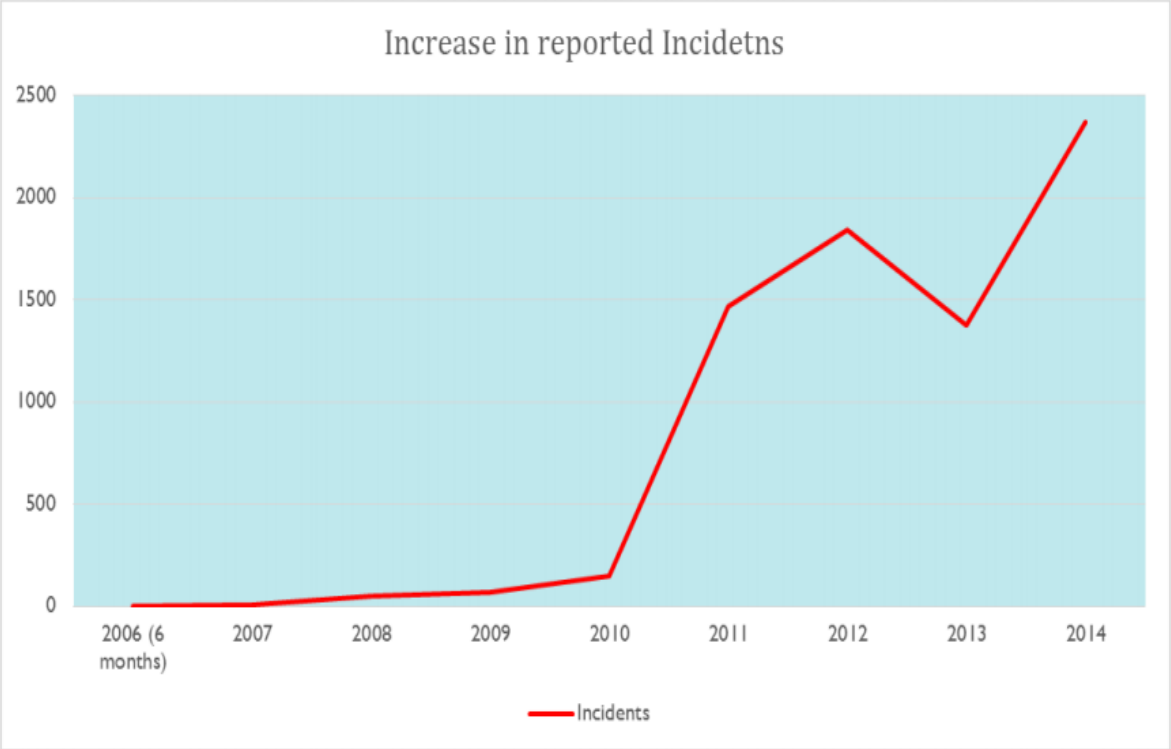

Figure 1: **Reported Cyber Incidents (Palliyaguru, 2015)**

Sri Lanka has established several laws relevant to cyber threat and copyright laws based on the English law. Both English law and Sri Lankan laws have common landscapes in relation to the digital media. Furthermore, there is no any difference between national law and the international law relating to cybersecurity in Sri Lanka. There are several legislations which passed by the Sri Lanka parliament namely Information and Communication Technology Act (No.27 of 2003), Computer Crimes Act (No. 24 of 2007), Payment Devices Frauds Act (No.30 of 2006) and Electronic Transactions Act (No. 19 of 2006). The technological framework for electronic signatures and authentication technologies and certificate authority established in September 2013. Sri Lanka has reported an increasing number of cyber incidents after introducing the internet. The Table 2 shows different types of incidents reported to Sri Lanka CERT during 2016 (APCert, 2016). Sri Lanka CERT has become the national cybersecurity in Sri Lanka, aim to protect the cyber threats and to coordinate defensive measurements and to protect the nation's information infrastructure (Sri Lanka Cert, ND).

Type of Incidents	Number of Incidents in the year 2016
Phishing	23
Abuse/Privacy violation	32
Ransomware	10

Spams	12
Financial frauds	16
Malicious software issues	11
Website compromise	10
Email threat	16
Intellectual property violation	7
DoS/DDoS	4
Social Media incidents	2200
Total	2341

Table 2: **Cyber Security Profile Incidents (2016) (AP, 2016)**

2. Data Collection and Analysis

Use of quantitative data is conducted in the positivist worldview of determinism, where causes determine effects or outcomes (Creswell, 2013). The study has focused on the relationship between cybersecurity and the level of awareness of higher education students attached to the national universities of Sri Lanka. Furthermore, the sample was selected from all 15 public universities in Sri Lanka. These higher education institutions are state-owned and they represent the highest number of students and high-quality graduates to the country. A part of the focused population can represent a sample that should be selected methodically for a study (Cooper, 2003). The sampling involves the selection of the units to be perceived on the basis of the researcher's individual decision about which ones will be the most suitable or representative (Babbie, 2005). The population is all the undergraduates those who are following all the kinds of degree programs in Sri Lanka. Random sampling technique was used to select a sample from all the undergraduates and from different degree programs. A student who is following a degree program can be identified as a unit of analysis of this study. This study randomly selected first, second, third and final year students of government universities to create the study sample.

Quantitative analysis is the gathering of information that can be articulated in mathematical appearance. This amalgamates information that is quantifiable and is capable of containing arithmetic consequences, monetary figures, or demographic facts which is a deductive compete of the association flanked by speculation and study.

In order to conduct the research, primary data was used to measure the level of awareness of students. Data was collected through a questionnaire which was developed in order to disclose key indicators that are related to demographics and behaviors to achieve research objectives.

3. Results

Respondent's answers were presented by using descriptive statistical methods. Frequency distributions tables revealed the number of respondents concerned. The frequency plots for each item was expressed in percentage and presented in tabular forms, diagrams such as pie charts and bar charts. These techniques were used to discover and summarize

the attributes of the sample to provide descriptive information. It was helpful to analyse the current situation of the selected dimensions. This analysis assists to achieve the objective of this research that is analyzing the existing situation of Sri Lankan university students. The questionnaire was distributed mainly among the 15 government universities in the simple random method. Therefore, the sample distribution of this study was illustrated in Table 3 to get a clear picture of the sample population. The population for the survey was 88,855 students at undergraduate level across 15 universities in Sri Lanka.

According to the capacity of respondents, 44% of respondents were male and, 56% of respondents were female. Further, the percentage values of gender categories were illustrated in the table. Respondents' level of the students in the sample profile has shown the Table 3. The majority (34%) of the respondents are in the final year, 3rd year and 2nd year, which is respectively 27% and 22% of students, while students who are in the 1st year were only 17%. According to the selected sample, 28 of respondents represent the Management and Accounting degree programs, 36 Science and Computing, 19 Social Sciences, and 23 Engineering degree programs. Further percentage values of study fields were illustrated in the above. It was clear that the majority of the respondents (36%) represent Science and Computing degree programs. Similarly, all other respondents fall into the 1st, 2nd, 3rd and 4th-year students those who are able to take partial experience decisions. The survey respondents' majority (40%) was from the Western province and 12% from Sabaragamuwa province, 8% in the Northern province,10% Central province and 7% from North Western province.

Demographic	Dimensions	Frequency	Percentage
Gender	Male	68	56
	Female	53	44
Study Field	Management and	28	28
	Science and Computing	36	36
	Social Sciences	19	19
	Engineering	23	23
	Other	15	15
Academic year	First year	21	17
	Second year	26	21
	Third year	33	27
	Final year	41	34
Province	Western	49	40
	Southern	11	9
	Uwa	7	6
	Northern	10	8
	Sabaragamuwa	14	12
	North Western	9	7
	Central	12	10
	North Central	5	4
	Eastern	4	3

Table 3: **Demographic Information**

4. Discussion

4.1 Online Incidents

As indicated by figure 2; the participants were asked to identify a number of online incidents they experienced. The statistics showed that majority (96) are spam emails and secondly 87 computer virus incidents. 12 cyberbullying victimizations and only 3 said that sexual solicitation. However, both male and female students are thought that they have experienced 42 unknown online incidents.

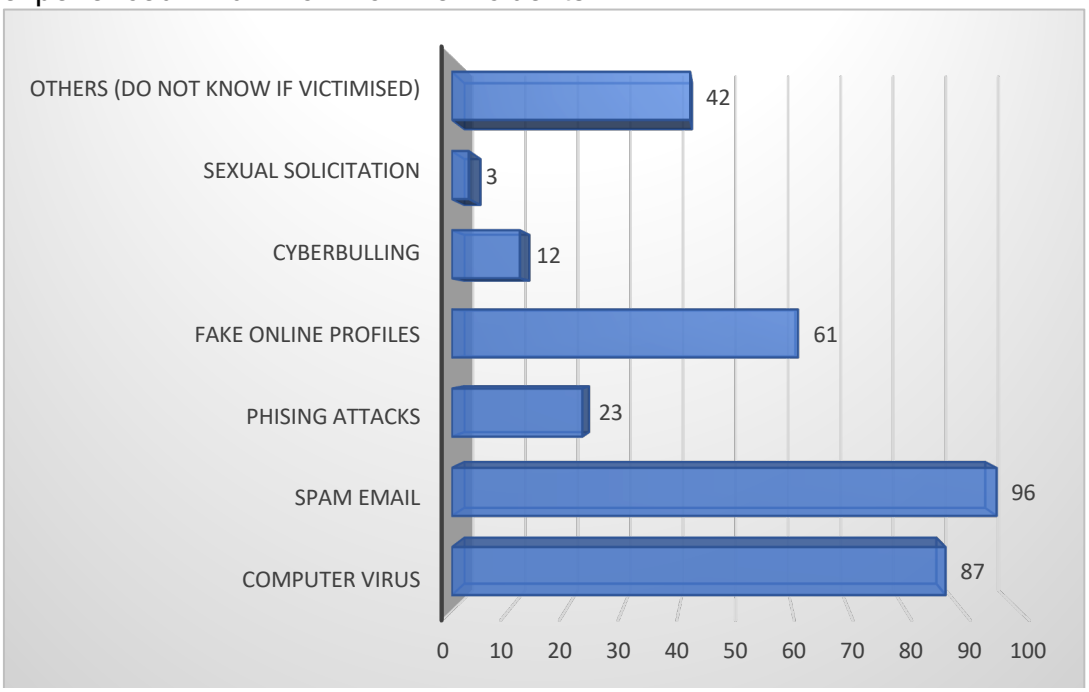

Figure 2: **Online Incidents**

4.2 Gender wise Cybersecurity Awareness

When considering survey respondent Figure 3 appears to be more aware of certain kinds of cybersecurity threats. For example, 51 male and 32 female knew how to update antivirus, while all most more than half of the male and female students are aware with enabling firewalls, social media privacy setting, identify spam emails and opening trust email attachments., However, both male and female students represent low-level belief that one's personal information is of no value to hackers with combination of 21 male and 9 female. Further, only a few said that adding unknowns to social media and given the password to others. This data suggests that overall awareness of students are higher and male students are in a position to understand the importance of threats rather than female.

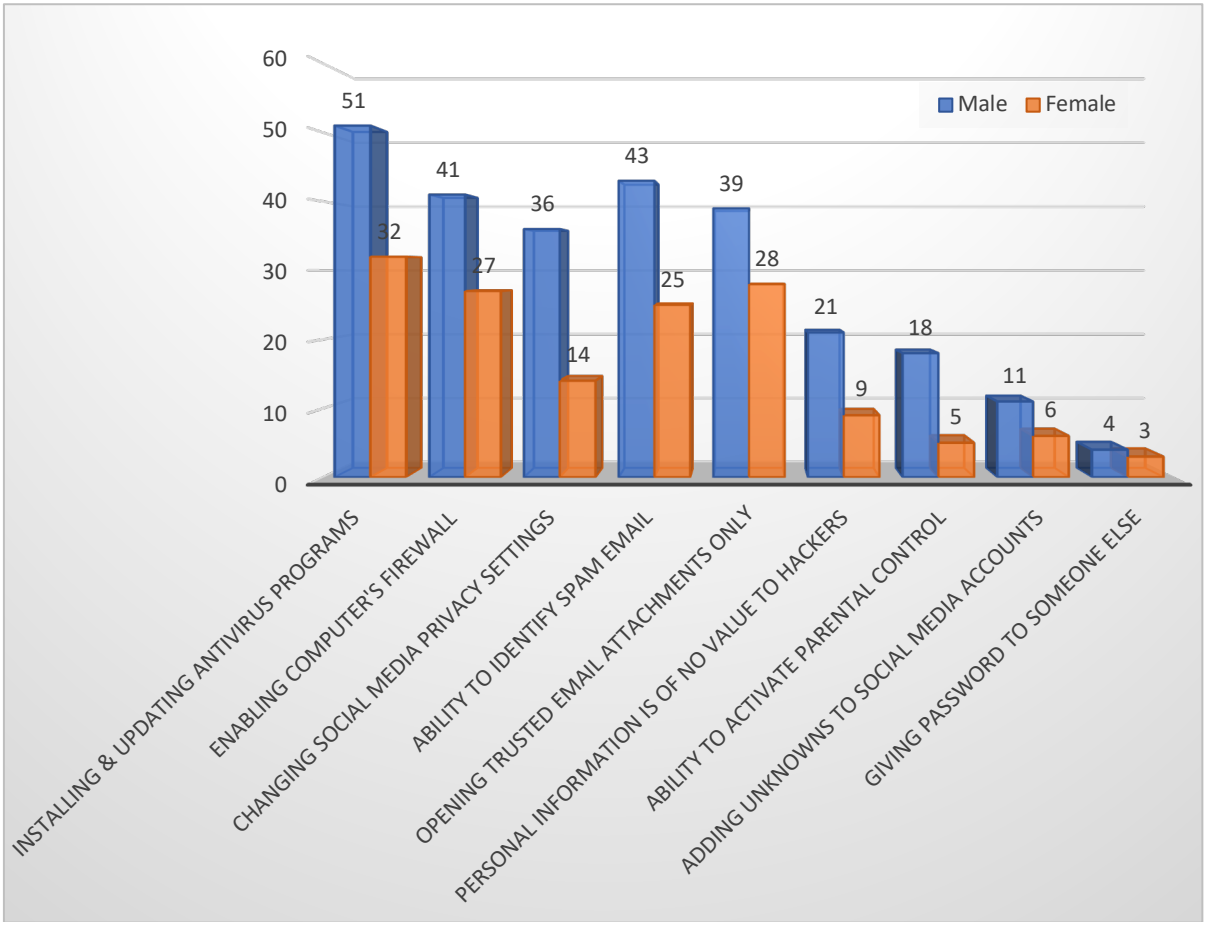

Figure 3: **Gender wise Cybersecurity Awareness**

4.3 Level of Cybersecurity Awareness

According to Figure 4, respondents' statistics indicate that their level of cybersecurity awareness. The majority of the respondents (39%) has moderate level cybersecurity awareness, secondly, 30% denotes a high level of awareness and further 9% respondents are of the view that it is high level. Only 15 respondents are of the opinion of that Organizational Influences is not at a satisfactory level while 16% indicates a high awareness and only 6% has very low-level knowledge. As a whole, 78% respondents of the sample said it is a moderate or high-level position. Generally, the respondents were well aware of the most common types of security enhancement tools and techniques. When considering gender gaps for cybersecurity awareness of students, the above chart shows the highest gap is 22% for Changing social media privacy settings and the lowest gap is 5% for Giving the password to someone else. According to this, unbiased to gender, the cybersecurity and cyber safety such as Installing & updating antivirus programs, Enabling computer's firewall and Changing social media privacy settings, Ability to identify spam email are widely used among higher education students. However, it is somewhat surprising that 35% of the respondents believed that their information is of no use to hackers. This shows that the users are unaware of the value of personal information, especially in the hands of third parties or cybercriminals who could misuse such information in various ways.

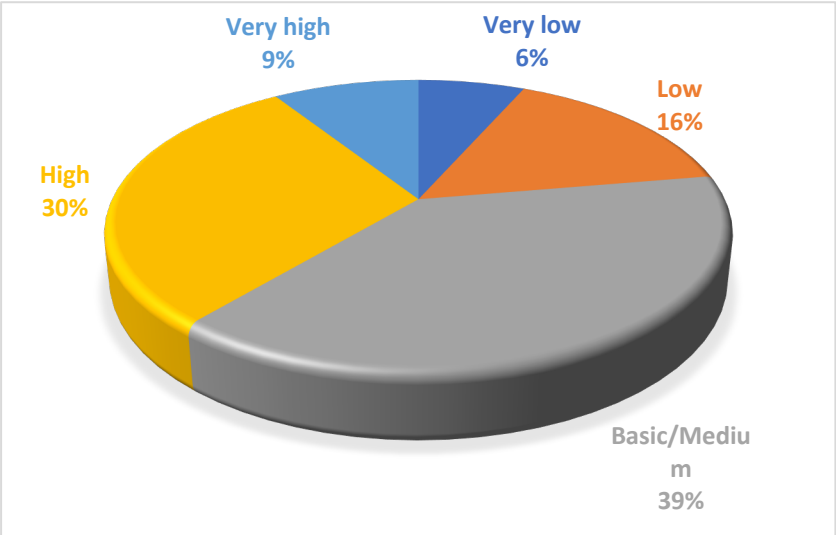

Figure 4: **Level of Cybersecurity Awareness**

5. Conclusion and Contribution

Overall, the survey participants seem well aware of the most common type of security enhancement tools and techniques namely using antivirus programs, enabling firewalls, opening trusted email attachments only and ability to identify spam email. However, both male and female students' awareness of new threats are not adequate and necessity of capacity building to improve student's knowledge on the online victimization. For instance, the value of privacy needs to be aware such as third-party involvement and cybercrimes such as misuse of personal information in a different manner.

Findings of the present work concluded that there is a significant difference between the awareness level of male and female users of internet services and it was established that the male students are more aware for cybersecurity in comparison to their female counterparts. The conclusion showed that the level of cybersecurity awareness among university students in Sri Lanka is not significantly low, but there are some knowledge gaps with new threats.

Cybersecurity is, however, a complex subject area usually surrounded by a lot of misuse of online victims. For this reason, awareness creation is avoidable in the fight against cybercrime. This, therefore, necessitates the need for a national cybersecurity awareness policy that focuses on students as key stakeholders in the education sector.

5.1 Limitation

The study suffered from several limitations that included, limitation in sampling procedure, participants not being open in responding to the questionnaires. Some of the participants' interaction was also hindered by the fact that they would be victimized. To overcome these challenges, the researcher explained the objectives of the study and assured them that the information collected would be treated with great confidentiality. Nevertheless, these limitations do not diminish the significance of the reported results as a whole.

5.2 Contribution and Future work

In this study, the significant amount of information was gathered from the respondents with respect to their opinions on cybersecurity in order to increase the level of student awareness. This research was more specific to students' experience and level of awareness and very detailed on the topic compared to other previous researches conducted by the foreign universities. This was the first research which conducted in Sri Lankan universities and in order to have more information on this topic it is best for this research to be continued and thus a further research should be conducted on students' experience and level of awareness in all academic institutions in Sri Lanka.

References

Abeysekara, E R D Liyanarachchi, M Wijesinghe, W S Jayarathne, N Wijethunga, M T N Perera, M 2012, 'Cyber Terrorism; is Sri Lanka Ready', General Sir John Kotelawala Defence University, Sri Lanka.

AP (Asian Pacific) Cert, 2016 Annual Report, URL: https://www.apcert.org/documents/pdf/APCERT_Annual_Report_2016.pdf site accessed: 20/08/2018.

Atkinson, S Furnell, S Phippen, A, 2009, 'Securing the next generation: enhancing e-safety awareness among young people'. Computer Fraud & Security, issue 7, pp. 13–19.

Babbie, E 2005, The Practice of Social Research, Belmont, California: Wadsworth.

CIA (nd) Sri Lanka, URL: https://www.cia.gov/library/publications/the-world-factbook/geos/ce.html, site accessed: 20/08/2018.

Christensen, J 2003, 'Solving the cybersecurity problem: The role of the Department of Homeland Security', http://www.wise-intern.org/journal/2003/jchristensen.pdf 26/09/18.

Computer Crime Act-No. 24, 2007, URL: http://www.slcert.gov.lk/Downloads/Acts/Computer_Crimes_Act_No_24_of_2007(E).pdf, 18/08/2018.

Connolly, C Maurushat, A Vaile, D Dijk, D V 2011, 'An overview of international cyber-security awareness raising and educational initiatives', A Galexia research report, with assistance from the Cyberspace Law and Policy Centre at the University of New South Wales – Australian Communications and Media Authority (ACMA), Australia.

Cooper, D R Schindler, P S 2003, Business Research Methods, 8th ed. New York: McGraw Hill.

Creswell, J. W., 2013, Research design: qualitative, quantitative, and mixed methods approach (4th ed.). Thousand Oaks, CA: Sage Publications, Inc.

Daily News, 2016, URL: http://dailynews.lk/2016/11/18/local/99473, site accessed: 20/08/2018.

Geer, D 2015, 'Six key areas of investment for the science of cyber security', The Futurist, no. 1, p. 10, viewed 4 October 2018, <http://ezproxy.deakin.edu.au/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edsgao&AN=edsgcl.428176483&authtype=sso&custid=deakin&site=eds-live&scope=site>.

Electronic Transaction Act-No. 19, 2006, URL: http://www.slcert.gov.lk/Downloads/Acts/ElectronicTransactionActParliamentver(E).pdf, 19/08/2018.
Joode, A D 2011, 'Effective corporate security and cybercrime'. Network Security, issue 9, pp. 16–18.
Kritzinger, E Bada, M Nurse, J R C 2017, 'A study into the cybersecurity awareness initiatives for school learners in South Africa and the UK', 10th World Conference on Information Security Education- WISE10, 29-31 May 2017, Rome (Italy).
Kritzinger, E Padayachee, K 2013,'Engendering an e-safety awareness culture within the South African context'. In AFRICON, 2013, pp. 1–5. IEEE.
Kritzinger, E Solms, S H V 2012, 'A framework for cybersecurity in Africa'. Journal of information assurance and cybersecurity, vol. 2012, doi: 10.5171/2012.322399.
Kritzinger, E Solms, S H V 2010, 'Cybersecurity for home users: A new way of protection through awareness enforcement'. Computers & Security, vol. 29, no. 8, pp. 840–847.
Lange, M D Solms, R V 2011, 'The importance of raising e-safety awareness amongst children', In Proceedings of the 13th annual conference on World Wide Web applications, p. 14.
Lange, M D Solms, R V 2012, 'An e-Safety educational framework in South Africa'. In Proceedings of the Southern Africa Telecoms and Network Applications Conference.
Palliyaguru, R 2015, 'Assessing the Threat of Cybercrime', GLACY International Conference, Sri Lanka.
Riem, A 2001, 'Cybercrimes of the 21st Century', Computer Fraud & Security, issue 4, pp. 12–15.
Senarathna, I, & Warren, M. 2017, A Sri Lankan hacking case study. In Valli, C. (Ed.). 2017, The Proceedings of 15th Australian Information Security Management Conference, Edith Cowan University, Perth, Western Australia. (pp.64-69)
Stone, K 2013, Keeping children and young people safe online: balancing risk and opportunity, http://www.eastrenfrewshire.gov.uk/CHttpHandler.ashx?id=11003&p=0: 26/09/2018.
Sri Lanka Cert (nd) http://www.slcert.gov.lk/aboutUs.php, URL: http://www.slcert.gov.lk/aboutUs.php, site accessed: 07/09/18.
Thomson, K L Solms, R V Louw, L 2006, 'Cultivating an organizational information security culture', Computer Fraud & Security, issue 10, p. 7–11.

DECEIVING AUTONOMOUS DRONES: SOME IMPLICATIONS?

W Hutchinson
Security Research Institute
Edith Cowan University, Western Australia
Email: w.hutchinson@ecu.edu.au

Abstract
This speculative paper examines the concept of deceiving Autonomous Drones that are controlled by Artificial Intelligence (AI) and can work without operational input from humans. Management and control by humans differ from that of AI even though at the superficial level, they have similar processes. This paper examines the potential of autonomous drones, their implications and how deception could possibly be a defence against them and /or a means of gaining advantage. It posits that although no truly general, autonomous drone exists at the moment, the development of AI and other technologies could expand the capabilities of these devices which will inevitably confront society with a number deep ethical, legal and philosophical issues. This exploration of autonomous drones and the concept of deception surfaces contradictions as deeper deception really fools the consciousness which is still not fully understood. It raises the question of whether drones or intelligent robots generally can truly have consciousness thus enabling true deception. It does not provide definitive answers but, hopefully, exposes a number of issues that will stimulate further research in this general area.

Keywords
Deception, robots, drones, artificial intelligence, security.

1 INTRODUCTION

In the last two decades, the term ‘drone’ usually meant a flying robot but has since been expanded to include any mobile robot. In this paper, ‘drone’ and ‘robot’ are used interchangeably. They are now found in the aerial, terrestrial, aquatic and space environments. Combined with artificial intelligence and a myriad of sensors, they have become formidable weapons and surveillance platforms (see Dougherty, 2015 for the range involved). In fact, defence against them is difficult for all but the most well-resourced entities. This phenomenon stimulated the start of this research which concentrates on autonomous rather than just automatic robots. The US Department of Defence (US DOD, 2014, p15) gives a simple explanation that an autonomous robot as:” … when the aircraft [drone] is under remote control, it is not autonomous. And when it is autonomous, it is not under remote control.” In other words, it is independent of humans for its operating actions.

When considering the ‘intelligence’ and ‘knowledge’ aspects of this topic, it is useful to look at the types of systems that have been developed as these

types of systems. Cummings (2017) states there is a hierarchy of knowledge systems starting with skills-based behaviours, then rules-based, then knowledge-based and finally expertise-based. Skills-based relies on the perception-cognition-action loop and can be automated without much difficulty. As the need for complexity increases, multiple and compound processes can be accomplished by Rules-based learning. The next two levels of system require a higher level of learning where Knowledge-based reasoning is needed where the stored set of rules does not match the existing environment so a new set of rules have to be created. Expert-based systems use judgement and intuition. Although the move from automated to autonomous systems changes at the rules-based level; it is really at the Expert level that solutions to the ambiguities in the environment can start to be trusted. Cummings (*ibid*) contends that, there are no truly reliable autonomous systems relying on Knowledge-based or Expert based systems, in operation at the moment. Hence, whilst there are many automated systems there are no truly, fully autonomous ones.

The other underlying rationale and emphasis of this paper is human 'security'. Security is fundamentally based on two approaches – overwhelming the opposition ('force') or deceiving them. Defensive security can involve such passive defensive approaches as obstacles (ranging from physical obstruction to nested passwords) to more dynamic factors such as 'honeypots'. Offensive approaches can be physical defence or active deception. This paper examines the latter. An assumption is made that almost all security measures, both offensive and defensive, involve some deception. Such is the surveillance capability (and increasingly weaponization) of drones that the security function of their targets can often be severely compromised. Thus, to protect a targeted asset means the drone and its sensors and command and control systems (C2) must be compromised by destruction or such means as manipulation of parts, hacking the C2 systems or physical approaches such as dazzling: see (Bennett and Waltz, 2007:17-66 for the various methods that can be used). However, on the surface, deception as a strategy would appear to be a plausible approach despite the dominance of the drone especially its system's sensory range especially if a human controller was in the decision loop. The drone's sensors and digital systems would probably have a much faster decision processing than that of a human controller. However, if a human pilot was involved then the known deceptive techniques could be employed to fool the pilot, the drone and, ultimately, its mission. These techniques would partially rely on the corruption of the data coming from the sensors and, also, the manipulation of the cognitive abilities of the human controller. The latter techniques have been documented widely (examples are: Harrington, 2009; Malin *et al*, 2017). However, the potential advent of autonomous drone systems with no human mission control would give the advantage to the drone system with its superior sensory and processing speeds. It should be noted that although that some truly autonomous drone systems (with simple parameters of action) are in existence and deployed, few will admit their operational status. Autonomous drones are increasingly attractive. To the military and industry, they are a source of 24/7 workhorses without the expensive costs of pilots and associated problems of trauma with their human controllers observing

the results of their activities. Certainly, there are humanitarian concerns but the economic and strategic/tactical viewpoints seem to be increasingly over-riding these issues (Walsh, 2018).

The willingness for political and management systems to consider human-less controlled systems of massive destructive capability can be illustrated by a Cold War example (Smith, 2008). This plan was considered by the Soviet system and 'nearly' implemented. It consisted of a crewless ship, packed with nuclear material and a cobalt nuclear device (effectively a globally effective radiation enhanced 'dirty bomb') and was to cruise Arctic waters. Sensors on this ship would register any excessive radioactivity and when the level of radiation passed a predetermined measurement, it would be assumed by the system that the Soviet Union, its leadership and its population had been destroyed. This would cause the ship-based control system to detonate the cobalt bomb, and contaminate the whole globe with radioactivity (Smith, 2008). Even at a superficial level, holes in this system are apparent. It was not implemented but was seriously considered although, it should be pointed out that certain commentators note that this doomsday machine is still in existence, and armed and ready to go. Some say that it was a double bluff by its creators to control internal power groups who might be tempted to attack the West (Keim, 2007; Torchinsky, 2017). Like the Mutually Assured Destruction (MAD) option below; what would be the point of starting a war if your own side would be destroyed? This mindset was also present in the West. It can be seen that despite irreversible consequences, trust in in an autonomous system was thought to be worth the risk.

This is also true of the MAD strategy of the West where total global destruction was possible and very nearly occurred on a number of occasions. It was avoided simply by humans over-riding the technologically driven decision-making systems. These 'automated' systems were put there because of the speed required for decisions to be taken (Ellsberg, 2017). It should be pointed out that the decisions that interrupted the prescribed series of events, were not a part of the normal process but because humans were aware of the consequences, the over-ride could occur. However, it does illustrate circumstances where these systems are regarded as serious options for high impact decisions for the human race. At the ethical level, Haas and Fischer (2017) discuss the use of autonomous drones in targeted killing of human targets, of course, comparatively fewer victims are affected than those above.

It is relevant that the United States' contemporary military doctrine - *the Third Offset Strategy* - relies heavily on modern technology and artificial intelligence including autonomous learning systems to gain an asymmetric advantage which will increase reliance on automated decision-making systems (Miranda, 2018). Hence, the significance of considering the impact of deploying these machines. The CSIS (2017) further explain that this strategy has many other facets such as evaluation, innovation, training and purchasing practices. However, its key areas in technological emphasis are:

- Autonomous learning systems
- Human machine collaborative decision making
- Assisted human operations,
- Advanced and unmanned systems operations
- Networked-enabled autonomous weapons, and
- High speed projectiles

It can be seen that autonomy for drones and AI developments are high on the priorities. It is interesting to note that Payne (2018) states that AI is the third offset.

To explain the structure of this paper the sections are as follows. Section 2 provides an overview of the processes involved in human deception, whilst section 3 outlines the types of deception that can be used to execute a deception plan. Section 4 examines the theory and possible tactics of deceiving a robot. Section 5 moves into the topic of 'swarming' drones and the possible future structures of swarms. Section 6 examines the idea of robots/drones as sentient objects and the dilemmas posed to society by this concept. Section 7 lists a series of possible research areas exposed by this paper in the area of deception theory and drones whilst, Section 8 concludes the paper some relevant comments and this is followed by the References used in the production of this paper.

2 THE PROCESS OF DECEIVING HUMANS

Humans have been deceiving other humans since history began, it appears to be a necessity for group and individual survival (see: Clarke and Mitchell, 2018; Bell and Whaley, 1991; Godson and Wirtz, 2002). The process of human deception is based on changing an individual's or group's sense of reality and guiding a series of internal assumptions to create a sense of reality that is beneficial to the deceiver. This involves manipulating the inputs to the human and, where possible, to add/deprive or generally manipulate the inputs to the human sensory systems and provide an environment where by the cognitive systems come up with beneficial behaviours and / or beliefs to the deceiver. Clark and Mitchell (2018, p.9) define deception as:

> *… a process intended to advantageously impose the false on a target's perception of reality.*

This paper is mainly concerned with deception in a military or security context which is perceived in the West to be at the intersection of the disciplines of Counter-intelligence/ Intelligence and Psychological Operations. However, there are potential attacks on civilian autonomous drones involved as well, in such activities as road and water surface transport as well as tipping trucks used in mining and smelting – thus taking it into the realm of industrial security.

Recent discoveries in cognitive science have shown how dynamic human systems are such as, the non-permanence of memories and how imprecise the beliefs of reality can be for a model-based explanation between humans and computers in their differences (Whitworth and Ryu, 2009). Amin and Malik (2013) give a good, experimentally based description of the types of

human memory based on the fundamental units of short and long term memory.

A simplified version of the human process of developing beliefs and behaviours is:

- An external physical or imagined event
- Analogue data produced from environment or device
- Senses register the event and its environment
- Data are processed by the nervous system, input data interpreted by brain and put into context
- Memories are 'updated' and data added to the knowledge base, and
- Beliefs, Behaviours and Memories are updated. (Hutchinson, 2006).

Deceiving humans is a complex process that requires formulating individual and social scenarios to 'massage' the cognition of the target to change its worldview. It involves understanding the complexity and dynamism of the human target and maintaining that deception in a context where it is still credible, and where the target is potentially aware of deceptive behaviours by the attackers.

Human perceptions of reality are governed by the health of the body (especially the brain and nervous system) and the accuracy of the body's sensors and those secondary artificial sensors used as input to the individual's perceptual system as well as the human's relationship and/or stress with their general relationship with the external environment. This is further complicated by the observation that 95% of brain activity is unconscious. Unconscious cognitive biases such as anchoring (focusing on the first factor encountered), clustering (observing phantom patterns that confirm preconceptions) and confirmation bias (preferentially using information that matches your preconceptions) also complicate the situation– a longer list of these biases can be found in Young (2018). These provide some of the material to deceive individuals. Other phenomena which involve groups such as Groupthink (Janis, 1982) can be exploited by the deceiver.

However, human consciousness is an important factor in a sophisticated deception. Contemporary experts are still discussing whether intelligent machines can have consciousness (Barrat, 2013, pp.45-46). As Clark (2014, p. 24) observes "simulation is not the same as instantiation".

3 TYPES OF DECEPTION

There have been a number of ways the principles of deception have been classified. Bennett and Waltz (2007, p.59) state that there are four principles:

- **Truth**: deception works in the context of what is true
- **Denial** denies the deceived with real and accurate data
- **Deceit** provides the target with false, wrong, or misleading data, and
- **Misdirection** which manipulates the target's attention and focus.

Whaley (1969/2007) developed tactics for using deception. Here deception was broken up into Level 1 Dissimulation (Hiding the Real) and Level 2 Simulation (Showing the False). Level 2 is always a part of Level 1.

Dissimulation (Hiding):

- **Masking** (basically means blending in for example, camouflage)
- **Repackaging** (where something is given a new 'wrapping')
- **Dazzling** (obscuring pattern confounding the target for example, using codes, or physical smoke confounding the target)

Simulation (Showing the False):

- **Mimicking** (copying/imitating pattern)
- **Inventing** (producing replicas, which have one or more characteristics of reality)
- **Decoying** (creating an alternative pattern misdirecting the attacker)

The above tactics cannot be used effectively unless they are based on a solid set of objectives imbedded in a strategy. This should lead to a dynamic 'story' that reinforces the deception. The environment of the story might well change as might the assumptions about the external environment. However, it must always be credible. Humans can be very inventive and amend the story and to make credible changes to the external environment. AI driven drone are more problematic as the environments and assumptions made, also change. The dynamic brain of humans can rationalise and compensate for unexpected changes that occur. Whilst there is a lag in the human brain, cognitive dissonance can often handle this. Only superior design and programming can cope with this sort of fundamental change. Kirk (2017) brings up the concepts of reflexivity and autonomy, each of these would need to be resolved before the nature of robot autonomy rather than automatic functionality could be determined.

All of these theoretical aspects of organised deception were formulated to ultimately deceive human targets although, it was assumed some sophisticated technology would also be applied to the process. However, the direct target in this paper is concerned purely with deceiving autonomous drones and so the emphasis and assumptions made are different although many of the strategies can be the same. The unanswered question is: can a drone be truly deceived? In other words, do these deceptive processes just disrupt the automated behaviour of the drone or truly have some effect on the higher autonomous function of the drone's control system?

4 THE PROCESS OF DECEIVING AUTONOMOUS DRONES

Deceiving drones requires an indirect knowledge of:

- the human input to the design of the robot
- the sensor equipment
- the logic of the management and control system initially designed by the creators
- any updated logic derived from the software itself as the software learns by experience; much like living organisms do.

Drones need intelligent processes which can access its worldview (both of which can be changed by internal processes or external attacks on its

sensors or directly to its C2 systems (both the ‘processes’ and its stored ‘worldview’).

AI relies on probabilistic processes so is similar to human learning but cannot use induction efficiently (Cummings, 2017). At the moment, a drone’s technical technological and cognitive strength is in the area of deduction. The C2 is probably more predictable than that of a human but not necessarily. Drones gain information of their surroundings using their sensors (just like humans) and send signals from their microcontrollers to motors (Wahde, 2007).

The series of processes of autonomous drones are similar to that of human beings, that is:

- External event
- Event sensed
- Situation abstracted
- Digitised and processed
- Interpreted (in context?)
- Behaviour and memory updated. [Action instigated].

The deceiver can potentially use manipulative techniques at each stage, for instance:

- Camouflage the event
- Disturb the sensing process by signal manipulation or sensory saturation
- Hacking the software abstraction and interpretation processes
- Hacking the hardware digitisation processes such as memory or buffering updates.

Before continuing, some details of the structure of the drone should be examined. The working environments can vary enormously from space to the deep sea as can the type of mission (surveillance or/and armed attack, and/or transport). Regardless of the variety of size, mission and other variables, the commonality of each will be assumed for this discussion. Common factors are assumed to be:

1) A management and control system (previously called C2 above) which consists of a ‘brain’ that is, an AI system with appropriate dynamic and permanent memory driving the physical functions in the drone). This control system consists of dynamic internal logic which manages the actuators and power source(s) as well as communicating with the AI system. Whilst dynamic learning by AI systems has improved in recent years, most non-classified systems have not yet been able to emulate the self-taught abilities of a human child (Kwon, 2018).
2) A power source to drive motive and internal systems.
3) A series of actuators.
4) A navigation system driven by the control system and an appropriate series of sensors controlled by the AI systems.
5) A communication system able to communicate internally, and with other drones or appropriate destinations.

6) Proprioceptors for the measurement of the robot's (internal) parameters (it should be noted here that the communication to a base in non-autonomous drones is their weakest security link);
7) A series of sensors for both internal monitoring (proprioceptors) and the external environment (exteroceptors)

Whilst this information can be used for an attack strategy on an autonomous drone, this paper is about the deception function which can be used to destroy or manipulate an entity. However, it is often more beneficial to use the target for a higher deceptive purpose to maximise the attack. For instance, to 'take over' the navigation system allowing the drone to land or crash in an area where the drone/wreckage can be examined and reverse engineered. The potential attack methods are numerous, for example, GPS tampering, and hacking the control system. This multiplies when sensor types increase and tend to be drone- and sensor-target specific. At the moment the sensors tend to be cameras (light and heat sensitive) but they can extend to anywhere on the electro-frequency spectrum. Other sensors can include Acoustic (especially useful in underwater drones), Nuclear (these can identify different materials and the types of radiation present), Chemical (for instance, nerve agents), Biological (for example, signatures for microscopic disease vectors) and Biometric Signatures. (Clark, 2011) This paper only attempts to outline the principles but the previous examples, hopefully, enable the reader to see the potential. An interesting drone is the stealthy and autonomous robot spider illustrated in Deakin (2010, p.452). The potential number of sensors illustrates the intelligence capabilities of drones, and hence the potential of deceiving such a device. Whether, the attack is on the control system by hacking, the destruction of the communication system, or the manipulation of sensors will be determined by the desired outcome, technical feasibility and the likelihood of being exposed by the target.

As the number the number of drones increase so do their threats and counter-measures. The number of proposed and enacted drones given by the work of Jacobsen (2015) certainly shows the increasing threats of autonomous robots as well as the benefits to certain functions. Of course, the control system might make the drone automatic, but it does not necessarily make it autonomous. This ability is the one of the emphases of this paper. To truly deceive something rather than just disrupt it, the target drone needs to have 'awareness' and 'consciousness'. These are both problematic. Holland and Goodman (2003) imply that as engineers make the control systems produce more intelligent behaviour then 'consciousness' might evolve. However, this is a quite a contentious point. Obtaining the attributes of consciousness would make the deceptive tactics listed in Section 3 possible.

5 THE IMPACT OF SWARMING PRACTICES

There is a need here to bring in the phenomenon of 'swarming'. This is the coordinated use of various drones which might be of different types, 'intelligence', size, and capabilities so they can act in unison. This complicates the act of deception enormously. This use of swarming techniques where numerous drones are used for one purpose is increasing

interest in this technique. The decreasing cost of smaller drones (Hambling, 2015) plus the built-in redundancy of swarms makes the use of many drones for an attack much more appealing. Also, it can make deception much more difficult as some drones that are disabled will still leave others to carry out the mission. At present, these 'swarming drone systems' seem to be considered for underwater protection of valuable assets such as submarines, and providing surveillance for military units at a cheap cost. The initial use of 'tethered' drones linked to a 'mothership' gives protection to the controlling vehicle and its crew which, in turn, gives extra surveillance facilities and, possibly fire-power as well as cover by providing sacrificial drones to the central control. This concept developed into autonomous swarms whereby each drone was independent but kept communications with other drones and acted like one entity much like a flock of birds (Singer, 2009). This implementation gives the group a lot more power and is much more difficult to deceive unless its elements are very simple. However, these self-organising swarms can lose some members without losing too much functionality – deceiving the swarm will be harder than deceiving the individual. Nevertheless, swarms, because they need to link up with each other, are more vulnerable to infection from malware and ironically this could be a weak point where software encouraging a deceptive move could be inserted.

Underwater drones do have a communication problem especially when not tethered to a 'mother ship' as communication signals are attenuated by the water medium. Signals are sent by radio and acoustic means or by light (blue has been used up to now). However, this has been partially overcome by using each drone being arranged into lines and passing the signal from one to another thereby extending the range much as classical network technology does.

The concept of swarms came out of a need to find asymmetric approaches to developing terrorist and insurgent approaches to war. From the US perspective, the enemy in the early 21st century tended to be relatively small dispersed groups compared to conventional forces. Although these tactics were not really new (see Arquilla and Ronfeldt, 2000), they did seem to be needed to compensate for the large, hierarchical forces which did not prove as flexible and speedy as these small groups. With the development of military drones, and with the continuing advancement of them, came the technological ability to produce smaller and more flexible varieties. As this development advanced, the increased communications and AI techniques allowed an ever-increasing potential of these machines to provide to develop drone swarms. The extension of network theory allowed the development of intelligent swarms which broadly can be hierarchical or networked (in an organisational sense).

Swarms can be designed by much larger entities as well, and the development of swarming systems can allow each element to work independently and come together in a swarm when needed so groups of drones can be expanded or decreased as the problem being tackled varies. Hence, drones of various abilities and form as well as environmental function, can be coordinated as necessary. This ability is very powerful, and

would require a deceiver to work at a population level rather than targeting an individual drone.

6 DRONES AS SENTIENT OBJECTS

In November 2017, a humanoid robot - Sophia - was awarded full citizenship Saudi Arabia (Jakarta Post, 2017). This development raises enormous implications for the use of drones. The recognition of drones and human is becoming closer.

An interesting point brought up in one of the discussions given with Sophia is 'her' claim that she could not be killed as her essence was still in the cloud. Is this immortality? Would destroying an autonomous drone with a passport be considered murder? Potentially, the acceptance of a robot as a citizen is a revolutionary event. Watching Sophia is an interesting introduction to the dilemmas posed by autonomous drones. In fact, the interviews shown in the reference raises more questions than can be explore in this paper.

7 FURTHER RESEARCH

There is a need to understand why these machines are being designed/produced and the foreseeable implications of this development should be investigated. The philosophical and reasoning processes for designing these drones and their ability to deceive and recognise deception should be programmed into them. The use of bottom-up as well as top-down thinking should be used and the deduction-retroduction-abduction-induction cycle (Waltz, 2003) should be investigated to raise the level of autonomous drone to the 'expertise' level (Cummings, 2017). Only then can autonomous drones on life threatening missions be trusted.

The design of communication topologies and 'nervous' systems of swarms could be investigated to mimic nature and develop systems that better suit the technology and nature of the task for instance, the decentralised but radially distributed system of most jellyfish (Kasuki and Greenspan, 2013) or the strange distribution of the octopus Godfrey-Smith (2017) which has a central brain but a distributed smaller brain in each of its eight tentacles – the central brain can take over all of the others but sometimes as required each 'tentacle brain' can over-ride this control. The potential of this format for tethered drones is enormous. If networks of any types and capacities drones could be connected together using this topology and could function like the system of the octopus; the system could work independently then when necessary attach to be a coordinated single system only to disperse the processes when necessary

The main potential areas of research are in the implications of sentient drones and their impact of the legal, social, organisational, ethical, philosophical and work practices as well as their impact on our understanding of deception and the human cognitive environment as well as the development of AI. The use of the theme of deception is useful as it includes both the physical and the cognitive and encompasses such areas as philosophy, consciousness studies, robotics, AI and engineering.

8 CONCLUSIONS

The use of autonomous drones will possibly increase in the future creating a security threat to many because of their surveillance abilities as well as their armed potential. As the number of autonomous drones increase and their variety multiplies, the shaping of the human environment in terms of social behaviour, crime, privacy, and a multitude of other factors will need to be considered. This increase and dynamism of technology has profound implications of all sections of society. Elon Musk who said AI is like: "summoning the demon" (Bartlett, 2018, p.111) is a summary warning from a noted pundit. Drones and AI are disruptive technologies and their impacts should be considered. This paper is an attempt in the narrow area of deceptive practices to advance the possible scenario in the contemporary environment to protect against undesired effects of autonomous drones in the areas of surveillance (spying and privacy) and threats of force.

As a defence against negative effects of drones, deception is a potential tool. However, drones themselves will be increasing in their own potential to deceive using new technologies such as holograms and neuro-hacking (Malin, *et al*, 2017). These need to be considered in the ongoing debate about the increasing use of drones especially autonomous models.

This research started as an attempt to find methods to minimise the negative effects of drones – a means of defence. It developed from that into an examination of deception and the relationship between humans and intelligent machines. There is much to be learned but awareness is essential. The intent of this paper is to provide a background to further stimulate investigation into the area.

REFERENCES

Amin, H. and Malik, A.S, (2013) Human memory retention and recall processes: A review of EEG and fMRI studies, *Neurosciences*, **18**,4: 330-344.

Arquilla, J and Ronfeldt, D (2000) *Swarming and the Future of Conflict*, Santa Monica, RAND.

Barrat, J. (2013) *Our Final Invention: Artificial Intelligence and the End of the Human Era,* New York, Thomas Dunne Books.

Bartlett, J. (2018) *The People vs Tech*, London, Penguin Radom House.

Bennett, M., Waltz E. (2007) *Counter deception: Principles and Applications for National Security*. Norwood, MA, Artech House.

Bell, J.B., Whaley, B. (1991) *Cheating and Deception*, New Brunswick, USA. Transaction Publishing.

Clark, A. (2018) *Mindware,* (*second edition*), New York, Oxford University Press.

Clark, R.M. (2011) *The Technical Collection of Intelligence,* Washington, CQ Press.

Clark, R.M., Mitchell, W.L. (2018) *Deception and Counter Deception*, Washington, SAGE.

Cummings, M.L. (2017) *Artificial Intelligence and the Future of Warfare,* London, Chatham House.

CSIS (2017) *Assessing the Third Offset Strategy*, Washington, Center for Strategic and International Studies. March 2017.

Deakin, R.S. (2010) *Battlefield Technologies: Networked-Enabled Information Dominance*, Boston, Artech House.
Dougherty, A.J. (2015) *Drones*, London, Amber Books.
Ellsberg, D. (2017) *The Doomsday Machine*, London, Bloomsbury.
Gerwehr, S., Glenn, R.W. (2000) Unweaving the Web: Deception and Adaptation in *Future Urban Operations,* Santa Monica, RAND.
Godfrey-Smith, P. (2016) *Other Minds: The Octopus, The Sea, and the Deep Origins of Consciousness,* New York, Farrar, Straus and Giroux.
Godson, R, Wirtz, J.J (2002) *Strategic Denial and Deception,* Transaction Publishers, New Brunswick.
Goodman, M. (2015) *Future Crimes*, London, Transworld Publishers
Haas, M.C. and Fischer, S.C. (2017) The evolution of targeted killing practices: Autonomous weapons, future conflict, and the international order, *Contemporary Security Policy*, **38**:2, 281-306.
Hambling, D. (2015) *Swarm Troopers*, Archangel Ink
Harrington, B. (Ed.) (2009) *Deception: From Ancient Empires to Internet Dating*, Stanford, Stanford University Press.
Holland, O, Goodman, R (2003) Robots with Internal Models*, Journal of Consciousness Studies*, **10**:4/5:2-45.
Hutchinson, W. (2006) Information Warfare and Deception, *Informing Science*, **9**: 213-223.
Jakarta Post (2017, Oct 29) *Meet Sophia: The first robot declared a citizen by Saudi Arabia.* Online: https://www.youtube.com/watch?v=E8Ox6H64yu8 [Accessed 05 November, 2018]
Jacobsen, J. (2015) *The Pentagon's Brain*, New York, Little Brown and Company.
Janis, I.L. (1982) *Groupthink* – 2nd Edition, United States, Houghton Mifflin.
Katsuki, T. and Greenspan, R.J. (2013) Jellyfish Nervous Systems, Current Biology,**23**, 14:592-594
Keim, B. (2007) *Soviet Doomsday Device Still Armed and Ready*. Online: https://www.wired.com/2007/09/soviet-doomsday/ [Accessed 06 August, 2018]
Kirk, R. (2017) *Robots, Zombies and Us: Understanding Consciousness,* London, Bloomsbury Academic
Kwon, D. (2018) Self Taught Robots, *Sc. Am*., **318**, 3: 20-26 [March, 2018]
Malin, C.H., Gudaitis, T., Holt, J.H., Kilkger, M. (2017) *Deception in the Digital Age*, London, Academic Press.
Miranda, J. (2018) The Third Offset, in: *Modern War*, **36**, pp. 48-57, Bakersfield CA, Strategy and Tactics Press.
Payne, K (2018) Artificial Intelligence: A Revolution in Strategic Affairs? *Survival*, **60**,5:7-32
Singer, P.W. (2009) *Wired for War: The Robotics Revolution and Conflict in the 21st Century*, New York, Penguin Press.
Smith, P.D. (2008) *Doomsday Men: The Real Dr Strangelove and the Dream of the Superweapon*, London, Penguin.
Torchinsky, J. (2017) *The Soviets Made a Real Doomsday Device in the '80s and the Russians still have it today*, Online: https://www.wired.com/2007/09/soviet-doomsday/ [Accessed 06 August, 2018]

USDOD (2014) *Unmanned Systems Integrated Roadmap* FY2013-2038, Washington, Pentagon.
Wahde, M. (2007) *Autonomous Robots*, Chalmers University of Technology, <on-line>
Walsh, T. (2018) *2062 – The World that AI made*, Carlton, La Trobe University Press.
Walsh, T. (2017) *It's Alive*, Carlton, La Trobe University Press.
Waltz, E. (2003*) Knowledge Management in the Intelligence Enterprise*, Boston, Artech House.
Whaley, B. (1969/2007) *Stratagem: Deception and Surprise in War*, Boston, Artech House.
Whitworth, B and Ryu, H. (2009) A Comparison of Human and Computer Information Processing, in: *Encyclopedia of Multimedia Technology and Networking - Second Edition,* M.Pagani (ed.), Hershey, New York, Information Science Reference.
Young, E. (2018) Lifting the Lid on the Unconscious, *New Scientist*, **3188** (28 July, 2018), pp. 34-39.

Collaborative Task Allocations in Warfare Multi-UAVs With Time Coupling Constraints

Yaozhong ZHANG[1], Lan CHEN[1] , Frank Jiang*[,2], Jiandong Zhang[1]

[1]*School of Electronics and Information, Northwestern Polytechnic University, China*
[2]*School of IT, Deakin University, Burwood, Australia*
**Corresponding author: Frank.Jiang@deakin.edu.au*

Abstract: In this paper, based on the characteristic of SEAD (Suppression of Enemy Air Defense), a heterogeneous multi-UAV (Unmanned Aerial Vehicle) cooperative task assignment model with coupling constraints is proposed. We improved the firefly algorithm and put forward a discrete firefly algorithm which is based on the differential evolution operator. It adopted the segmented integer encoding, step update strategy and combined mutation, crossover and selection to reconstruct the individual. The simulation results show that discrete firefly algorithm can effectively solve the cooperative task assignment problem of multi-UAVs under the coupling task environment.

Keywords: Suppression of Enemy Air Defense (SEAD); coupling tasks; firefly algorithm; task assignment.

Introduction

UAVs (Unmanned Aerial Vehicles) are characterized with good stealth, strong autonomy and recovery. In addition, UAVs can replace the "boring, bad, dangerous" task of the pilot, reducing the casualties and minimising the cost of equipment. The military usages of the UAVs are increasing, and it is gradually turning to become the main combat force from the executive assistant in the mission [1][2]. Cooperative Multi-UVAs in the coupled task environment is a complex constrains problem of decision and optimization, and it is also a typical NP-Hard problem. Under the premise of meeting the various tactical targets, the main research problem is how to assign the tasks to each UAV and specify the specific execution order and time, so as to make the overall operational effectiveness as best as possible while meeting the various constraints [3][4].

Alighanbari and Kuwata [5] established a structure that can solve the task allocation problem with time series constraints, by using the mixed integer linear programming to obtain the optimal solution, and the suboptimal solution is obtained by the taboo searching algorithm, which greatly improves the efficiency of the solution. The literature [6] [7] considered the priority constraints in the single-task plan, which stipulates that a task with a lower priority must be assigned after a task with a higher priority, but it does not consider an exponential increase in the optimization time as the task type increases, it results in a difficult implementation of a large-scale

dynamic task planning. Mclain and Beard [8] also considered the multi-UAV cooperative space constraints, task execution timing series constraints, time constraints and other relations in types of collaborative constraints, but also brought a huge problem space search algorithm. Choi and Whitten [9] researched the problem of task planning with coupled constraints in complex tasks, it points out that centralized method is more conducive to solving the problem of coupling task than distributed method, but the centralized structure requires higher communication and computing.

Through the analysis of relevant literature, there are many achievements have been done in the research of collaborative task decision-making with coupling constraints. However, there are still fewer research on the cooperative mission assignment problem of multiple UAVs with time coupling constraints and special coupling constraints. Based on the SEAD combat background, this paper establishes a multi machine cooperative task allocation model in coupled constrained environment, describes the problem of multi machine cooperative task allocation, and proposes a hybrid discrete firefly algorithm based on differential evolution operator. By improving the standard firefly algorithm, the performance of the algorithm is greatly improved.

Problem Statement

Scenario: Assuming M numbers of UAV in three types (A, B, C), defining $U = \{U_1, U_2, \cdots, U_M\}$ as the sets of UAVs, and defining $U_S = \{U_1, U_2, \cdots, U_{M_a+M_b}\}$ as the sets of UAVs which can execute the reconnaissance missions, and defining $U_A = \{U_{M_a+1}, U_{M_a+2}, \cdots, U_{M_a+M_b+M_c}\}$ as the sets of UAVs which can execute the attract missions, the configuration information of UAVs is showing as the following Table 1. Supposing that there are N numbers if air defense suppression targets on a plane battlefield, and defining $T = \{T_1, T_2, \cdots, T_N\}$ as the sets of targets, the location information of each target is known, and in order to completely destroy the enemy's air defense system, it requires our UAVs to execute three tasks of confirming, attacking and damaging assessment for each air defense target in turn. The T_{jh} represents the mission h of the target T_j, $h = 1$ represents confirming task, $h = 2$ represents attracting task, $h = 3$ represents damaging assessment task. After completing the mission, each UAV must land at the designated base.

Table 1 **Details of UAVs**

Type of UAVs	Function	Quantity	Number
A	Reconnoiter	M_a	$U_1, U_2, \cdots, U_{M_a}$
B	Reconnoiter /Attack	M_b	$U_{M_a+1}, U_{M_a+2}, \cdots, U_{M_a+M_b}$
C	Attack	M_c	$U_{M_a+M_b+1}, U_{M_a+M_b+2}, \cdots, U_{M_a+M_b+M_c}$

Objective function

The minimum maximum range of UAV is chosen as the task planning index, which minimizes the maximum range of the UAVs, and guides the task allocation strategy to minimize the direction of each UAV, that is

$$F = \min(\max_{i=1,2,\cdots,M} Voy_i), \quad i = 1,2,\cdots,M \quad (1)$$

In this equation, Voy_i represents the mission range of U_i, $i = 1,2,\cdots,M$, M is the number of UAV.

It is assumed that the altitude and speed of UAV during mission execution are constant, range of the UAV U_i is:

$$Voy_i = v_i \times (eT_j - sT_j) + dis(T_j, BP) \quad (2)$$

In this equation, v_i is the speed of U_i; eT_j is the time that target T_j completes three tasks; sT_j is the time that T_j starts executing first task time; $dis(T_j, BP)$ is the Euclidean metric between target T_j and base, the equation is:

$$dis(T_j, BP) = \sqrt{(x_{BP} - x_{T_j})^2 + (y_{BP} - y_{T_j})^2} \quad (3)$$

x_{BP}、y_{BP}、x_{T_j}、y_{T_j} are the position coordinates of the base BP and target T_j.

$$sT_j = \begin{cases} dis(p_i, T_j) / v_i & n_i = 1 \\ eT_{(j-1)} + T_{i((j-1),j)} & \text{otherwise} \end{cases} \quad (4)$$

p_i represents the initial position of U_i, $T_{i((j-1),j)}$ represents the time required for U_i to fly from T_{j-1} to T_j, and the equation is:

$$T_{i((j-1),j)} = dis((j-1), j) / v_i \quad (5)$$

Constraint condition

(1) Each target contains three tasks to combat tasks, and each task must be executed

$$\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{h=1}^{3} x_{ijh} = 3N \quad (6)$$

Decision variables are $x_{ijh} \in \{0,1\}$, the values are as follows:

$$x_{ijh} = \begin{cases} 1 & U_i \text{ performs the task } h \text{ on } T_j \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

(2) Each task on each target can only be executed once

$$\sum_{i=1}^{M} x_{ijh} = 1 \quad (8)$$

(3) Each UAV is assigned at least one task, that is

$$\sum_{j=1}^{N}\sum_{h=1}^{3} x_{ijh} \geq 1 \quad (9)$$

(4) Range constrains

$$Voy_i \leq Voy\max_i \quad (10)$$

$Voy\max_i$ represents the maximum range of UAV.

(5) Constraints of weapon load resources in constrained range

$$\sum_{j=1}^{N} x_{ij2} \le R_i \qquad (11)$$

R_i is the number of weapons loaded for U_i.

(6) Timing constraint: this constraint is to satisfy the order constraint of three tasks in the task of acknowledgement, strike and damage assessment, which can be expressed as:

$$sT_{jh} + x_{ijh} * t_{ijh} \le eT_{jh} \qquad (12)$$

$$eT_{jh} \le sT_{j(h+1)} \qquad (13)$$

sT_{jh} represents target T_j starting time when executes task *h*, eT_{jh} represents the finishing time when T_j executes task h, t_{ijh} is the time consumption on task h.

(7) Time interval constraints: this constraint is designed to ensure that UAVs are affected by blast and smoke when they are damaged in the damage assessment of the target. There is a certain time interval between the damage assessment and the strike task. At the same time, in order to make the evaluation effective and effective, the time interval is not more than a certain limit. The constraint can be expressed as:

$$eT_{j2} + Inter\_\min \le sT_{j3} \qquad (14)$$

$$eT_{j2} + Inter\_\max \ge sT_{j3} \qquad (15)$$

$Inter\_\min$ represents the minimum time interval between strike task and damage assessment task

$Inter\_\max$ represents the maximum time interval between striking task and damage assessment task.

(8) Task simultaneous execution: the constraint refers to a certain tactical consideration or a task that requires two or more targets at the same time in order to achieve a particular effect, such as simultaneous strike. This article describes the simultaneous execution of constraints in a task as follows:

It is assumed that task A is executed at t_a, task B executed at t_b, given a fixed time interval $\inter$, if $|t_a - t_b| \le \inter$, that task A and task B meet tasks and perform constraints simultaneously. Assuming that both the target T_i and the target T_j must attack at the same time, they can be expressed as:

$$|sT_{i2} - sT_{j2}| \le \inter \qquad (16)$$

(9) Task priority constraints: this constraint refers to a certain level of priority for a target of two or more, and the higher priority needs to be executed before the target with lower priority is executed. As for some tactical consideration, the target T_k must be damaged before the T_l is confirmed, and it can be expressed as:

$$eT_{k3} \le sT_{l1} \quad (17)$$

Hybrid Discrete Firefly Algorithm Based on Differential Evolution Operator
The principle of standard firefly algorithm: Firefly Algorithm (FA) is propo sed by Xin-She Yang in 2008, it originates from the simulation of the sw arm behavior of firefly in nature, and it is a new high-level meta heuristic optimization algorithm. The mathematical description of the standard fir efly algorithm is as follows:

(1) The expression of the firefly: it is assumed that in the D dimension se arching space, there are groups of N numbers of fireflies in a group, the number i firefly's location is:

$$X_i = (x_i^1, x_i^2, x_i^3, \cdots, x_i^D) \quad i = 1, 2, 3, \cdots, D \quad (18)$$

In this equation, x_i^j represents the location of number i in j dimension, the initial location of each firefly is produced randomly.

(2) Absolute brightness: I_i is the absolute brightness of number i firefly, and also represents the objective value ate the location of fireflies. It is generally set as the absolute brightness value of firefly i in $X_i = (x_i^1, x_i^2, x_i^3, \cdots, x_i^D)$ is equal to the value of that objective function, that is:

$$I_i = f(X_i) \quad (19)$$

(3) Relative brightness: the relative brightness of firefly i to firefly j is expressed in I_{ij}, due to the brightness of firefly i decreases with the increase of the distance and the absorption of air, so the I_{ij} is :

$$I_{ij}(r_{ij}) = I_i e^{-\gamma r_{ij}^2} \quad (20)$$

In this equation, I_i is the absolute brightness value of firefly i; γ is the optical absorption coefficient, which is usually set as a constant numb er; r_{ij} is the distance between firefly i and firefly j, that is:

$$r_{ij} = \| X_i - X_j \| = \sqrt{\sum_{k=1}^{D} (x_{i,k} - x_{j,k})^2} \quad (21)$$

(4) Attraction Calculation: the attraction force of the firefly i to the firefly j is expressed is β_{ij}, supposing that the relative brightness I_{ij} of firefly i to firefly j is proportional, so by the definition of relative brightness, t he attraction of the firefly i to the firefly j is:

$$\beta_{ij}(r_{ij}) = \beta_0 e^{-\gamma r_{ij}^2} \quad (22)$$

In that equation, r_{ij} is the distance between firefly i and the firefly j; β_0 is the initial attraction force in r=0 position.

(5) Location update: if the fitness value of firefly i is better than firefly j, so the firefly i will move in the direction of firefly j,and its location update formula is:

$$X_j(t+1) = X_j(t) + \beta_{ij}(r_{ij})(X_i(t) - X_j(t)) + \alpha\varepsilon \quad (23)$$

In that equation, t is the number of iterations of the algorithm, α is a constant, ε is a random number vector derived from Gauss

distribution, uniform distribution or other distribution. Population initialization

In order to prevent the population from getting into the local optimal solution prematurely, the cosine similarity between the fireflies is compared in the process of random initialization, when the similarity exceeds the threshold of a certain set ξ, so re-initialized one of the fireflies. The cosine similarity formula is:

$$\cos(X_i, X_j) = \frac{X_i \cdot X_j}{\|X_i\| \times \|X_j\|} \quad (24)$$

Individual reconstruction

In order to further strengthen the information communication between the firefly individuals in the population, the firefly algorithm does not fall into the local optimal solution too early, introduces the differential evolution operator, and realizes the cooperation and competition between the fireflies by three kinds of operation of mutation, cross and selection.

(1) Mutation operation:

An improved differential evolution algorithm based on "local neighbourhood variation" and "overall neighbourhood mutation" are used to perform mutation operations on the task allocation part of fireflies.

1) The formula of local neighbourhood variation is:

$$L_i^t = X_i^t + \alpha_c \bullet (X_{n\_best}^t - X_i^t) + \beta_c \bullet (X_p^t - X_q^t) \quad (27)$$

In this equation, L_i^t is the mutated individual, $X_{n\_best}^t$ is the best individual in the neighbourhood in X_i^t, X_p^t and X_q^t are randomly selected neighbourhood individuals, α_c and β_c are local scaling factors, make $\alpha_c = \beta_c = F$, from the research in [11], F=0.85.

2) The formula of overall neighbourhood variation is:

$$g_i^t = X_i^t + \alpha_g \bullet (X_{best}^t - X_i^t) + \beta_g \bullet (X_{r1}^t - X_{r2}^t) \quad (28)$$

In this equation, g_i^t is the original firefly individual X_i^t which is a new individual after the t iteration, X_{best}^t is the best individual of whole firefly population in the t iteration, α_g and β_g are the dither scaling factors which are based on fixed scaling factor F, the value is gotten by this formula:

$$\alpha_g = \beta_g = 0.0001 * rand + F \quad (29)$$

3) The final mutation operator is obtained by combining the "local neighbourhood variation" and "global neighbourhood variation" by weighting, and get the final mutated factor:

$$V_i^t = \omega_g \bullet g_i^t + (1 - \omega_g) \bullet L_i^t \quad (30)$$

For the individual task sequencing part, the neighbourhood search-based mutation method is adopted, and its operation steps are as follows:

Step 1: In the individual task sequencing part, randomly choose r bits and all neighbourhoods of their ranking are generated.

Step 2: The fitness function of all neighbourhoods is calculated, and the best individual is selected as the progeny and the original individual is replaced.

(2) Cross operation:

In order to better balance the global search ability and local search ability of the difference operator, the paper [12] uses cross probability factor with exponentially increasing iterations:

$$Cr = Cr_{\min} + (Cr_{\max} - Cr_{\min}) \times \exp(-a \times (1 - t/T)^b) \quad (31)$$

In this equation, Cr is cross probability factor, $Cr_{\min}$ and $Cr_{\max}$ are minimum cross rate and maximum cross rate, and $Cr_{\min} = 0.4$, $Cr_{\max} = 0.6$, $a = 40$, $b = 4$, T is set as maximum number of iterations, t is current number of iterations. The value of the new individual U_i^t is:

$$u_{i,j}^t = \begin{cases} v_{i,j}^t & rand \le Cr \quad or \quad j = j_{rand} \\ x_{i,j}^t & otherwise \end{cases} \quad (32)$$

(3) Selection operation

Before choosing the operation, it is necessary to make illegal solution and correction to new individuals, so as to ensure that the new task assignment scheme after mutation and cross is effective and feasible. In order to maintain the constant number of offspring and to evolve towards a better direction, the next step of the algorithm is to choose greedy strategies, the individuals with better fitness values selected between the original individuals X_i^t and the new individuals U_i^t after mutation and crossover operation are retained to the next generation. The description of the selecting operation is:

$$X_i^{t+1} = \begin{cases} U_i^t & fitness(U_i^t) > fitness(X_i^t) \\ X_i^t & fitness(U_i^t) \le fitness(X_i^t) \end{cases} \quad (33)$$

The calculation of objective function:

After decoding the firefly individual, the task execution sequence of each unmanned aerial vehicle is obtained. Through the task sequence, the initial position of the UAV, the location of the target and the return base position, the expected execution time of each task in the sequence can be quickly calculated, but it is likely to be unable to meet the coupling constraints. In order to satisfy the coupling constraints, the coupling constraint matrix is introduced as $T^s \in R^{3N \times 3N}$, it represents that there is coupling constraint relationship between tasks. T^s represents the coupling constraint between task i and task j, the rule of value is:

$$T_{ij}^{s}=\begin{cases}\infty & \text{There is no special coupling constraint between} T_i \text{ and } T_j \\ 0 & T_i \text{ and } T_j \text{ have simultaneous constraints} \\ 1 & T_i \text{ must be executed before } T_j \\ -1 & T_i \text{ must be executed after } T_j\end{cases} \quad (34)$$

Simulation

Simulation environment: the PC involve in Intel 2.53FHz Main Frequency, 2G memory, Windows 7 operating system and Matlab2014a platform.

Case Study: it assumes that there are 7 UAVs and 10 targets to be destroyed in the mission scenario, the information of the UVAs is following as Table 2 below, and the targets information and reverting base information are following as Table 3 below. It supposes that the UAV completes the reconnaissance mission at 0.05h, the time to carry out the damage assessment task is 0.1h, and the time for carrying out the damage assessment task is 0.15h, the battlefield situation is shown as in Figure 3.

Table 2 **Initial information of UAV**

The number of UAVs	The type of UAVs	Initial position	Speed km/h	Load	Maximum voyage/km
1	Reconnaissance UAV	(0, 400)	120	/	5000
2	Reconnaissance UAV	(0, 150)	120	/	5000
3	Reconnaissance UAV	(150, 0)	120	/	5000
4	Reconnaissance UAV	(400, 0)	120	/	5000
5	Reconnaissance &Attack	(0, 0)	120	4	5000
6	Attack UAV	(0, 300)	120	6	5000
7	Attack UAV	(300, 0)	120	6	5000

Table 3 **Information of target and base location**

The number of targets	Target position (km,km)	The number of targets	Target position (km,km)
1	(600 , 600)	6	(1300, 950)
2	(350 , 900)	7	(1400, 650)
3	(520, 1250)	8	(900 , 300)
4	(800, 1400)	9	(1250, 480)
5	(1150, 1150)	10	(900 , 900)
base	(1500, 1500)		

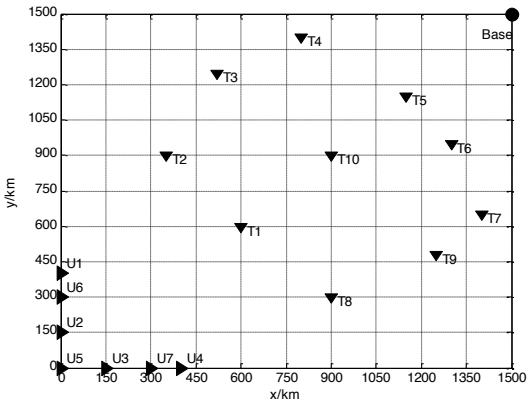

Fig. 1 **The battlefield environment**

Simulation results and analysis:
Based on the above battlefield assumption, the DE-DFA algorithm proposed in this paper is used to solve the problem, the specific parameters are configured as:

Population size $n = 50$, the maximum number of iterations are 100 times, $\beta_0 = 1$, $\alpha = 0.5$, $\omega_g = 0.4$, neighbourhood range of body variation is 5.

The results of optimal task allocation and the sequence of UAVs' tasks are shown by simulation, as shown in Table 2 and Table 3 are schematic diagram of multiple UAVs execution tasks under the optimal assignment results. Figure 2 is the Gantt chart of the assigning results.
As shown in Figure 3 and table 4, It can be seen that the DE-DFA algorithm can effectively solve the problem of multi UAV cooperative task allocation under the condition of time coupling constraints.

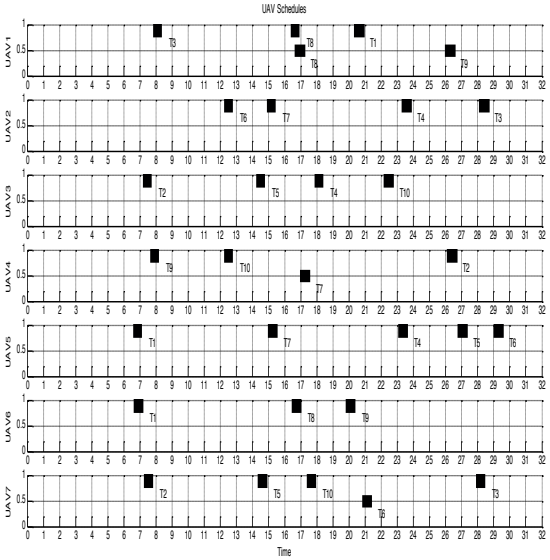

Fig.2 **Gantt chart of the assigning results**

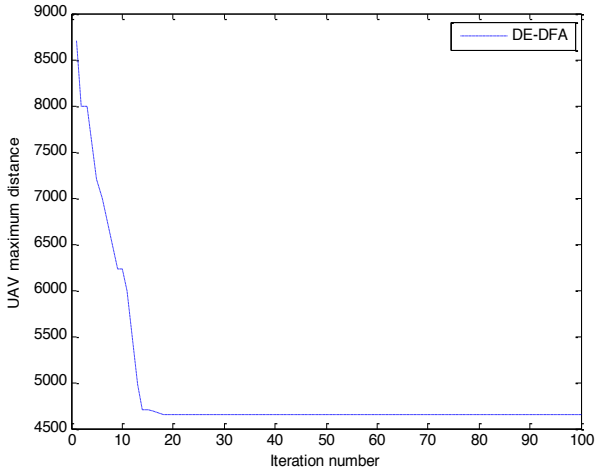

Fig. 3 **Convergence curve**

Table 4 **The results of optimal distribution**

Optimal distribution plan	[5 6 1 3 7 4 1 7 2 3 5 2 3 7 5 2 7 5 2 5 4 1 6 1 4 6 1 4 7 3 1 2 6 5 7 4 9 3 10 8 2 1 8 5 7 10 4 6 3 5 8 16 9 7 4 3 10 2]		
The task execution sequence and distance of UAVs	The number of UAVs	Task sequence	Flight distance
	1	7-22-24-3-27	4245.1km
	2	16-19-12-9	4459.9km
	3	4-13-10-30	3584.7km
	4	25-28-21-6	4506.4km
	5	1-20-11-15-18	4140.5km
	6	2-23-26	3499.4km
	7	5-14-29-17-8	4429.9km

Conclusion

This paper studies the task allocation problem of the multi-isomeric UAVs cooperative execution of SEAD, and fully considers the heterogeneous characteristic of multiple UAVs and the complex coupling constraints in SEAD, etc. also comes up with hybrid discrete firefly algorithm based on differential evolution operators and solves the task allocation problem of multiple UAVs in the coupling task environment. Through the simulation experiment, it can be seen that the algorithm can not only effectively solve the cooperative task allocation problem of multi UAV with time coupling constraints, but also can solve the problem of multi unmanned aerial vehicle cooperative task allocation with special coupling constraints and time coupling constraints.

Reference:

[1]Israel K, Nesbit R. Defense Science Board Study on Unmanned Aerial Vehicles and Uninhabited Combat Aerial Vehicles [J]. 2004.

[2]Clothier R A,Walker R A. Determination and Evaluation of UAV Safety Objectives [M]// The Future Internet. Springer Berlin Heidelberg, 2014:193-207.

[3]Long T.Research on distributed tasks allocation and coordination for multiple UCAVs cooperative mission control[D].Changsha: National University of Defense Technology, 2006
[4]Ye Y Y.Research on mission planning for cooperative UCAVs[D].Changsha:College of Mechatronic Engineering and Automation, National University of Defense Technology, 2005
[5]Alighanbari M, Kuwata Y, How J P. Coordination And Control Of Multiple Uavs With Timing Constraints And Loitering[C] //American Control Conference, 2003. Proceedings of the 2003. 2003:5311 - 5316.
[6]Jin Y, Minai A A, Polycarpou M M. Cooperative real-time search and task allocation in UAV teams[C]// IEEE Conference on Decision & Control. 2004:7-12 Vol.1.
[7]Hu J, Xie L, Lum K Y, et al. Multiagent Information Fusion and Cooperative Control in Target Search[J]. IEEE Transactions on Control Systems Technology, 2013, 21(4):1223-1235.
[8]Mclain T W, Beard R W. Coordination Variables, Coordination Functions, and Cooperative- Timing Missions[J]. Journal of Guidance Control & Dynamics, 2005, 28(1) :págs. 150-161.
[9]Choi H L, Whitten A K, How J P. Decentralized task allocation for heterogeneous teams with cooperation constraints[C]// Proceedings of the American Control Conference 2010:3057-3062.
[10]Yang X S, Sadat Hosseini S S, Gandomi A H. Firefly Algorithm for solving non-convex economic dispatch problems with valve loading effect [J]. Applied Soft Computing, 2012, 12(3):1180-1186.
[11] Dong J.Study on firefly algorithm and its application in path planning of underwater vehicles[D]. Harbin : Harbin Engineering University, 2013.
[12] SUN B,JI W,DU S,*et al.*Based on improved firefly algorithm research for multi-resource and multi-objective job-shop scheduling [J] .Modern Manufacturing Engineering, 2016(2):65-72.

Cloud forensics relationship between the Law Enforcement and Cloud Service Providers

Younis Al-Husaini, Matthew Warren and Lei Pan

Deakin University Centre for Cyber Security Research and Innovation, Deakin University, Geelong, Victoria, Australia.
yalhusai@deakin.edu.au; matthew.warren@deakin.edu.au; l.pan@deakin.edu.au.

Abstract. Cloud computing is one of the most important advances in computing in recent history. In contrast, cybercrime has developed side by side and rapidly in recent years, where the battle has exposed the exploitation of cloud computing by terrorist groups as a technique for fraud, stealing money and information, leaking secret documents, hacking government websites, recruiting new members and other activities. They are taking advantage of the gap between cloud service providers (CSPs) and law enforcement (LEAs), where LEAs cannot work without the cooperation of CSPs since their relationship is not one challenge that can be addressed, indeed should be on the legal, organisational and technical level which effect on the cooperation among them. Therefore, it is essential to enhance the Cloud forensics relationship between LEAs and CSPs. This research addresses the need for a unified collaborative model to facilitate proper investigations and explore and evaluate existing different models involved in the relationship between LEAs and CSPs as a participant in investigations. Moreover, it investigates how their relationship affects the path of the real-world forensic cases.

Keywords: Cloud Forensics, Incident Response, Law Enforcement Agents, Cloud Service Providers, Cloud Forensics Readiness.

1 INTRODUCTION

Cloud computing has become a revolution in the technology world, and has resulted in a massive expansion at the individual and institutional levels, because of the fact that communities are becoming more dependent on cloud computing services that are necessary to replace old systems to save time and cost, Gartner expects cloud services to exceed $ 300 billion by 2021 [1]. Governments are now turning to the application of cloud computing for the delivery of their systems and services. Due to this emerging technology, most of the current digital crime forensic cases have moved from local device storage to the cloud including smart devices connected to cloud environments. The result of this change is seen in the daunting challenges that Law Enforcement faces when conducting investigations involving data stored in the cloud. Hence, Digital Forensics (DF) had a new specialization named as "Cloud forensics (CF)".

The complex, dynamic and highly interconnected infrastructure in the cloud computing has reconsidered of traditional digital forensics introducing many difficulties and challenges for all stakeholders. This has affected the work of LEAs authorities globally due the required the collaboration from CSPs, and new processes needed to address and cope with new technologies used in the cyber. This paper aims to report and discuss the different existing models involved in Cloud Forensics Readiness (CFR), and the relationship between CAPs and LEAs Have been enhanced or studied by previous researches or not. In the remaining of this paper, Section 2 covers a literature review of cloud forensics and challenges. In Section 3 we introduce and discuss existing CFR models and relevant standards and finally, Section 4 shares additional conclusions and recommendations for future work.

2. DIGITAL FORENSICS

The concept of digital forensics is relatively new compared to other branches of forensic science that can be dated back to the early 1920s [2, 3]. The initial work focused on “the establishment of the popular practice of using the comparison microscope for bullet comparison in the 1920s” [4]. Since the emergence of electronic crimes, the notion appeared and evolved in parallel to various emerging crimes performed by cyber offenders around the globe. The impact of this problem has intensified with the accelerated development of information technology. Therefore, "digital forensics emerged in response to the growth of crimes committed by the use of computer systems either as an object of the crime, a tool used to perpetrate a crime or a repository of evidence related to a crime” [5], In 1984, the Federal Bureau of Investigation (FBI) laboratory started to develop programs dedicated to the examination of computer evidence [6]. The prime purpose of Digital Forensics is to facilitate the reconstruction of events and actions which are found to be criminal or helping to anticipate any malicious actions shown to be troublesome to planned operations. Therefore, the credibility of digital evidence is at the core of the digital forensic process because it is the means by which a forensic conclusion is either accepted or rejected [7]. Despite many assertions by researchers about the importance of an international standard for digital forensics [8-10], there is no international digital forensics standard to unify the process.

3. CLOUD COMPUTING

Cloud computing cannot be considered a new technological term, but just in 2007 only cloud computing was introduced to the public, after the announcement of Google and IBM cooperation in cloud technologies [11, 12]. Gartner predicts cloud services around the world to grow in 2018 to more than $ 186 billion, an increase of 21.4 per cent from 2017, as it is expected to exceed $ 300 billion by 2021 [1]. One of the first government initiatives in cloud computing was the initiative of the US government "Cloud First", which states that the federal agencies should consider the solutions cloud [13].

Cloud computing has been defended by NIST as “Model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction” [14]. Cloud computing services models are divided into three groups, (SaaS) Software as a Service, (PaaS) Platform as a Service, and (IaaS) Infrastructure as a Service [14-18]. Cloud computing deployment models have been classified into four types: Public Cloud Computing, Community Cloud Computing, Private Cloud Computing, Hybrid Cloud Computing [14].

4 CLOUD FORENSICS

Previous researchers have published numerous papers on cloud forensic with great interest in the cloud computing. However, despite all this research, no solutions were found to address the cloud forensics challenges [19]. Many researchers argue that the challenges of cloud forensic cannot be solved through technology alone, because there are regulatory and legal principles that must be solved side by side [20]. Numerous challenges that need to be solved in all these areas, many academic, technical, regulatory and legal researchers have begun to discuss these challenges.

The term of the cloud forensics is relatively new. It was the first researcher presented this term in 2011 [21] [11], who introduced organisational, technical and legal cloud forensics challenges. NIST has defended cloud computing as "Cloud computing forensics science is the application of scientific principles, technological practices and derived and proven methods to reconstruct past cloud computing events through identification, collection, preservation, examination, interpretation and reporting of digital evidence" [22].

4.1 Previous studies

Cloud forensics readiness has been well studied by many researchers [11, 23-34]. A theoretical framework was proposed with some suggested technical solution which will be reviewed below: The NIST Cloud Computing Forensic Science Working Group (NCC FSWG)[22] is one of the most critical reports that highlight the challenges and problems of cloud forensics. However, the report is relatively outdated, as the NCC FSWG has not issued any other report in this area, perhaps the lack of financial support for the group. Yet it is the most detailed and comprehensive comparison to other papers so far.

Another excellent report related to cloud forensic was published in 2016 by European Network and Information Security Agency (ENISA), which came under the title "Exploring Cloud Incidents" It provides an overview of the current state of cloud forensic and the incident response, and identify and analyse the current technical, legislative, organisational challenges

[34]. One of the CFR models is of interest by implementing a Botnet solution for monitoring the cloud environment and providing acceptable digital evidence that can be used within the courts, which is proposed by [30], yet this model needed to be standardised to integrate with other cloud processes. Another forensics readiness model which could be used by CSPs as a technique for DFR, which can help CSPs to control evidence is required for investigations. Although, the range of this framework is restricted to data examination in forensic analysis within the cloud infrastructure [25]. Remote access / central to the investigator to the cloud computing it was one of the proposed models for CFR by [31], which could support digital forensics investigators to do their work. A conceptual model for immigrating the organisations to the cloud environment suggested by [26], the idea to define the status of readiness of CSPs. The proposed model which involves a process tool, allows organisations to create the right decision and choose the suitable CSP. Highlighting the requirements of cloud forensics, the authors used a non Malicious Botnet to measure forensic readiness. Those suggested requirements include technical, operational and legal aspects based on the [32] standard. Once again the requirements need to be ascertained and tested to ensure the performance is not only effective but can be a standard to which all forensics investigations are conducted and will continue to evolve with future technologies [35]. A conceptual framework which is detected to help IaaS users activate forensics readiness. The framework shows how IaaS users can get the potential digital evidence without depending on CSPs. This model includes nine elements, containing the technical, legal and organisational forensics readiness core values [27].

A forensics-by-design framework for Cyber-Physical Cloud Systems (CPCS) suggested by [28], where they are highlighting the significance of forensic readiness. It contains six elements and assures us that a CPCS can be intended for facilitating forensics investigations. This framework can help investigators and accelerate the forensics investigations. CFR framework by highlighting the factors affecting CFR technical, legal and regulatory [11].

According to the previous literature review [11, 23-34], there is a consensus of researchers on the significance of cloud forensics for all the stakeholders, and they proposed a range of forensics readiness framework and situations, Table 1 shows the differences between those models.

Author	Year	Cited by	Framework	Solution	Trustworthiness	Segregation of Duties	SOPs & Standards	Law and regulations	Readiness	Accountability
Grobler	2010	41	X		M		D	D	W	
Elyas	2014	14	X		M		D	D	W	
Sibiya	2013	10		X					D	
Makutsoane	2014	4	X		M		D	M	P	D
Kebande	2016	6		X				D	D	
Moussa	2014	2	X		D		D	D	W	
Ab Rahman	2016	49	X				D	D	W	
Valjarevic	2013	6	X				W	M	W	
Kebande	2014	20		X			D	D	D	
Trenwith	2013	17		X			D		D	
ISO/27043	2015	*	X				W	P	W	
NIST	2014	*	X		W	P	D	D	M	M
ENISA	2016	*	X		D	M	W	W	D	D
Alenezi	2017	1	X		D		D	P	W	

	Proposed DF in Orgnizations	F	100% Fully covered
	Proposed CFR with CSPs	W	75% Well covered
	Covered Challenges: LEAs & CSPs	P	50% Partially covered
	Covered Relationship: LEAs&CSPs	D	25% Was discussed
*	International Organization	M	5% Only Mentioned

Table 1**: Existing Cloud Forensics Previous Works**

Table 1 presents a set of models proposed by the researchers [11, 23-34], whether the model is just a framework or recommended solutions, as well as the year of publication and the number of citations, and finally the six factors (trustworthiness, segregation of duties, standards, Laws and regulations, readiness, accountability). To further clarify the area covered by the research, the table distribute to four colors, where the lead colour indicates that the researcher covered the readiness of digital forensic in the organizations, and symbolizes light blue, which occupied the majority of the researcher's interests to include the forensic cloud readiness for CSPs, and the blue symbolizes for the covered the cloud forensic challenges for LEAs and CSPs, and in the last green colour and symbolizes to researchers who covered the relationship between LEAs and CSPs, where the researcher could not reach any privies study in this regard, which means there is a lack of studies in this area.

The researcher also assesses the level of coverage of each of the factors separately, Where the letter (F) means 100% Fully covered, (W) 75% Well covered, (P) 50% Partially covered, (D) 25% Was discussed, (M) 5% Only Mentioned. This assesses based on the researcher's evaluation of the previous studies; the assessment was based on the extent of its coverage of the factors.

Although some researchers have proposed technical solutions [25, 30, 31, 35] that may help the CSPs as a first responder, these suggested solutions need to be standardised and integrated with other factors. In general, there is a knowledge gap in the relationship between CSPs and LEAs on the legal, organisational and technical level to enhance trustworthiness and cooperation among them.

4.2 Why we Need to study LEAs & CSPs Relationships

There is a broad consensus among researchers [14, 16, 21, 34, 36-38] that the CSPs should cooperate with LEAs to make successful forensic investigations into cloud computing, as all powers are with the CSPs, and LEAs cannot operate without the cooperation of the CSPs. Although there is consensus on the importance of CSPs cooperation with LEAs, there is still no research study to highlight this relationship and identify the factors that need to be understood in order to enhance and streamline the relationship. Here we discuss some factors and their impact on the relationship between CSPs and LEAs:

Trustworthiness
Trustworthiness one of the significant challenges for all cloud stakeholders, especially on CSPs as a first responder and during any cloud event. The CSPs has different concerns and priorities from the LEAs [33] [16]. It is likely that its priority will be to restore the service rather than to evidence integrity. Also, the CSPs is likely to start the investigation procedures without taking appropriate precautions to ensure the integrity of potential evidence [39-44].

Segregation of Duties
The segregation of duties should be in place across all stages before, during and after the incident, starting from the readiness in both sides LEAs and CSPs, and assign the role and tasks For all cloud partners (CSPs, consumer, LEAs, judiciary) specially during cloud forensic investigations to avoid the conflicts of interest.

SOPs & Standards
Although there are a variety of different proposed models, there are no acceptable standards on how to govern the cloud forensics. Various organisations do their own model/SOPs, which were based on personal experience which manages the assessments and validation of cloud forensics, software and policies without any standards. Despite the existence of many studies and academic research, which propose frameworks and solutions to the CFR [11, 23, 24, 26-34], there are no international standards of cloud forensics.

Law and regulations
Laws and regulations must govern evidence obtaining processes from the cloud environment and the relationship between stakeholders and LEAs. The lack of response from CSPs, despite the existence of a search warrant, is one of the problems experienced by LEAs. The legal,

regulatory and most forensic researchers have confirmed this factor [11, 23, 24, 26-30, 32-34]. The sub areas are:

Jurisdiction

Despite the proliferation of connected devices around the world through the Internet, the web has made the world small and unbounded place, but law enforcement agencies are suffering from the problem of international jurisdiction in the virtual world, often the results are illegal, or no cooperation from international CSPs due to international jurisdiction [21, 45-49]

Lack of international agreements & laws

Worth mentioning, there are laws and treaties such as Budapest in 2001[50], which aim to unify the international efforts to combat digital crimes, which included many definitions of criminal acts, leaving each state to determine the punishment it deems appropriate. The Convention also specified a particular clause on the need for cooperation between members, whether at the level of evidence collection or extradition. Despite all of that, still, there is a lack of cooperation from CSPs located within different geographical boundaries (International CSPs), this is due to the differences in the laws and regulations in force in the two countries. [21, 45, 51]. To keep up-to-date of the dramatic development of cloud forensics, the Budapest Convention Council in 2014 established the T-CY Cloud Evidence Group (CEG), which aims to "explore solutions to access evidence in the cloud for criminal justice, including through mutual legal assistance". CEG summarised the problems and recommended solutions in its final report on 16 September 2016; this included the preparation of the second additional Protocol to the Budapest Convention. On June 8, 2017, the "Terms of Reference for the Draft Additional Protocol II to the Budapest Convention on Cybercrime" were approved. Which aims to enhance direct cooperation with CSPs in other jurisdictions about applications for subscriber information, conservation requests and emergency requests [52]. The project is expected to be completed at least two and a half years [53].

Readiness

All these various difficulties in the cloud forensics have prompted many organisations and governments to try to be forensically ready to conduct a forensic investigation in the cloud. Cloud forensics Readiness (CFR) has been defined as “A mechanism aimed at reducing the cost of carrying out an investigation in a cloud environment by providing any relevant information needed before setting up the investigation.” [11]

CFR is one of the main things that should be implemented and readily available in organisations to mitigate the challenges in the cloud forensics investigation and the gap between LEAs and CSPs, and all the privies studies confirmed that [11, 23-34].

Accountability

Accountability is the responsibility of individual’s actions towards cloud forensics. Accountability is one of the most effective elements for cloud forensics toward the governing the relationship between all the

stakeholders and LEAs. Table 1 illustrates the lack of the previous studies which highlight this factor, where was mentioned only by [26, 33, 34].

5 CONCLUSION AND FUTURE WORK

This research seeks to study the challenges of cloud forensics and to highlight the relationship between CSPs and LEAs. A wide range of existing research in CFR has a framework and prototype for solutions was quantified. However, as previous literature has shown, there is a knowledge gap in the relationship between the CSPs and LEAs in cloud forensics from a legal and technical perspective, the research highlights the importance of this relationship in order to achieve a high level of readiness, segregation of duties and responsibility.

To achieve the objectives, the researcher will use an interpretive qualitative case study approach, and the relationship between the CSPs and LEAs will be investigated. This research is expected to contribute to the body of knowledge of cloud forensics as the development of new theory, and a practical contribution is expected, where new insights and vision to decision and policy makers, CSPs and LEAs to improve this relationship.

REFERENCES

[1] Gartner. (2018). *Gartner Forecasts Worldwide Public Cloud Revenue to Grow 21.4 Percent in 2018*. Available: https://www.gartner.com/newsroom/id/3871416

[2] V. Baryamureeba and F. Tushabe, "The enhanced digital investigation process model," in *Proceedings of the Fourth Digital Forensic Research Workshop*, 2004, pp. 1-9.

[3] A. Agarwal, M. Gupta, S. Gupta, and S. C. Gupta, "Systematic digital forensic investigation model," *International Journal of Computer Science and Security (IJCSS),* vol. 5, no. 1, pp. 118-131, 2011.

[4] Troopers. (2018). *Crime Laboratory System: Forensic Science History*. Available: https://www.troopers.ny.gov/Crime_Laboratory_System/History/Forensic_Science_History/

[5] S. Ballou, *Electronic crime scene investigation: A guide for first responders*. Diane Publishing, 2010.

[6] W. G. Kruse II and J. G. Heiser, *Computer forensics: incident response essentials*. Pearson Education, 2001.

[7] S. R. Selamat, R. Yusof, and S. Sahib, "Mapping process of digital forensic investigation framework," *International Journal of Computer Science and Network Security,* vol. 8, no. 10, pp. 163-169, 2008.

[8] X. Du, N.-A. Le-Khac, and M. Scanlon, "Evaluation of Digital Forensic Process Models with Respect to Digital Forensics as a Service," *arXiv preprint arXiv:1708.01730,* 2017.

[9] M. D. Kohn, M. M. Eloff, and J. H. Eloff, "Integrated digital forensic process model," *Computers & Security,* vol. 38, pp. 103-115, 2013.

[10] M. Reith, C. Carr, and G. Gunsch, "An examination of digital forensic models," *International Journal of Digital Evidence,* vol. 1, no. 3, pp. 1-12, 2002.

[11] A. Alenezi, R. K. Hussein, R. J. Walters, and G. B. Wills, "A Framework for Cloud Forensic Readiness in Organizations," *2017 5th IEEE International Conference on Mobile Cloud Computing, Services, and Engineering (MobileCloud),* pp. 199-204, 2017.
[12] M. A Vouk, "Cloud computing–issues, research and implementations," *Journal of computing and information technology,* vol. 16, no. 4, pp. 235-246, 2008.
[13] V. Kundra, "Federal cloud computing strategy," 2011.
[14] P. Mell and T. Grance, "The NIST definition of cloud computing," 2011.
[15] E. E. D. Hemdan and D. H. Manjaiah, "A cloud forensic strategy for investigation of cybercrime," in *2016 International Conference on Emerging Technological Trends (ICETT)*, 2016, pp. 1-5.
[16] S. Alqahtany, N. Clarke, S. Furnell, and C. Reich, "A forensic acquisition based upon a cluster analysis of non-volatile memory in IaaS," in *Anti-Cyber Crimes (ICACC), 2017 2nd International Conference on*, 2017, pp. 123-128: IEEE.
[17] S. Hraiz, "Challenges of digital forensic investigation in cloud computing," in *2017 8th International Conference on Information Technology (ICIT)*, 2017, pp. 568-571.
[18] B. K. Raju, G. Meera, and G. Geethakumari, "Cloud forensic investigation: A sneak-peek into acquisition," in *2015 International Conference on Computing and Network Communications (CoCoNet)*, 2015, pp. 348-352.
[19] Y. Zhao and B. Duncan, "Could Block Chain Technology Help Resolve the Cloud Forensic Problem?," *CLOUD COMPUTING 2018,* p. 49, 2018.
[20] L. De Marco, "Forensic Readiness Capability for Cloud Computing," 10187642 Ph.D., University College Dublin (Ireland), Ann Arbor, 2015.
[21] K. Ruan, J. Carthy, T. Kechadi, and M. Crosbie, "Cloud forensics," in *IFIP International Conference on Digital Forensics*, 2011, pp. 35-46: Springer.
[22] P. Mell and T. Grance, "Nist cloud computing forensic science challenges," *Draft Nistir,* vol. 8006, 2014.
[23] C. Grobler, C. Louwrens, and S. H. von Solms, "A framework to guide the implementation of proactive digital forensics in organisations," in *Availability, Reliability, and Security, 2010. ARES'10 International Conference on*, 2010, pp. 677-682: IEEE.
[24] M. Elyas, S. B. Maynard, A. Ahmad, and A. Lonie, "Towards a systemic framework for digital forensic readiness," *Journal of Computer Information Systems,* vol. 54, no. 3, pp. 97-105, 2014.
[25] G. Sibiya, T. Fogwill, H. S. Venter, and S. Ngobeni, "Digital forensic readiness in a cloud environment," in *AFRICON, 2013*, 2013, pp. 1-5: IEEE.
[26] M. P. Makutsoane and A. Leonard, "A conceptual framework to determine the digital forensic readiness of a Cloud Service Provider," in *Management of Engineering & Technology (PICMET), 2014 Portland International Conference on*, 2014, pp. 3313-3321: IEEE.
[27] A. N. Moussa, N. B. Ithnin, and O. A. Miaikil, "Conceptual forensic readiness framework for infrastructure as a service consumers," in *Systems, Process and Control (ICSPC), 2014 IEEE Conference on*, 2014, pp. 162-167: IEEE.
[28] N. H. Ab Rahman, W. B. Glisson, Y. Yang, and K.-K. R. Choo, "Forensic-by-design framework for cyber-physical cloud systems," *IEEE Cloud Computing,* vol. 3, no. 1, pp. 50-59, 2016.
[29] A. Valjarevic and H. Venter, "A Harmonized Process Model for Digital Forensic Investigation Readiness," in *IFIP International Conference on Digital Forensics*, 2013, pp. 67-82: Springer.

[30] V. R. Kebande and H. S. Venter, "A cloud forensic readiness model using a Botnet as a Service," in *The International Conference on Digital Security and Forensics (DigitalSec2014)*, 2014, pp. 23-32: The Society of Digital Information and Wireless Communication.
[31] P. M. Trenwith and H. S. Venter, "Digital forensic readiness in the cloud," in *Information Security for South Africa, 2013*, 2013, pp. 1-5: IEEE.
[32] ISO/27043, "Information technology - Security techniques - Incident investigation principles and processes," *ISO/IEC,* 2015.
[33] NIST, "Cloud Computing Forensic Science Challenges," no. Cloud Computing Forensic Science Challen, 2014.
[34] D. Liveri and C. Skouloudi, "Exploring Cloud Incidents," *The European Network and Information Security Agency (ENISA),* pp. 1-14, 2016.
[35] V. R. Kebande and H. Venter, "Requirements for achieving digital forensic readiness in the cloud environment using an NMB solution," in *11th International Conference on Cyber Warfare and Security: ICCWS*, 2016, p. 399.
[36] S. A. Almulla, Y. Iraqi, and A. Jones, "A state-of-the-art review of cloud forensics," *Journal of Digital Forensics, Security and Law,* vol. 9, no. 4, pp. 7-28, 2014.
[37] S. Zhang, L. Wang, and X. Han, "A KVM virtual machine memory forensics method based on VMCS," in *Computational Intelligence and Security (CIS), 2014 Tenth International Conference on*, 2014, pp. 657-661: IEEE.
[38] G. Grispos, T. Storer, and W. B. Glisson, "Calm before the storm: the challenges of cloud," *Emerging digital forensics applications for crime detection, prevention, and security,* vol. 4, no. 1, pp. 28-48, 2013.
[39] M. E. Alex and R. Kishore, "Forensics framework for cloud computing," *Computers & Electrical Engineering,* vol. 60, pp. 193-205, 2017.
[40] B. Grobauer and T. Schreck, "Towards incident handling in the cloud: challenges and approaches," in *Proceedings of the 2010 ACM workshop on Cloud computing security workshop*, 2010, pp. 77-86: ACM.
[41] Y. Chen, V. Paxson, and R. H. Katz, "What's new about cloud computing security," *University of California, Berkeley Report No. UCB/EECS-2010-5 January,* vol. 20, no. 2010, pp. 2010-5, 2010.
[42] M. Creeger, "Moving to the edge: a CTO roundtable on network virtualization," *Communications of the ACM,* vol. 53, no. 8, pp. 55-62, 2010.
[43] N. Convery and K. Ferguson-Boucher, "Storing information in the cloud," *Bulletin of Information and Records Management Society, Issue,* pp. 3-5, 2010.
[44] K. K. R. Choo, "Cloud computing: challenges and future directions," 2010.
[45] K. Ruan, J. Carthy, T. Kechadi, and I. Baggili, "Cloud forensics definitions and critical criteria for cloud forensic capability: An overview of survey results," *Digital Investigation,* vol. 10, no. 1, pp. 34-43, 2013.
[46] K. Ruan and J. Carthy, "Cloud computing reference architecture and its forensic implications: A preliminary analysis," in *International Conference on Digital Forensics and Cyber Crime*, 2012, pp. 1-21: Springer.
[47] K. Ruan, J. James, J. Carthy, and T. Kechadi, "Key terms for service level agreements to support cloud forensics," in *IFIP International Conference on Digital Forensics*, 2012, pp. 201-212: Springer.
[48] R. Adams, "The emergence of cloud storage and the need for a new digital forensic process model," *Cybercrime and cloud forensics: Applications for investigation processes,* pp. 79-104, 2012.

[49] J. I. James, A. F. Shosha, and P. Gladyshev, "Digital forensic investigation and cloud computing," in *Cybercrime and cloud forensics: Applications for investigation processes*: IGI Global, 2013, pp. 1-41.
[50] O. E. COUNCIL, "Convention on Cybercrime," *Budapest, November,* vol. 23, 2001.
[51] I. Orton, A. Alva, and B. Endicott-Popovsky, "Legal process and requirements for cloud forensic investigations," *Cybercrime and Cloud Forensics: Applications for Investigation Processes,* pp. 186-229, 2012.
[52] CCDCOE, "Council of Europe Ponders a New Treaty on Cloud Evidence," 2017-06-28 2017.
[53] CouncilofEurope. (2017). *Cybercrime: towards a Protocol on evidence in the cloud*. Available: https://www.coe.int/en/web/human-rights-rule-of-law/news/-/asset_publisher/qI77sHb3q28L/content/cybercrime-towards-a-protocol-on-evidence-in-the-clo-1

Understanding the influence of Individual's Self-efficacy for Information Systems Security Innovation Adoption: A Systematic Literature Review

Mumtaz Abdul Hameed
Technovation Consulting and Training Private, Limited 33, Chandhani Magu
Male'. Maldives
Email: mumtazabdulhameed@gmail.com

Nalin Asanka Gamagedara Arachchilage
School of Engineering and Information Technology
University of New South Wales, The Australian Defence Force Academy
Australia
Email: nalin.asanka@adfa.edu.au

Abstract

Information Systems security cannot be fully apprehended if the user lacks the required knowledge and skills to effectively apply the safeguard measures. Knowledge and skills enhance one's self-efficacy. Individual self-efficacy is an important element in ensuring Information Systems safeguard effectiveness. In this research, we explore the role of individual's self-efficacy for Information Systems security adoption. The study uses the method of Systematic Literature Review using 42 extant studies to evaluate individual self- efficacy for Information Systems security innovation adoption. The systematic review findings reveal the appropriateness of the existing empirical investigations on the individual self-efficacy for Information Systems security adoption. Furthermore, the review results confirmed the significance of the relationship between individual self-efficacy and Information Systems security adoption. In addition, the study validates the past administration of the research on this subject in terms of sample size, sample subject and theoretical grounds.

Keywords: Innovation Adoption Process; Information System Security; IS Security Adoption; Self-Efficacy; User Acceptance of Innovation

1. Introduction

Information Systems (IS) assets (information and computer resources) are at risk from a variety of threats, including virus, worms, Trojans, spyware, scare-ware, crime-ware, key-loggers, botnet, DDoS, browser-hijackers, pharming, phishing etc. [8]. Such attacks commonly referred to as 'IS security threats' mainly intended to improperly disclose, modify or delete sensitive information and maliciously destruct and destroy computer resources [23]. New prospect the internet has presented to the

users have in fact, offered criminals and individuals with a vicious mind-set to misuse IS assets aimlessly.

To thwart IS security threats and safeguard organisational IS assets in general, a combination of measures is taken such as the installation of anti-virus, anti-spyware and anti-phishing software, setting up firewalls, maintaining and restricting access controls, using intrusion detection and prevention systems and by putting in encryption and content filtering software [33, 38, 49]. These measures offer a technological or technical solution to the problem, but by no means reasonable to efficiently safeguard IS security threats completely [3, 30, 49, 50, 56, 73, 74]. So as to survive with increased threats and to effectively protect IS assets, non-technical solutions such IS security policies have likewise been employed [53]. Research has established the view that organisations and individuals who opt for technical as well as non-technical measures to protect their IS assets are more likely to attain success in safeguarding IS resources [47, 56, 65]. In IS literature an innovation is referred as an idea, a product, a process or a technology that is new to an individual or organisation [25, 27]. Hence, technical and non- technical IS security measures may collectively be referred as IS security innovations.

Although both technical and non-technical IS security measures are important, several research had pinpointed behaviour of individual user within an organization as one element of ambiguity in securing IS assets [5, 16, 56, 65]. With all the technical and non- technical IS security measures at one's disposal, efficient use cannot be realized if the end user lacks the required knowledge and skills to adequately apply the measures. If the end-users of organisational IS does not understand the importance of IS security practices and are not eager to accept the policies, then those safeguards measures become ineffective [30]. Given that the security attacks are increasingly widespread and more organized than ever, it is important to gauge the knowledge of users to detect and prevent such attacks.

When an individual possesses the necessary knowledge about the effectiveness of a safeguard measure in providing protection from IS threats, that individual is more likely to adopt preventive behaviour or action [38, 51, 68]. Chan et al. [10] stated that acquisition of knowledge related to an IS countermeasure builds one's self-confidence in dealing with threats. According to IS literature, computer self- efficacy is the term that relates individual's self- confidence and ability to successfully use a computer or IS to accomplish a specific task [4, 13]. Computer self-efficacy have also been cited as essential in determining one's intention to engage in current or future use of an IS.

Prior research on IS indicates a significant positive relationship between individual's IS self-efficacy and the usage of ISs [60]. Also, individuals IS self-efficacy has found to be a significant determinant for IS security adoption [30, 53, 66]. Eastin and LaRose [20] state that self-efficacy overcomes the fear many novice users experience in an event of threat and enhances the ability to cope with any attack. Arachchilage and Love [4], identified self-efficacy as an important determinant of the IS

security threat avoidance behaviour and a key element in ensuring safeguard effectiveness.

This research attempts to examine the role of an individual's self-efficacy in IS security innovation adoption. To this end, the study reviewed past literature on IS security to establish the relationship between self-efficacy and IS security adoption. The research makes three main contributions to theory and practice. First, using a review of IS security literature, the research verifies the significance of examining the effect of individual self-efficacy on IS security adoption. Secondly, the analysis carried out established the existing savvy of the role of individual self-efficacy for IS security innovation adoption. Finally, the study approves the significance of individual self-efficacy for IS security innovation adoption.

The remainder of this paper is organised as follows. The 'Theoretical Background' section illustrates the basics of self-efficacy relating to IS security. In the subsequent section 'Research Questions', we presented 4 research questions for the study. The 'Research Methodology' section, briefly discusses the method employed to examine the influence of the relationship between self-efficacy and IS security innovation adoption. In Section 5, we presented the result obtained from the data analysis. Finally, in Section 6, we discussed the finding of the study results, in addition, conclusion was also presented in Section 6.

2. Theoretical Background

The focus of IS security is to protect and safeguard organization's IS assets from vulnerabilities [1]. The main challenge for organization's IS security is to protect unauthorized access of information sources [21] and to defend computer resources against malicious attacks. As a result, organizations allocate a substantial amount of resources to safeguard their IS assets from IS security threats [23].

Various solutions have been developed in response to IS security and these solutions targets both technical and non-technical problem areas [3, 5]. With all the IS security measures at one's disposal, the efficient use cannot be realised if the end user lacks the required knowledge and skills to adequately apply the measures. Banu and Banu [8] indicated that IS security attacks over the internet are successful because of many inexperienced and unsophisticated users. Additionally, social engineering attacks are now much more concealed as such naive users are more inclined to incautiously divulge passwords and other sensitive and classified information. Lack of awareness of the users regarding the maliciousness of crimes over the internet in effect has opened a fertile ground for cyber-criminals to conduct IS security attacks. Even in the present-day, a number of users are ignorant that their personal information is actively being targeted by cyber- criminals. Given that the security attacks are increasingly widespread and more organized than ever, it is important to develop the knowledge of users to detect and prevent such attacks.

According to Rogers [51], when individuals possess the requisite knowledge about the effectiveness of mechanisms that provide protection from a threat, they are more likely to adopt that measure. In other words, a person who is knowledgeable about IS security is more likely to assess IS security risks and accordingly employs security innovations effectively to address those risks [41]. Individual's knowledge has a co- relation to one's self-efficacy to perform a behaviour [3].

Bandura [7] defined self-efficacy as the judgment of one's ability to organize and execute given types of performance. Hence, in the context of this research, self- efficacy is referred as a belief in one's ability to thwart IS security threats and one's capability to safeguard IS assets from IS security attacks. Tamjidyamcholo et al. [59] noted that a high level of self-efficacy in a person will make them much more self-assured about their abilities and strengthens their motivation. Hence, when users are knowledgeable about IS security threats, they have more self-confidence to take relevant actions to thwart attack by adopting preventive behaviour.

Researchers often utilised Bandura [7]'s theory of self- efficacy to measure individual's self-confidence. The fundamental of this theory is in understanding the relationship between one's belief and one's willingness to engage in behaviours necessary to successfully accomplish a task. The theory also explains the process an individual experience as he or she encounters a new challenge together with the judgments, evaluations, and appraisals made based on the knowledge learnt [6].

3. Research Questions

This paper considered the existing IS security literature to determine the importance of individual self-efficacy for IS security innovation adoption. The analysis focused specifically on investigating, the following research questions:

RQ1: What are the demographics of the extant studies of individual self-efficacy on IS security innovation adoption including the year of study, sample groups, sample size, countries?

RQ2: What are the theoretical foundation used in the existing studies of individual self-efficacy on IS security innovation adoption?

RQ3: Is there a difference in investigating individual self-efficacy for different types of security innovations?

RQ4: What are the results of the studies that examine the relationship between individual self-efficacy and IS security innovation?

4. Research Methodology

A finding of an individual study is not sufficient to generalise on a particular issue, while to reach an overall outcome, findings of a number of independent studies on a subject can be combined [24]. A technique known as a Systematic Literature Review (SLR) may be used to identify, analyse and interpret all available evidence related to a specific research question [27]. To meet our research objectives and to address the research questions, we carried out a SLR to study the role of self- efficacy for IS security innovation adoption. SLR improves the likelihood of generating a clearer, more objective answer to the research questions. As SLRs considers study design (sampling strategy and data collection methods), data and analytical methods used, the reviews are effective at gauging the robustness of evidence. The use of SLR procedure enabled the study to obtain an overall conclusion regarding the relationships between individual self-efficacy and IS security adoption.

To ensure a thorough coverage of academic articles related to IS security adoption, we conducted an extensive literature search of IS Journals using Google Scholar and multiple large-scale and reputable digital libraries and databases including Web of Science, IEEE Xplore, Science Direct (Elsevier), ACM Digital Library, Wiley Online Library, ProQuest, EBSCO, Springer LINK and Emerald Management Xtra. These sources contain ample high-quality journal articles and conference papers. The search focused only on peer- reviewed journal and conference articles.

To determine which of the articles were really relevant to the research objectives the study established, an inclusion and exclusion conditions. The study selection criteria for the SLR were: (C1) it should be an empirical study on IS security adoption, (C2) the study should examine individual self-efficacy as a dependent variable, and finally, (C3) the study examines the relationship between individual self-efficacy and IS security innovation adoption.

The initial search yielded 544 citations by following inclusion and exclusion criterion C1. To accomplish the inclusion and exclusion criterion C2, the abstracts of all 544 were manually scanned to identify if the articles examine individual self-efficacy. Number of articles identified as potentially relevant were 112. By applying inclusion and exclusion criterion C3 for these 112 articles, 39 articles with 42 studies were found eligible for the SLR. The 42 studies that meet all 3 criteria examined the effect of individual self-efficacy for the adoption of IS security innovations.

5. Results

We conducted a statistical analysis using frequencies and percentages to combine and summarize the variables collected.

5.1. Distribution of Studies by Year

Table 1 shows the literature distribution by publication year of the studies. Data from the SLR shows that self- efficacy has been considered in the IS security innovation adoption literature since 2004.

Year	No. of Studies
2004	1
2005	1
2006	0
2007	3
2008	2
2009	8
2010	6
2011	2
2012	6
2013	3
2014	3
2015	0
2016	6

Table 1**: Literature distribution by publication year.**

The academic discussion of individual self-efficacy on IS security adoption has mostly taken place during the last 12 to 14 years. Table 1 shows that the number of articles over time has increased and during this period, the topic has increasingly attracted among the scholarly researchers. The distribution of studies by publication year suggests that examining individual self-efficacy for IS security innovation adoption is an increasingly emerging discourse. Also, SLR confirms that individual self-efficacy for IS security innovation adoption is still an active IS tract, as there were 6 articles published in the year 2016.

5.2. Distribution of Sample Groups in the studies

The result of this analysis provided some clarification to RQ1.

Subject Groups	No of Studies
Individual	18
Organisation	2
Student	18
Mixed	2

None	2

Table 2: **Distribution of sample groups used in the studies.**

Table 2 illustrates the number of studies that employ different sample groups in the studies to examine individual self-efficacy for IS security innovation adoption. Results suggest that the majority of studies conducted their studies by engaging individuals by adopting convenience sampling or by using student subjects. The analysis also helped explain RQ1.

5.3. Distribution of Sample size in the studies

SLR analysed sample size of the reviewed studies to further elucidate RQ1. Among the 42 studies considered in the SLR, 40 studies utilised survey methodology. In this 40 studies, a total of 13841 participants was included, with an average sample size of 346. Table 3 showed that the study employing smallest and largest sample were 77 and 988 participants, respectively. Approximately, 67% (two third) of the studies use more than 200 participants in their assessment.

Description	**No. of.**
Studies with sample	40
Smallest sample size	77
Largest sample size	988
Sample Size 0 - 100	1
Sample Size 101 - 200	12
Sample Size 201 - 300	9
Sample Size 301 - 400	3
Sample Size 401 - 500	5
Sample Size 501 - 600	4
Sample Size 601 - 700	2
Sample Size 701 - 800	1
Sample Size 801 - 900	0
Sample Size 901 -	3

Table 3**: Distribution of sample size of the studies**.

5.4. Distribution by countries

As a final appraisal to RQ1, we analysed the moderating effect of the country of study. Table 4 visually indicates that almost half of the studies were produced in the USA. The studies covered Asia, Europe and North America with a representation of 8 different countries.

Country	No. of Studies
Canada	3
China	2
Finland	4
Malaysia	3
Singapore	2
South Korea	2
Taiwan	3

Table 4: **Distribution of country of the studies**

6. Theories Used in the Reviewed Studies

In response to RQ2, we analysed the theoretical foundation for each reviewed literature. To examine the relationship between self-efficacy and IS security innovation adoption, reviewed studies used a number of different theories. Table 5 shows the different theoretical model exploited in the reviewed studies.

Theories	No. of Studies
Protection Motivation Theory (PMT)	22
Theory of Planned Behaviour (TPB)	6
Theory of Reasoned Action (TRA)	5
Social Cognitive Theory (SCT)	5
Deterrence Theory (DT)	4
Technology Acceptance Model	3
Technology Threat Avoidance	2
Cognitive Evaluation Theory (CET)	1
Coping Theory (CT)	1
Decomposed Theory of Planned	1
Extrinsic Motivational Model (EMM)	1
Health Belief Model (HBM)	1
Instrinsic Motivation Model (IMM)	1
Rational Choice Theory (RCT)	1
Social Bond Theory (SBT)	1

Table 5: **Different theories used in the studies.**

PMT is the most widely used theory to determine the relationship between self-efficacy and IS security adoption. More than half of the reviewed studies used PMT or PMT integrated with other theories. Reviewed literature suggests that apart from PMT, the Theory of Planned Behaviour (TPB), Theory of Reasoned Action (TRA) and Social Cognitive Theory (SCT) are among the most widely used theories in examining self-efficacy on IS security innovation adoption.

7. Types of Innovation

According to the classification of Zmud [71] we defined the type of innovation as process and product. For this study, process innovation involves establishing a new system, method or policies that changes the IS security operational processes, whereas product innovation are new products introduced to enhance IS security. Different factors determine the adoption of process and product innovation and the extent to which these factors impact on the adoption process [61]. We differentiate the reviewed studies into two sets of process and product innovation and examine some demographics including sample size, sample groups for each group of the studies. Also, we examine if there is any difference in the application of theories for the studies that examine process and product innovations. Table 6 highlights the difference in study practices for process and product security innovations. The result of this analysis would address to RQ3.

Description	**Process**	**Product**
No of Studies	24	18
Total sample size	895	4887
Sample Group		
Individual	13	5
Organisation	2	0
Student	8	10
Mixed	0	2
None	1	1
Theories used		
Protection Motivation Theory	12	10
Theory of Planned Behaviour	4	2
Theory of Reasoned Action	5	0
Social Cognitive Theory (SCT)	5	0
Deterrence Theory (DT)	4	0
Technology Acceptance Model	2	1

Table 6: **Distribution of studies using different security innovations.**

Also, it is evident from the results that most of IS security process innovation studies utilises individuals as a subject, whereas, most of IS security product innovation studies employs student participants.

As for the theories used for two groups of studies, process innovation studies tend to combine PMT with the theoretical basis of either TRA, SCT or Deterrence Theory (DT) compare to studies examining product innovations.

8. Significance

The relationship between independent and dependent variables is usually evaluated in term of 'test of significance', highlighting their relationship [25, 26]. 'Test of significance' and various other 'effect sizes' such as correlation co-efficient provided by quantitative studies can be aggregated to find an overall outcome [27]. Effect size when considered in terms of significance is frequently referred as weak, moderate or strong significance [24]. Hunter et al. [32] and Hameed and Counsell [25], however, suggested that aggregation of 'test of significance' results from different studies could produce a misleading outcome. This is because, there is no rule for determining the value of the correlation that interprets as weak, moderate or strong significance.

For the study, we extracted from the reviewed studies the correlation co-efficient values of the relationship between self-efficacy and IS security innovation adoption. We interpreted the correlation co-efficient values under a single classification to obtain the test of significance for our assessment. We adopted the correlation value referred by Hameed and Counsell [24] and Hameed and Counsell [26], which categorises: a correlation value between 0 and ±0.09 as insignificant, ±0.10 and ±0.29 as weak significance, ±0.30 and ±0.49 as moderate significance, ± 0.5 and ± 0.69 as strong significance, ±0.70 and ±0.89 as the very strong significance and ±0.9 and ±1.0 near perfect. Based on the above classification we coded the correlation co- efficient of individual studies and aggregated resulting tests of significance to obtain the overall assessment of the relationship between self-efficacy and IS security innovation adoption.

Among the 42 studies considered in the SLR, 35 studies provided correlation co-efficient for the relationship between individual self-efficacy and IS security innovation adoption. Table 7 summarizes the results of an aggregated test of significance for the relationship between self-efficacy and the adoption of IS security innovation.

Significance	No. of Studies
Insignificant (0.00 to ±0.09)	3
Weak Significance (0.10 to ±0.29)	7
Moderate Significance (0.30 to ±0.49)	16

Strong Significance (0.50 to ±0.69)	7
Very Strong Significance (0.70 to ±0.89	2
Perfect (0.10 to ±1.00)	0

Table 7**: Aggregated test of significance for the studies.**

9. Discussion and Conclusion

This SLR aimed to understand the role of individual self-efficacy on IS security innovation adoption. The results highlighted that individual self-efficacy is a significant attribute of IS security innovation adoption. The SLR results of the distribution of studies by publication year suggest that researchers have started examining the effect of individual self-efficacy on IS security innovation adoption since 2004. This is the period where online social media and social networking became a mainstream concept with the launching of Facebook on February 2004. These social media emerge as a target for scams; exposing individual and organisational data at risk. More people put their personal information online, offering a huge opportunity for cyber criminals to exploit. Thus, IS security innovation adoption has speedily been under scrutiny since the rise of social media and researcher started examining individual self-efficacy as one of the key predictors for IS security innovation adoption.

Studies that examined the influence of individual self- efficacy for IS security innovation adoption has explored for different sample groups. The SLR findings showed that the research on the relationship between individual self-efficacy and IS security innovation adoption based their studies on convenience samples of both students and non-students. The findings indicate that approximately half of the reviewed literature used student subjects. Using student subjects for experimental research as a substitute for another group has been widely criticised for having little external validity and generalisability. The ethical concerns of student participation revolve mainly around the issue whether the participant serve with their own consent. Also, it has been argued that student samples are fundamentally biased in age, experience, and intellectual ability. However, the studies reviewed in the SLR provided no justification for their chosen subject sample nor did acknowledge any limitations for the use of a student sample. Hence, the effect of individual self- efficacy for IT security innovation adoption bears no significance for the difference in sample groups.

The results of SLR showed that the average sample size of the studies is approximately 350 participants. A study that has a sample size which is too small may have an unrealistic chance of yielding a useful information. Larger sample sizes have the obvious advantage of providing more data for researchers to work with and provide more accurate mean values and a smaller margin of error. Thus an appropriate determination of the sample size used in a study is a crucial step in the design of a study. The sample size used in the majority of the studies reviewed in the SLR deemed appropriate. This commends of the soundness of the selected

studies for the SLR. In addition, it provides evidence on the correctness of the results of the reviewed studies that examine the relationship between individual self-efficacy and IT security innovation adoption.

In order to identify if culture moderates the relationship between individual self-efficacy and IS security innovation adoption, we explored the distribution of country of the reviewed studies in the SLR. Deans et al. [18] states that culture influences usage of IT in different countries. In a meta-analysis of TAM, Schepers and Wetzels [52] used western and non-western as a moderating factor in the context of culture. They divide the studies conducted in Europe, North America, Australia and New Zealand as western and the rest of world as non-western. The SLR represents a diverse culture which belongs to both western and non-western groups. Hence, the SLR indicates that the overall results of existing literature that considers the influence of individual self-efficacy for IS security innovation adoption is not biased towards one particular culture.

The SLR also explored the theoretical foundation exploited in examining individual self-efficacy for IS security innovation adoption by the reviewed studies. The result of the SLR identified PMT as the principal model. In a meta-analysis study, Floyd et al. [22] described PMT as one of the most powerful explanatory theories predicting individual intentions to adopt safeguard measures. PMT is useful in analysing and exploring recommended actions or behaviours to avert the consequences of threats such as IS security attacks. Apart from PMT, SLR identified SCT, TRA, and TPB as other models utilised in examining the effect of individual self-efficacy for IS security innovation adoption. SCT [7] posits that one's confidence in their ability to perform it a behaviour successfully will produce positive valued outcomes. The main tenet in the TRA is that an individual's behavioural intention in a specific context depends on attitude toward performing the target behaviour and on subjective norm. The TRA holds that the practical impact of subjective norm on the behavioural intention is that an individual may choose to perform a specific behaviour, even though it may not be favourable to him or her to do so [64]. TPB is an extension of TRA at the same time adopt the efficacy expectancies of SCT into consideration.

In this study, we identified if there is a difference in investigating individual self-efficacy for different types of security innovations. In order to analyse, we categorised IS security innovations as product and process to access the scenario. The results show that the average sample size used for IS security process innovation studies (373 participants) is higher than the product innovation studies (271 participants). One explanation is that process innovation involves replacing the entire system or work procedure, whereas product innovation does not involve change of an entire system. Also, it is evident from the results that most of IS security process innovation studies utilises individuals as subjects, whereas most of IS security product innovation studies employ students. One probable explanation could be that process innovation such IS security policies are mostly adopted in an organisational setting for which the sample subjects would most probably be non-students.

Finally, the SLR analysed the correlation co-efficient for the relationship between individual self-efficacy and IS security adoption behaviour to aggregate the tests of significance of the reviewed studies. In terms of the percentage, 92% of the studies found self-efficacy as significant (correlation value between ±0.10 to ±1.00) attribute in IS security innovation adoption. Also, approximately 71% of the studies we considered verified the association between self-efficacy and IS security adoption as moderate significance (correlation value between ±0.30 to ±0.49) or strong significance (correlation value between ±0.50 to ±0.69). Hedges and Olkin [31], Hameed and Counsell [24] and Hameed and Counsell [26] suggested that it would be within reason for a study to consider an established relationship to exist between two variables when a majority of prior studies had found statistically significant results. Hence, results of aggregated tests of significance indicate that individual self-efficacy is an important predictor of IS security innovation adoption.

This study offers several contributions to the IS security management literature. The study contributes to the field of IS security by empirically endorsing the influence of individual self-efficacy for IS security innovation adoption. Additionally, to recognise the current understanding of the subject, we gathered almost all their existing studies that examine individual's self- efficacy for IS security innovation adoption.

The most important theoretical implication is that this study using SLR verifies the significance of self- efficacy for IS security innovation adoption. Another key implication of this study is the importance of spreading IS security knowledge among the users for safeguarding IS assets. On one hand, knowledge has a simple positive effect on self-efficacy, which affects the individual's security behaviour. On the other hand, knowledge allows users to assess a security technology fairly and improve the quality of decision making. IS security literature has emphasised on the need to pay attention to security education, awareness and training initiatives and interventions. Therefore, we suggest that organizations create appropriate education, training and security awareness programs that ensure employees possesses up-to-date knowledge of IS security as well as facilitate conditions that will improve their individual self-efficacy regards IS threats. This study has certain limitations. The major limitation of this analysis was the inadequacy of studies that examined individual self-efficacy on IS security innovation adoption. The result of the SLR would be more accurate and better explained if analysed with more studies.

References

1. Alshboul, A. 2010. “Information Systems Security Measures and Countermeasures: Protecting Organisational Assets from Malicious Attacks,” Communications of the IBIMA, pp 9p.
2. Anderson, C.L., and Agarwal, R. 2010. “Practicing Safe Computing: A Multimedia Empirical Examination of Home Computer User Security Behavioral Intentions,” MIS Quarterly (34:3), pp 613-643.
3. Arachchilage, N.A.G., and Hameed, M.A. 2017. “Integrating Self-efficacy into a Gamified Approach to Thwart Phishing Attacks,” In: The Proceedings of 5th International Conference on Cybercrime and Computer Forensics (ICCCF), arXiv:1706.07748.
4. Arachchilage, N.A.G., and Love, S. 2014. “Security Awareness of Computer Users: A Phishing Threat Avoidance Perspective,” Computers in Human Behavior (38), pp 304-312.
5. Arachchilage, N.A.G., Love, S., and Beznosov, K. (2016). “Phishing Threat Avoidance Behaviour: An Empirical Investigation,” Computers in Human Behavior (60), pp 185- 197.
6. Bandura, A. 1983. “Self-efficacy Determinants of Anticipated Fears and Calamities,” Journal of Personality and Social Psychology (45), pp 464-469.
7. Bandura, A. 1977. “Self-Efficacy: The Exercise of Control,” NY: W. H. Freeman and Company.
8. Banu, M.N., and Banu, S.M. 2013. “A Comprehensive Study of Phishing Attacks,” International Journal of Computer Science and Information Technologies (4:6), pp 783-786.
9. Bulgurcu, B., Cavusoglu, H., and Benbasat, I. 2010. “Information Security Policy Compliance: An Empirical Study of Rationality-Based Beliefs and Information Security Awareness,” MIS quarterly 34(3), pp 523-555.
10. Chan M., Woon I.M.Y., and Kankanhalli A. 2005. “Perceptions of Information Security at the Workplace: Linking Information Security Climate to Compliant Behavior,” Journal of Information Privacy and Security (1:3), pp. 18-41.
11. Chenoweth, T., Minch, R. and Gattiker, T. 2009. “Application of Protection Motivation Theory to Adoption of Protective Technologies,” In: The Proceedings of the 42nd Hawaii International Conference on System Sciences.
12. Chou, H., and Chou, C. 2016. “An Analysis of Multiple Factors Relating to Teachers' Problematic Information Security Behavior,” Computers in Human Behavior (65), pp 334-345.
13. Compeau, D.R., and Higgins, C.A. 1995. “Computer Self- efficacy: Development of a Measure and Initial Test,” MIS Quarterly (19:2), pp 189–211.
14. Cox, J. 2012. “Information Systems User Security: A Structured Model of the Knowing–Doing Gap,” Computers in Human Behavior (28) pp 1849-1858.
15. Crossler, R. E. 2010. “Protection Motivation Theory: Understanding Determinants to Backing up Personal Data,” In: The Proceedings of the 43rd Hawaii International Conference on System Sciences.

16. Crossler, R.E., Johnston, A.C., Lowry, P.B., Hu, Q., Warkentin, M., and Baskerville, R. 2013. "Future Directions for Behavioral Information Security Research," Computers and Security (32), pp 90–101.
17. D'Arcy J. and Hovav, A. 2004. "The Role of Individual Characteristics on the Effectiveness of IS Security Countermeasures," In: The Proceedings of Americas Conference on Information Systems 2004. 176. http://aisel.aisnet.org/amcis2004/176
18. Deans, C.P., Karawan, K.R., Goslar, M.D., Ricks, D.A. and Toyne, B. 1991. "Identification of Key International Information Systems Issues," Journal of High Technology Management Review (2:1), pp 57-81.
19. Dinev, T., Goo, J., Hu, Q., and Nam, K. 2009. "User Behaviour Towards Protective Information Technologies: The Role of National Cultural Differences," Information Systems Journal (19), pp. 391–412
20. Eastin, M.S and LaRose, R. 2000. "Internet Self-Efficacy and the Psychology of the Digital Divide," Journal of Computer- Mediated Communication (6:1).
21. Feruza, Y.S., and Kim, T. 2007, "IT Security Review: Privacy, Protection, Access Control, Assurance and System Security," International Journal of Multimedia and Ubiquitous Engineering (2:2), pp 17-32.
22. Floyd, D.L., Prentice-Dunn, S., and Rogers, R.W. 2000. "A Meta-Analysis of Research on Protection Motivation Theory," Journal of Applied Social Psychology (30:2), pp 106-143.
23. Hameed M.A., and Arachchilage N.A.G. (2016). "A Model for the Adoption Process of Information System Security Innovations in Organisations: A Theoretical Perspective," In: The Proceeding of the 27th Australasian Conference on Information Systems, arxiv.org/abs/1609.07911.
24. Hameed, M.A., and Counsell, S. 2012. "Assessing the Influence of Environmental and CEO Characteristics for Adoption of Information Technology in Organizations," Journal of Technology Management and Innovation (7:1), pp 64-84.
25. Hameed, M.A., and Counsell, S. 2014a. "Establishing Relationship between Innovation Characteristics and IT Innovation Adoption in Organisations: A Meta-analysis Approach," International Journal of Innovation Management (18:1), p 41.
26. Hameed, M.A., and Counsell, S. 2014b. "User Acceptance Determinants of Information Technology Innovation in Organisations," International Journal of Innovation and Technology Management (11:5), p 17.
27. Hameed, M.A., Counsell, S., and Swift, S. 2012. "A Meta- analysis of Relationships between Organisational Characteristics and IT Innovation Adoption in Organisations," Information and Management (49:5), pp 218-232.
28. Hanus, B. and Wu, Y. A. 2016. "Impact of Users' Security Awareness on Desktop Security Behavior: A Protection Motivation Theory Perspective," Information Systems Management (33:1), pp 2-16.

29. Herath, T., Chen, R., Wang, J., Banjara, K., Wilbur, J., and Rao, H. R. 2014. "Security Services as Coping Mechanisms: An Investigation into User Intention to Adopt an Email Authentication Service," Information Systems Journal (24:1), pp 61-84.
30. Hearth, T., and Rao, H.R. 2009. "Protection Motivation and Deterrence: A Framework for Security Policy Compliance in Organizations," European Journal of Information Systems (18:2), pp. 106-125.
31. Hedges, L.V., and Olkin, I. 1985. "Statistical Methods for Meta-Analysis," Academic Press, Orlando
32. Hunter, J.E., Schmidt F.L., and Jackson, G.B. 1982. "Meta-Analysis," Beverly Hills, CA: Sage.
33. Ifinedo, P. 2012. "Understanding Information Systems Security Policy Compliance: An Integration of the Theory of Planned Behaviour and the Protection Motivation Theory," Computers and Security (31), pp 83-95.
34. Ifinedo, P. 2014. "Information Systems Security Policy Compliance: An Empirical Study of the Effects of Socialisation, Influence, and Cognition," Information and Management (51:1), pp 69-79.
35. Johnston, A.C., and Warkentin, M. 2010. "Fear Appeal and Information Security Behaviors: An Empirical Study," MIS Quarterly (34:3), pp 549-566.
36. Lai, F., Li, D. and Hsieh, C. 2012. "Fighting Identity Theft: The Coping Perspective," Decision Support Systems (52), pp 353-363.
37. Lee, D., Larose, R., and Rifon, N. 2008. "Keeping Our Network Safe: A Model of Online Protection Behaviour," Behaviour and Information Technology (27:5), pp 445–454.
38. Lee, Y., and Larsen K.R. 2009. "Threat or Coping Appraisal: Determinants of SMB Executive's Decision to Adopt Anti- malware Software," European Journal of Information Systems (18:2), pp 177-187.
39. Lee, Y., Lee, J.Y. & Liu, Y. 2007. "Protection Motivation Theory in Information System Adoption: A Case of Anti- Plagiarism System," In: The Proceedings of Americas Conference on Information Systems 2007. 62. http://aisel.aisnet.org/amcis2007/62
40. Liang, H., and Xue, Y. (2010). "Understanding Security Behaviors in Personal Computer Usage: A Threat Avoidance Perspective," Journal of the Association for Information Systems (11:7), pp 394-414.
41. Lui, S.M., and Hui, W. 2011. "The Effects of Knowledge on Security Technology Adoption: Results from a Quasi- experiment," In: The Proceedings of the 5th International Conference on New Trends in Information Science and Service Science.
42. Marett, K., Harris, R.B. and McNab, A.L. 2011. "Social Networking Websites and Posting Personal Information: An Evaluation of Protection Motivation Theory," Transactions on Human-Computer Interaction (3:3), pp 170-189.
43. Meso, P., Ding, Y., and Xu, S. 2013. "Applying Protection Motivation Theory to Information Security Training for College Students," Journal of Information Privacy and Security (9:1), pp 47-67.

44. Milne, G.R., Labrecque, L.I., and Cromer, C. 2009. “Toward an Understanding of the Online Consumer’s Risky Behavior and Protection Practices,” The Journal of Consumer Affairs (43:3) pp 449-473.
45. Mohamed, N., and Ahmad, I. 2012. Information Privacy Concerns, Antecedents and Privacy Measure Use in Social Networking Sites: Evidence from Malaysia,” Computers in Human Behavior (28), pp 2366–2375.
46. Ng, B.Y., Kankanhalli, A., and Xu, Y. 2009. “Studying Users' Computer Security Behavior Using the Health Belief Model,” Decision Support Systems (46:4), pp 815-825.
47. Pahnila, S., Siponen. M., and Mahmood M.A. 2007a. “Employee’s Behavior Towards IS Security Policy Compliance,” In: The Proceedings of 40th Hawaii International Conference on System Sciences, p. 1561
48. Pahnila, S., Siponen, M. and Mahmood, A. 2007b. “Which Factors Explain Employees’ Adherence to Information Security Policies? An Empirical Study,” In: The Proceedings of Pacific Asia Conference on Information Systems 73. http://aisel.aisnet.org/pacis2007/73
49. Rhee, H., Kim, C., and Ryuc, Y, C. 2009. “Self-efficacy in Information Security: Its Influence on End Users’ Information Security Practice Behaviour,” Computers & Security (28), pp 816-826.
50. Rhodes, K. 2001. “Operations Security Awareness: The Mind has No Firewall,” Computer Security Journal (18:3), pp 27- 36.
51. Rogers, R.W. 1983. “Cognitive and Physiological Processes in Fear Appeals and Attitude Change: A Revised Theory of Protection Motivation,” In: J. Cacioppo and R. Petty (Eds.), Social Psychophysiology. New York: Guilford Press, pp 153- 176.
52. Schepers, J., and Wetzels, M. 2007. “A meta-analysis of the technology acceptance model: Investigating subjective norm and moderation effects,” Information and Management (44), pp 90-103
53. Siponen, M., Mahmood, M.A., and Pahnila, S. 2014. “Employees’ Adherence to Information Security Policies: An Exploratory Feld Study,” Information & Management (51), pp 217-224.
54. Siponen, M. T., Pahnila, S., and Mahmood, A. 2007. “Employees’ Adherence to Information Security Policies: An Empirical Study,” In: T he Proceedings of the International Federation for Information Processing IFIP SEC 2007 Conference 2007.
55. Son, J.Y. 2011. “Out of Fear or Desire? Toward a Better Understanding of Employees' Motivation to Follow IS Security Policies,” Information & Management (48:7), pp 296-302.
56. Stanton, J., Stam, K., Mastrangelo, P., and Jolton, J. 2005. “Analysis of End User Security Behaviors,” Computers and Security (24:2), pp 124–133.
57. Gupta, B. B., Arachchilage, N. A., & Psannis, K. E. (2018). Defending against phishing attacks: taxonomy of methods, current issues and future directions. *Telecommunication Systems*, *67*(2), 247-267.
58. Sun, J.C., Yu, S., Lin, S.S.J., and Tseng, S. 2016. “The Mediating Effect of Anti-phishing Self-efficacy Between College Students’ Internet Self-

efficacy and Anti-phishing Behavior and Gender Difference," Computers in Human Behavior (59), pp 249-257.
59. Tamjidyamcholo, A., Baba, A. S. B., Gholipour, R. and Yamchello, H. T. 2013. "Information Security Professional Perceptions of Knowledge-Sharing Intention in Virtual Communities under Social Cognitive Theory," In: The Proceedings of the 3rd International Conference on Research and Innovation in Information Systems – 2013.
60. Tamjidyamcholo, A., Bin Baba, M.S., Tamjid, H., and Gholipour, R. 2013. "Information Security - Professional Perceptions of Knowledge-Sharing Intention under Self- efficacy, Trust, Reciprocity, and Shared-language," Computers and Education (68), pp 223-232.
61. Torkzadeh, R., Pflughoeft, K., and Hall, L. 1999. "Computer Self-efficacy, Training Effectiveness and User Attitudes. An Empirical Study," Behaviour and Information Technology (18:4), pp 299-309.
62. Tornatsky, L.G. and Fleischer, M. 1990. "The Process of Technological Innovation," Lexington Books.
63. Tsai, H.Y. S., Jiang, M., Alhabash, S., LaRose, R., Rifon, N. J., and Cotten, S. R. 2016. "Understanding Online Safety Behaviors: A Protection Motivation Theory Perspective," (59), pp 138-150.
64. Vance, A., Siponen, M., and Pahnila, S. 2012. "Motivating IS Security Compliance: Insights from Habit and Protection Motivation Theory," Information and Management (49) pp 190-198.
65. Venkatesh, V. and Davis, F. D. 2000. "A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies," Management Science (46:2), pp 186–204.
66. Vroom, C., and von Solms, R. 2004. "Towards Information Security Behavioural Compliance," Computers & Security (23), pp 191-198.
67. Warkentin, M., Johnston, A.C. Shropshire, J., and Barnett, W. D. 2016. "Continuance of Protective Security Behavior: A Longitudinal Study," Decision Support Systems (92), pp 25– 35.
68. Wei, L., and Zhang, M. 2008. "The Impact of Internet Knowledge on College Students' Intention to Continue to Use the Internet," Information Research (13:3), p 348.
69. Woon, I., Tan, G., and Low, R. 2005. "A Protection Motivation Theory Approach to Home Wireless Security," In: The Proceedings of International Conference on Information Systems, p 31.
70. Workman, M., Bommer, W., and Straub, D. 2008. "Security Lapses and the Omission of Information Security Measures: A Threat Control Model and Empirical Test," Computers in Human Behavior (24), pp 2799-2816.
71. Yoon, C., Hwang, J. W., and Kim, R. (2012). "Exploring Factors that Influence Students' Behaviors in Information Security," Journal of Information Systems Education 23(4), pp 407-417.
72. Zmud, R. W. 1982. "Diffusion of Modern Software Practices: Influence of Centralization and Formalization," Management Science (28:12), pp 1421–1431.
73. Arachchilage, N. A. G., & Love, S. (2013). A game design framework for avoiding phishing attacks. *Computers in Human Behavior*, *29*(3), 706-714.

74. Arachchilage, G., & Asanka, N. (2012). *Security awareness of computer users: A game based learning approach* (Doctoral dissertation, Brunel University, School of Information Systems, Computing and Mathematics).

APPENDIX

SDY NAME	YER	SAM		CNTRY	Theories	INN TYP	COR
Herath and Rao (2009)	2009	ORG	312	USA	PMT, DT, DTPB	PRC	0.51
Ng et al. (2009)	2009	MIX	134	Singapore	HBM	PRD	0.4
Mohamed and Ahmad (2012)	2012	SDT	340	Malaysia	PMT, SCT	PRC	0.419
Son (2011)	2011	IND	602	USA	EMM, IMM	PRC	0.23
Workman et al. (2008)	2008	IND	588	USA	PMT	PRC	
Rhee et al. (2009)	2009	SDT	415	USA	SCT	PRC	0.363
Johnston and Warkentin (2010)	2010	MIX	215	USA	PMT	PRD	0.342
Bulgurcu et al. (2010)	2010	ORG	464	Canada	TPB, RCT	PRC	0.395
Yoon et al. (2012)	2012	SDT	202	South	PMT	PRC	0.1
Ifinedo (2012)	2012	IND	124	Canada	TPB, PMT	PRC	0.32
Ifinedo (2014)	2014	IND	124	Canada	TPB, SCT, SBT	PRC	0.24
Anderson and Agarwal (2010)	2010	IND	594	USA	PMT	PRD	0.44
Anderson and Agarwal (2010)	2010	IND	101	USA	PMT	PRD	0.38
Chou and Chien Chou (2016)	2016	IND	505	Taiwan	PMT	PRD	0.05
Warkentin et al. (2016)	2016	SDT	253	USA	PMT	PRD	0.888
Siponen et al. (2014)	2014	IND	669	Finland	TRA, CET	PRC	0.243
Tamjidyamcholo et al. (2013)	2013	IND	138	Malaysia	TRA, SCT	PRC	0.566
Lee et al. (2008)	2008	SDT	273	USA	PMT	PRD	0.6
Vance et al. (2012)	2012	IND	210	Finland	PMT	PRC	0.47
Chan et al. (2005)	2005	IND	104	Singapore		PRC	0.4
Herath et al. (2014)	2014	SDT	134	USA	TAM, TTAT	PRD	-0.08
Marett et al. (2011)	2011	SDT	522	USA	PMT	PRC	0.51
Lui and Hui (2011)	2009	SDT	752	China	TAM	PRD	0.082
Wei and Zhang (2008)	2008	SDT	279	China	TAM	PRC	0.32
Sun et al. (2016)	2016	SDT	411	Taiwan		PRD	0.52
Sun et al. (2016)	2016	SDT	411	Taiwan		PRD	0.45
Liang and Xue (2010)	2010	SDT	152	USA	TTAT	PRD	0.283
Dinev et al. (2009)	2009	SDT	332	USA	TPB	PRD	0.39
Dinev et al. (2009)	2009	SDT	227	South	TPB	PRD	0.35
Hanus and Wu (2016)	2016	SDT	229	USA	PMT	PRC	0.65
Lai et al. (2012)	2012	SDT	117	USA	CT	PRC	-0.186
Meso et al. (2013)	2013	SDT	77	USA	PMT	PRD	0.784
Siponen et al. (2007)	2007	IND	917	Finland	PMT, DT, TRA	PRC	0.407
Tamjidyamcholo et al. (2013)	2013	SDT	138	Malaysia	PMT	PRC	0.565
Tsai et al. (2016)	2016	IND	988	USA	PMT	PRC	0.26
Chenoweth et al. (2009).	2009	IND	204	USA	PMT	PRD	
Crossler (2010)	2010	IND	112	USA	PMT	PRD	
D'Arcy and Hovav (2004)	2004	NON			DT	PRC	
Cox (2012)	2012	IND	106	USA	TPB	PRC	0.43
Lee et al. (2007)	2007	NON		USA	PMT	PRD	
Milne et al. (2009)	2009	IND	449	USA	PMT, SCT	PRC	
Pahnila et al. (2007)	2007	IND	917	Finland	PMT, DT, TRA	PRC	

[YER - Year], [SAM G - Sample Group: IND -Individual; ORG - Organisation; SDT - Student; MIX - Mixed; NON - None], [SAM S - Sample Size], [CNTRY - Country], [Theories: PMT - Protection otivation Theory; TPB - Theory of Planned Behaviour; TRA - Theory of Reasoned Action; SCT - Social Cognitive Theory; DT - Deterrence Theory; TAM - Technology Acceptance Model; TTAT - Technology Threat Avoidance Theory; CET - Cognitive Evaluation Theory; CT - Coping Theory; DTPB - Decomposed Theory of Planned Behaviour; EMM - Extrinsic Motivational Model; HBM - Health Belief Model; IMM - Intrinsic Motivation Model; RCT - Rational Choice Theory; SBT - Social Bond Theory], [INN TYP - Innovation Type: PRC - Process; PRD - Product], [COR - Correlation]

Political Cyber Operations

Matthew Warren
Deakin University Centre for Cyber Security Research and Innovation,
Deakin University, Geelong, Victoria, Australia.
matthew.warren@deakin.edu.au

Abstract
Social media impacts all aspects of society from citizens to businesses but also political parties. The paper proposes a new social media engagement model that evaluates political cyber operations and the success of such campaigns during elections. The paper will use the Cook Islands 2018 general election to validate the model.

Keywords: Social Media, Elections, Cook Islands and Information Operations.

1) Introduction

Social media has been defined as "a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of user-generated content" (Kaplan & Haenlein, 2010). Social media impacts all aspects of society from citizens to businesses but also political parties.

Social Media offers real challenges for political parties as there is increased acceptance of social media by voters. It also means that political discussions are conducted in a public forum and voters have the ability to contribute to the discussion. This means that political parties may have little control over the discussion or even lose control of the discussion that occur online. The means that social media has real challenges for political parties.

So why is social media so important for political parties. It is important because of the large and rapidly increasing number of users (voters) using social media and their increased online expectations. It is also important because users (voters) have expectation around the use of technology to engage with a variety of organisations and individuals, social media has become the accepted standard due to its of widespread use and easy of use, there is the expectation that users (voters) can engage with political parties.

From a political party perspective, social media provides a cost-effective medium to reach-out to large number of users (voters), it provides a rich two way engagement with users (voters) and by its nature creates interaction. Social media also offers a business benefits for political parties, by using social media they could engage with many more users (voters) rather than traditional media, so it means their investment in social media could give greater returns.

Another key aspect of the use of social media by political parties is that it allows them to influence voters and they way that could vote, this is also known as information operations. Information operations also known as

influence operations, includes the collection of tactical information about an adversary as well as the dissemination of propaganda in pursuit of a competitive advantage over an opponent. (Waltzman, 2017).

The paper presents a model that allows for the assessment of information operations by political parties and uses an election campaign to validate the model.

2) Research Question
The researchers have identified a number of key research questions that they wanted resolved. The research questions are:

- What impact does social media have upon political parties engagement with voters (influence operations);

- What impact does social media have upon leaders of political parties engagement with voters (influence operations);

In order to answer these questions, a proposed Social Media (SM) voter engagement model (VEM) (SM-VEM) had been developed to determine the levels of engagements that occurred between political parties and stake holders during a particular period of time, e.g. during an election. The model will be validated based upon election data from the 2018 Cook Islands election.

3) *Model Development*
There was no model identified that could answer the research questions that were defined at the start of the paper. Therefore the researchers had to develop a new model. The development of the SM-VEM Model was based on the following steps of development:

Stage 1 Assessment of the Honeycomb model;
Stage 2 Analysis of other social media management methods;
Stage 3 Consideration of political use of social media;
Stage 4 Development of SM-VEM model.

3.1) Stage 1 Assessment of the Honeycomb Model
One of the first engagement models that was developed was the Honeycomb model, this was developed to understand how consumers engage via social media with businesses (Kietzmann et al, 2011). The honeycomb model is based upon seven functionality blocks identified in the Honeycomb model and further explained in their follow-up work (Kietzmann et al, 2011, Kietzmann et al, 2012). The seven function blocks of social media used in the model are (Kietzmann et al, 2011):

- Identify - the extent to which users reveal their identities in a social media setting;
- Conversation – the extent to which users communicate with other users in a social media setting;
- Sharing - the extent to which users exchange, distribute, and receive content;

- Presence – the extent to which users can know if other users are accessible;
- Relationship - the extent to which users can be related to other users;
- Reputation - represents the extent to which users can identify the standing of others, including themselves, in a social media setting;
- Groups – the extent to which users can form communities and sub communities. The more 'social' a network becomes, the bigger the group of friends, followers, and contacts.

This model represented an the first model to determine engagement with social media and voters,

3.2) *Stage 2 Analysis of other social media mangagement methods;*
The next stage was to look at research that cover social media engagement across a whole number of sectors. A number of different social media engagement models were analysed to determine the engagement themes that were common to those models. The engagement models looked at social media engagement from a wide range of areas, e.g. higher education, medical usage, banking sector, heath promotion crisis and disaster management and marketing. The themes and reference sources are presented in Table 1.

Themes	Sources
Branding / Authenticity	Anderson (2011), Maxwell (2012), Bottles and Sherlock (2011), Patino et al. (2012), Dutta (2010), Porter et al. (2011), Voss and Kuma (2012), Thomas and Thomas (2012), Li and Bernoff (2008), Panagiotopoulos et al (2015), Senadheera (2015).
Listening	Dutta (2010), Li and Bernoff (2008), Maxwell (2012), Patino et al. (2012), Temin (2012), Voss and Kuma (2012), Panagiotopoulos et al (2015), Senadheera (2015), Houston (2014).
Engagement	Anderson (2011), Lefever (2012), Li and Bernoff (2008), Brenner (2012), Maxwell (2012), Dutta (2010), Patino et al. (2012), Heiberger and Junco (2011), Porter et al. (2011), Thomas and Thomas (2012), Voss and Kuma (2012), Panagiotopoulos et al (2015), Senadheera (2015), Houston (2014), Petch (2004).
Visibility	Anderson (2011), Klososky (2012), Li and Bernoff (2008), Dutta (2010), Maxwell (2012), Temin (2012), Voss and Kuma (2012), Senadheera (2015).

Relationship	Anderson (2011), Klososky (2012), Li and Bernoff (2008), Bottles and Sherlock (2011), Brenner (2014), Patino et al. (2012), Dutta (2010), Porter et al. (2011), Heiberger and Junco (2011), Temin (2012), Thomas and Thomas (2012), Voss and Kuma (2012), Senadheera (2015), Houston (2014), Petch (2004).
Trust	Anderson (2011), Klososky (2012), Brenner (2014) Li and Bernoff (2008), Bottles and Sherlock (2011) Patino et al. (2012), Li and Bernoff (2008), Li and Bernoff (2008), Voss and Kuma (2012), Houston (2014).
Organisational Impact	Panagiotopoulos et al (2015), Senadheera (2015).

Table 1: **Thematic analysis of focused Social Media Engagement Models**

Based upon the assessment of the key themes that were identified at Table 1, the following themes areas were determined:

- Branding / Authenticity – extent to which the organisation are identified as being that organisation, e.g. you can tell from the social media channels that company A is company A;
- Listening – extent to which organisations can receive questions from users via social media;
- Engagement – extent to which organisations responds to questions or post new information via social media;
- Visibility – the ease by which users can find organisations social media channels;
- Relationship – represents the extent to which users can be related to other users (as per the Honeycomb Model).
- Trust – the trust that users have upon information being posted.
- Organisational Impact – the agility of the organisation to change their social media strategy.

3.3) Stage 3 Social Media Engagement Areas and Themes

The next stage was to compare the Honeycomb model (Stage 1) against the other models that been identified (Stage 2) to identify a common functions and general themes, these are shown in Table 2.

Honeycomb Model Function Blocks	*Other Models Themes*
Identify	Branding / Authenticity
Conversation	Listening
Sharing	Engagement
Presence	Visibility
Relationship	*Relationship*
Reputation	Trust
Groups	Organisational Impact

Table 2: **Social Media Engagement Areas and Themes**

What has been identified are a number of key areas that relate to the function of social media and number of themes that organisations should relate to when they are using social media. The only common theme is that of the development of relationships via social media (highlighted in bold), which importantly relates to the extent that users can be identify to other users.

We now need to consider how social media is used by political parties to engage with voters and whether there are unique or common attributes.

3.4) Voter Engagement via Social Media

The next stage was to identify key research themes relating to voter engagement via social media. This stage looked at studies into political social media usage around the world and looked specifically at the following countries: Australia, USA, Brazil, Sweden, UK, Ireland and New Zealand. The themes that were identified were:

Themes	Sources
Interactive	Macnamara et al, (2012), Grussell and Nord (2012), Bruns and Highfield (2013), Sauter and Bruns (2013).
Engagement	Macnamara et al, (2012), Flew (2008), Bruns and Highfield (2013), Bruns and Highfield (2013), Vitak and Zube (2011), Gilmore and Howard (2013), Newman (2010), Lynch and Hogan (2012), Bruns and Highfield (2015), Cantijoch (2012)

Channels	Macnamara et al, (2012), Flew (2008), Bruns and Highfield (2013), Newman (2010).
Relationship	Bruns and Highfield (2013), Gong and Lip (2009), Newman (2010).
Organisational Strategy	Macnamara et al, (2012), Flew (2008), Bruns and Highfield (2013), Sauter and Bruns (2013), Vitak and Zube (2011), Gilmore and Howard (2013), Grusell and Nord (2012), Newman (2010), Bruns and Highfield (2015), Cantijoch (2012).

Table 3. **Voter Social Media Engagement**

Based upon the analysis of the studies (as shown by Table 3), the common voter engagement themes via social media that were identified, were:

- Interactive – the extent to which voters communicate with other voters in a social media setting;
- Engagement – extent to which political parties respond to questions posted via voters on social media;
- Channels – the different identifiable social media channels that different political parties use;
- Relationship - the extent to which voters can be related to other voters;
- Organisational strategy – the extent to which political parties develop and evolve an engagement strategy.

3.5) Stage 4 Development of Voter Engagement Model
The next stage was to merge the information form stage 1,2,3 and develop into a single model. This comparison is shown in table 4.

Political Social Media Themes (stage 3)	Honeycomb Model Function Blocks (stage 1)	Other Models Themes (stage 2)
Channels	Identify	Branding / Authenticity
Interactive	Conversation	Listening
Engagement	Sharing	Engagement
Organisational Strategy		Organisational Impact
Relationships	Relationship	Relationship

	Reputation	Trust
	Presence	Visibility
	Groups	

Table 4: **Comparison of Social Media Engagement Areas and Themes**

So the researchers developed the following model to model political social media engagement, as shown by Table 5.

Functionality	Application of the functionality in relation to Voter Engagement
Channels	This is a functionality provided by social media technologies for users to form an online presence.
Listening	Conversation functionality is present key social media technologies.
Engagement	Online communities formed on key social media technologies.
Relationships	All key social media technologies support the formation of non-reciprocal relationships.
Reputation	Considering the public nature of the communities formed on social media.
Organisational Strategy	Extent that organisations have developed a strategy or adapt a strategy.

Table 5: **Development of SM-VEM Model**

The next step was to link the stages of the SM-VEM Model to the social media aspects. This linkage is shown in Table 6.

Functionality	Data Type	Measured Constructs/Observed themes		
		Facebook	Twitter	YouTube
Listening	Quantitative	Wall Posts (Hard to collect this information)	Tweets	
Engagement	Quantitative			Video Uploads
Relationships	Quantitative	Page Likes	Followers	Subscribers
Reputation	Qualitative	Verified account and links through official websites.		
Channels	Qualitative	Brand name, logo, colours, and contact information.		

Organisational Strategy	Qualitative	Description of organisational strategy over time.

Table 6: **Mapping of key social media functionalities to SM-VEM model**

4) Case Study

To validate the SM-VEM model a real life case study was selected. The case study was the 2018 Cook Islands Election and data was collected between 27th May – 15th June, 2018. The Cook Islands are a country in the South Pacific, the population is 17,000 people of which 10,000 people live in the capital Rarotonga (Cook Islands Government, 2018a).

The Cook Islands has a single parliament that consists of 24 MPs and as of 12th June, 10,917 people enrolled to vote in the Cook Islands election. The pre 2018 parliament consisted of the following break down of MPs (Cook Islands Government, 2014):

Cook Islands Party 13;
Democratic Party 9;
One Cook Islands 2;
Titikaveka Oire No Seats;
Independents No Seats.

5) Assesment of SM-VEM Model

The data collection took the form of collecting social media data, namely Facebook, Twitter, and YouTube data that was collected during the course of the election will be used to validate the model.

5.1) Channels

The first stage was to collect information about the political parties and leaders and the social media platforms that used. This information is presented in Table 7.

Party	Web	Facebook	Twitter	YouTube
Cook Islands Party (CIP)	x	x		
Democratic Party (DP)	x	x		x
Alternative Must Ravenga Openga (AMRA)				
Titikaveka Oire (TO)				
One Cook Islands (OCI)				
Heny Puna (Leader CIP)	x	x		
Tina Browne (Leader DP)				

Table 7: **Cook Islands Political Parties / Leaders use of Social Media**

This highlights the extent to which the organisation are identified as being that organisation, e.g. you can tell from the social media channels that organisation A is organisation A.

This level is also concerned by which users can find organisations social media channels. This is the ease of which voters can find the political parties social media channels. This is shown in Table 7 and shows that CIP and their leader have an official presence on Facebook, on the Web, the DP have an official presence on Facebook, on the Web and YouTube but the leader of the DP has no online presence. In this regards CIP, DP and the Leader of the CIP has definable social media pages. An example of the CIP facebook page is shown in Figure 1.

Figure 1: **Cook Islands Party Facebook Page**

5.2) Listening
This is the extent to which organisations can receive questions from users via social media. The aim of this part of the model is to assess Twitter usage. As none of the political parties or leaders used Twitter, this part of the model could not be assessed.

5.3) Engagement
This is the extent to which organisations responds to questions or post new information via social media with a particular with a particular focus on videos and social media post.

The DP had a dedicated YouTube Channel and during the course of the election posted no videos via that channel and had no new subscribers.

Both the DP and CIP used Facebook to post videos. The DP during the election posted 3 videos with a total of 8900 views which related to an average views of 2966 per video. The CIP during the election posted 15

videos in Facebook, which had a total of 36857 views which related to average of 2457 views per video. The total Cook Islands electoral was 10,917 voters so it showed that the CIP had the greatest impact with their videos via Facebook.

5.4) Relationship
This represents the extent to which users can be related to other users. This stage of the model relates to the use of Facebook and the level of engagement.

In terms of Facebook general engagement during the course of the election, the DP had 133 New Likes for their post and attracted 136 New Follows. The leader of DP (Tina Brown) had no dedicated Facebook page.

During the course of the election, the CIP had 299 New Likes for their post and attracted 307 New Follows. The leader of CIP (Henry Puna) had a dedicated Facebook page and during the election achieved 4 New Likes and 4 New Follows.

5.5) Reputation
This stage relates to the trust that users have upon information being posted. None of the Facebook or YouTube pages of CIP and DP had been officially verified. The outcome was this that voters did not have that the level of assurance surrounding the accounts.

5.6) Organisational Strategy
This stage relates to agility of the organisation to change their social media strategy.

In terms of the election CIP and DP both made use of Facebook as their main social media account, the DP had a old YouTube Channel which they did not use during the course of the election.

Another observation of that the leader of DP and the OCI, AMRA, TO parties did not have any online presence.

6) Discussion
In terms of the election the CIP perform better in terms of social media interaction on Facebook and video views but was not reflect necessary in physical votes. The DP had modest levels of social media engagement but increased their votes apart from the leader of DP losing their seat. It should be noted that the leader of DP had no online presence.

The research shows that social media did not have a huge impact upon the Cook Islands general elections and as such it cannot a key factor. To answer the research questions put forward:

- What impact does social media have upon political parties engagement with voters (influence operations);

  Limited impact upon the Cook Islands Elections.

- What impact does social media have upon leaders of political parties engagement with voters (influence operations);

  Limited impact but the leader of DP who had the most seats and who had no online presence lost her a seat. Could the lack of an online presence been a factor?

The outcome of the election was that the following members of parliament were elected (Cook Islands Government, 2018b):

Cook Islands Party 10 (-3);
Democratic Party 11 (+2);
One Cook Islands 1 (-1);
Titikaveka Oire No seats elected;
Independents 2 (+2).

The Cook Islands Party stayed in power with the support of OCI and Independents MPs that were elected.

7) Conclusion

The aim of the paper was to identify how political parties use social media and the development of a conceptual model to model how political information operations could occurred.

The SM-VEM model was developed and validated using the Cook Islands 2018 general election. The Cook Islands general election provided an example to prove the concept.

The next stage of development is the development of formal modelling linked to Pearson Correlation / Spearman Rank Correlation Coefficient Analysis focusing on key social media interactions, e g. Likes V Talking About on the different platforms.

8) References

Anderson, D.J. (2011), “The foray into social media: a clinician, and skeptic, sold”, Frontiers of Health Services Management, Vol. 28 No. 2, pp. 23-27.
Bottles, K. and Sherlock, T. (2011), “Who should manage your social media strategy?”, Physician Executive, Vol. 37 No. 2, pp. 68-72.
Brenner, J. (2014), “Pew internet: social networking Fact Sheet:”, URL: www.pewinternet.org/Commentary/2012/March/Pew-Internet-Social-Networking-full-detail.aspx, accessed, 15/12/15.
Bruns, A and Highfield, T (2013), ‘Political networks on Twitter’, Information, Communication & Society, vol. 16, no. 5, pp. 667-691.
Bruns, A and Highfield T (2015) Social media in selected Australian federal and state election campaigns, 2010-15, 16th Annual Meeting of the Association of Internet Researchers Conference (AOIR16), Phoenix, USA.
Cantijoch M (2012) Communication in the 2008 U.S. Election. Digital Natives Elect a President, Information, Communication & Society, vol 15,

No 2, pp:324-325.
Cook Islands Government (2014) Election Results http://www.mfem.gov.ck/preliminary-results-2014/53-statistics/other-information, accessed, 22/9/18.
Cook Islands Government (2018a) Cook Islands Demographic Profile, URL: http://www.cookislands.gov.ck/statistics/census-and-surveys/cook-islands-demographic-profile, accessed, 22/9/18.
Cook Islands Government (2018b) Cook Islands Election Results, URL: http://www.mfem.gov.ck/elections, accessed, 22/9/18.
Dutta, S. (2010), "What's your personal social media strategy?", Harvard Business Review, Vol. 88 No. 11, pp. 127-130.
Flew, T (2008) Not yet the Internet election: online media, political commentary and the 2007 Australian federal election. Media International Australia Incorporating Culture and Policy, pp. 5-13.
Heiberger, G. and Junco, R. (2011), "Meet your students where they are: social media", NEA Higher Education Advocate, URL: http://blog.reyjunco.com/pdf/HeibergerJuncoNEA.pdf, accessed, 15/12/15.
Houston, J., Hawthorne, J., Perreault, M., Park, E., (2014) 'Social media and disasters: a functional framework for social media use in disaster planning, response, and research', Disasters, 39, 1, pp.1–22.
Gilmore, J and Howard, N (2013) Does Social Media Make a Difference in Political Campaigns? Digital Dividends in Brazil's 2010 National Elections, Center for Communication and Civic Engagement, 6/5/2013 Working Paper University of Washington, USA.
Gong H and Lips M (2009) The Use Of New Media By Political Parties In The 2008 National Election, Victoria University of Wellington, New Zealand.
Grussell, M and Nord, L (2012) Three Attitudes to 140 Characters: The Use and Views of Twitter in Political Party Communications in Sweden Public Communication Review, Vol. 2 No. 2.
Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. Business Horizons, 53(1), 59–68. doi:10.1016/j.bushor.2009.09.003
Kietzmann, J.H., Hermkens, K., McCarthy, I.P. and Silvestre, B.S. (2011), "Social media? Get serious! Understanding the functional building blocks of social media", Business Horizons, Vol. 54 No. 3, pp. 241-251.
Kietzmann, J.H., Silvestre, B., McCarthy, I and Pitt, L (2012) Unpacking the social media phenomenon: towards a research agenda, Journal if Public Affairs, Vol 12, No 2, pp 109-119.
Klososky, S. (2012), "Social technology: the next frontier", Financial Executive, Vol. 28 No. 4, pp. 40-45.
Lefever, R. (2012), "Exploring student understandings of belonging on campus", Journal of Applied Research in Higher Education, Vol. 4 No. 2, pp. 126-141.
Li, C. and Bernoff, J. (2008), Groundswell: Winning in a World Transformed by Social Technologies, Harvard Business Press, Boston, MA.
Lynch, K., and Hogan, J. (2012). How Irish political parties are using social networking sites to reach generation Z: An insight into a new online social network in a small democracy, Irish Communications Review, vol,

13, pp 83-98.
Macnamara, J, Sakinofsky, P and Beattie, J (2012) *E-lectoral engagement: Maintaining and enhancing democratic participation through social media*, Report to the Australian Electoral Commission by Australian Centre for Public Communication, University of Technology Sydney, Australia.
Maxwell, C. (2012), “How to use social media to win new business”, Director, Vol. 65 No. 6, pp. 46-49.
Newman, N (2010) #UKelection2010, mainstream media and the role of the internet, Working Paper, Reuters Institute for the Studies of Journalism, Oxford University, UK.
Panagiotopoulos, P, Shan L, Barnett, J, Ragan A and McConnon A (2015) A framework of social media engagement: Case studies with food and consumer organisations in the UK and Ireland, International Journal of Information Management, Vol 35, pp-394-402.
Patino, A., Pitta, D.A. and Quinones, R. (2012), “Social media’s emerging importance in market research”, Journal of Consumer Marketing, Vol. 29 No. 3, pp. 233-237.
Petch, T. (2004), “Content analysis of selected health information websites: final report”, available at: www.sfu.ca/act4hlth/pub/working/ContentAnalysis.pdf (accessed November 16, 2012).
Porter, C.E., Donthu, N., MacElroy, W.H. and Wydra, D. (2011), “How to foster and sustain engagement in virtual communities”, California Management Review, Vol. 53 No. 4, pp. 80-110.
Sauter, T and Bruns, A (2013). Social Media in the Media: How Australian Media Perceive Social Media as Political Tools, Australian Research Council, Centre of Excellence for Creative Industries and Innovation, Brisbane, Australia.
Senadheera, V (2015), The adoption of social media by Australian banks to communicate with the public, Ph.D. thesis, Deakin University, Australia.
Temin, D. (2012), “What boards must know about social media”, Corporate Board, Vol. 33, No. 194, pp. 11-15.
Thomas, M. and Thomas, H. (2012), “Using new social media and Web 2.0 technologies in business school teaching and learning”, Journal of Management Development, Vol. 31, No. 4, pp 358-367.
Vitak, J, Zube, P, Smock, A, Caleb, T, Carr, M. Ellison, N and Lampe, C (2011). It’s Complicated: Facebook Users’ Political Participation in the 2008 Election, Cyber Psychology, Behavior, and Social Networking vol. 14, no 3.
Voss, K and Kuma, A (2013) The value of social media: are universities successfully engaging their audience, Journal of Applied Research in Higher Education, Vol 5, Issue 2, pp 156-172.
Waltzman, R (2017), The Weaponization of Information, RAND, URL: https://www.rand.org/content/dam/rand/pubs/testimonies/CT400/CT473/RAND_CT473.pdf, accessed 10/11/18.

Cyber Warfare: An Enquiry into the Applicability National Law to Cyberspace.

Helaine Leggat
MAICD, CISSP, CISM, CIPP/US, CIPP/IT
Principal Lawyer, Sladen Legal
L5, 707 Collins Street, Docklands 3008
hleggat@sladen.com.au
Disclaimer: These are the personal views of author.

Abstract
The *Tallinn Manual on the International Law Applicable to Cyber Warfare*[i] *(**Tallinn**)* sets out ninety-five 'black-letter rules' governing conflicts and the basis for each in treaty and customary law.

This talk considers the applicability of *national law* to cyberspace. Specifically, whether there is sufficient basis at a national law level to establish norms for acceptable behavior at an international level. The proposition being that it is time for a new kind of international cooperation in relation to cyber warfare and acceptable norms of behavior in cyberspace.

This talk will consider the applicability of current Australian and other national criminal and tort law by using hypothetical scenarios. Specifically, in relation to self-defence, conspiracy and corporate responsibility in the private sector.

The intention is to encourage experts in the national law of various jurisdictions to identify applicable national law to cyberspace, and to cooperate internationally to establish national rules equivalent to the Tallinn work.

Keywords
Tallinn Manual, cyberlaw, cyber security, cyber warfare, self-defence, conspiracy, corporate responsibility, national law, international law.

Introduction and Background
Charter of the United Nations of 1945

The Charter of the United Nations of 1945[ii] (**UN Charter**) records the wish to save succeeding generations from war; respect treaty obligations and international law; and maintain international peace and security. Specifically, it records that armed force shall not be used, save in the common interest.

The purposes of the UN Charter include taking collective measures for the prevention and removal of threats to peace and acts of aggression. Principles for achieving the objectives include settling international disputes by peaceful means that maintain international peace, security

and justice, and refrain from force against the territorial integrity or the political independence of any state. Significantly, Article 51[iii] recognises the inherent right of individual or collective self-defence if an armed attack occurs.[iv]

The Charter of the United Nations, it's Additional Protocol 1[v] (1977), and the Geneva Conventions[vi] (1864 - 1949) cater for pre-digital armed conflict. A voluminous and well documented record of military and academic literature on the evolution to digital conflict exists, which includes, what is now recognised, 'cyberwar' and 'cyber warfare'. This paper, like the Tallinn Manual uses the terms 'cyberwar' and 'cyber warfare' in a purely descriptive, non-normative sense.

The regulatory framework relating to cyberwar and cyber warfare referenced in this enquiry is recorded in Additional Protocol 1, as (i) comprising international agreements (treaty law), (ii) the principles of international law derived from established custom, (iii) the principles of humanity, and (iv) the dictates of public conscience.

This background is included for context, and as a starting point for what I call the *pre-digital system of legal order* that has supported human societies for decades, and which has disintegrated as a result of the internet. It is this disintegration that propels the intended outcomes of this paper. Namely, to cooperate internationally by recognising national rules that already exist, and agreeing that these become new international norms for behaviour in cyberspace.

Precedent

At the outset, I would like to point to precedent in support of my contention that private enterprise, under national law, has been, and can continue to be, instrumental in shaping new international norms based on agreement through the law of contract. I believe this points to international solutions in relation to cyber warfare.

Lex Mercatoria

During the Middle Ages merchants travelling across Europe to trade fairs, markets and seaports needed common ground rules to create trust and confidence for robust international trade. The differences amongst local feudal, royal and ecclesiastical law provided a significant degree of uncertainty and difficulty for the merchants operating in international markets[vii].

Custom and practice evolved into a distinct body of law known as *Lex Mercatoria, a* body of law independent of national laws which assured commercial participation and basic fairness in international trade relationships based on contract and consensus, despite the national law differences.

Lex Informatica

In the digital age participants travelling across information systems have confronted the same unstable and uncertain environments due to numerous national laws, changing rules and conflicting regulations which have arisen as a result of history, culture and religion, that are just as important for participants of the information society as the *Lex Mercatoria* was to merchants hundreds of years ago[viii].

Some twenty years ago, international consensus was reached by nations coming together to cooperate in the interests of international digital trade. The result was the recognition and facilitation of electronic transactions and communications as a result of the United Nations Commission of International Trade Law[ix] (UNCITRAL) model laws and conventions. To date, UNCITRAL has been responsible for two model laws and one convention which have shaped the modernisation and harmonisation of electronic commerce[x]. The connection with cyberwar is that it is the startling success of the digital economy, human nature (including greed opportunism and power politics), that have led to the breakdown of old norms and the need to find consensus on how we might restore trust and certainty to international relationships.

Social Media, Terms and Conditions

Social media behemoths, as private sector entities, regulate the behaviour of their enormous communities through the law of contract in the form of terms and conditions of the use of their platforms. This is *Lex Mercatoria* and *Lex Informatica* in operation. Billions of civilians from innumerable jurisdictions consent to behave in an acceptable manner.

My contention is that private sector entities can similarly cooperate to establish norms of behaviour that would result in new and acceptable forms of behaviour in cyberspace.

The Current Australian Position
Australian Department of Foreign Affairs and Trade

I have argued publicly and in discussions with the Australian Department of Foreign Affairs and Trade (**DFAT**) (under Julie Bishop's leadership), in the presence of personnel from the Attorney-General's Office, with international law experts, and with personnel from the Department of Prime Minister and Cabinet (prior to the formation of the new Department of Home Affairs), that twenty years after the first UNCITRAL model law[xi] it is time for a new model law or conventions that would result in cooperative and normative behaviour for cyberspace, including with respect to cyberwar.

My observation is that DFAT and the Australian Federal Legislature hold very different positions on the need for new information-related laws. DFAT, whose mandate is Foreign Affairs and Trade, is the authority empowered to enter into international agreements. However, the

Department's Cyber Engagement Strategy of 2017[xii], numerous public statements by the DFAT's Ambassador for Cyber Affairs, and discussions with me, make it clear that DFAT believes that Australia has 'enough laws' (on the subject of 'cyber affairs'), and 'no new laws are needed'. DFAT, as part of executive government is mandated to enter into arrangements that would result in model laws and conventions. These in turn would need to be ratified by the Australian Government and adopted into national Australian law. Clearly there is no appetite to do this.

For me, the result is that we need to work with what we have, - thus the approach outlined in this enquiry.

Australian Federal Government

The view of the Australian Federal Legislature is patently different. Driven by increasing levels of cybercrime and terrorism, a raft of data-related legislation has been promulgated and is under consideration[xiii]. Importantly, these are all relate to cyberwar, because they deal with information and information systems which comprise the new, fifth domain of conflict.

Geo-political Shifts and Hegemony

Geo-political shifts in power; hegemonic forces; withdrawal from international cooperation; scale and value of the information economy; freedom and power of civilians resulting from mobility and access to information; the resultant crack-down through surveillance by government on civilians, as well as the rise of executive government (specifically policing) have all contributed to the breakdown of pre-digital legal order.

While governments of Western nations call for partnership between government, private enterprise and academia, I see little being done by governments to facilitate this. They continue to cling to power and have an avaricious need for information collection. This, under the guise of public benefit through sharing, leads to the increased size of the prize for attackers. Information sharing remains largely asymmetric and unhelpful. It compels me to think of other ways to contribute to the establishment of new norms for behaviour in cyberspace. I believe that some measure of delegation of power will be beneficial to the overall safety and security of nation states, individually and collectively. Mechanisms exist in current Australian law, and other national laws which enables this – let the private sector share the national security load. The private sector is well placed and resourced to do this.

The Changing Face of Warfare (Hybrid[xiv] and Asymmetric)
Traditional Wars

Traditional forms of warfare (land, sea, air, space) and early forms of digital warfare (Stuxnet[xv]) were confined to conflict between states as sovereign entities. Conflict, even armed conflict, between sovereign

entities and their own corporate entities or civilians was not recognised as ‘war’.

International law (public) governs the relationships between international sovereign states and entities as largely ‘equal’, being horizontal in power. National laws govern the relationships between sovereign states and their own corporate entities and civilians. Here, power is not equal. These are vertical relationships of power, where the state has power over its corporate entities and civilians.

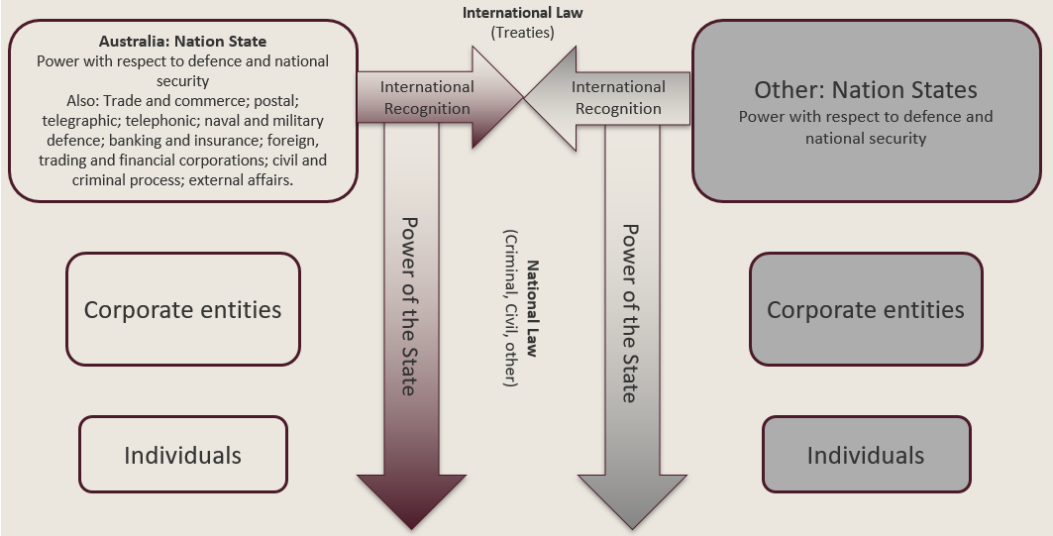

Figure 1: **Relationships of Power**

A breakdown in horizontal relationships in international law can lead to war, usually instigated at the behest of a government which is responsible for national security. A breakdown of vertical relationships in national law leads to law enforcement by the executive arm of the incumbent government.

The last decades have seen a breakdown of the system of global order, specifically in relation to the horizontal and vertical structures of power.

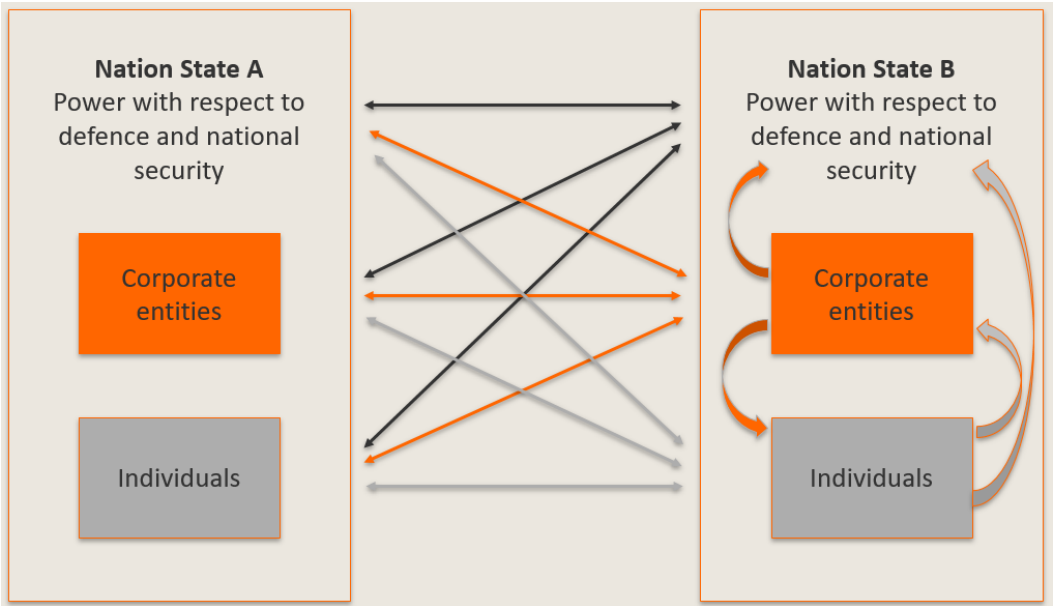

Figure 2: **Structures of Power**

New Wars

From the time that sovereign states began to trade in, and stockpile zero day exploits as digital weapons;[xvi] when sovereign states began to use cyberwar tactics;[xvii] when states sponsored proxies;[xviii] when criminal organisations and individuals became guns for hire;[xix] when technology service providers 'partnered' with governments to weaken trust;[xx] when foreign influence changed democracy;[xxi] when IP theft was measured in billions of dollars;[xxii] when threats to individual privacy became equal to national security,[xxiii] from then, traditional war, was 'new' war. Hybrid. Asymmetric.

If we are to have no new laws, we need to work with the laws we have in order to address new war situations. The stability of global order based on the rule of law, learned from two great wars has dissipated.

One result is that '*Private sector entities operate today on the front lines of cyber conflict, targeted by a variety of hostile actors that seek to steal and misappropriate their intellectual property, degrade their infrastructure, and disrupt their business activities. Despite this reality, the options available within the private sector for responding to cyber threats are outdated and constrained. The status quo is reactive in nature and advantages the attacker*'.[xxiv]

I believe that it is essential to empower private sector entities. I also believe it is imperative to do so in an ordered manner, under the rule of law and with respect to the law of different jurisdictions. This is why I

advocate an approach that recognises similarities in different national legal systems, and effectively seeks to establish a common standard.

The Enquiry

Approach

This enquiry into whether there is sufficient basis at a national law level to establish norms for acceptable behaviour at an international level has involved an examination of relevant aspects of:
Tallinn Manual on the Law Applicable to Cyber Warfare.
Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations.[xxv]
Hypothetical scenarios to introduce topics for consideration in relation to criminal, tort and common law, and self-defence, conspiracy and corporate responsibility in the private sector.
National Laws of the following Jurisdictions: Australia, New Zealand, USA, UK, China, Singapore and India.
The scope constraints in presenting this paper mean that we will be able to include only a sample of examples. Nevertheless, I believe this approach is compelling and points to the need for further work. If you are interested in viewing and discussing the supporting research, please contact me.

Cyber Attack, Framework and Scenarios

Definition of a Cyber Attack and Applicability

Tallinn defines a 'cyber attack' as 'a cyber operation, whether offensive or defensive, that is reasonably expected to cause injury or death to persons or damage or destruction to objects'. The definition applies equally in international and non-international armed conflict (Rule 92. Tallinn 2.0). i.e. the principles of international law outlined in the Tallinn Work applies to the scenarios. The Tallinn Work is included as a frame of reference to the examination of national laws, Clearly, there must be alignment between the two.

Summary of Tallinn Principles

Most Tallinn principles apply to international and non-international conflict. Some exclusions include neutrality and belligerent reprisals. This summary is indicative, not conclusive, of the principles in Tallinn 1 (T1) and Tallinn 2.0 (T2) that have parallel reference to the scenarios. Many of

the findings set forth in the conclusion of this paper depended upon interpretation. (By way of example, ‘unauthorised access’ is a reinterpretation of ‘trespass’ in property law).

Sovereignty. (T1. Rule 1) (T2. Rule 1).
Due diligence. (T2. Rule 6).
Jurisdiction. (T 1. Rule 2) (T2. Rule 8).
Responsibility. (T1. Rule 6) (T2. Rule 14).
Non-state actors rarely regulated by international law. (T2. Rule 33).
Control of infrastructure and adverse knowledge. (T1. Rule 5).
Legal responsibility for cyber operations. (T1. Rule 6).
Attribution not necessarily linked to routing through state. (T1. Rule 8)
Countermeasures to be proportionate. (T1. Rule 9) (T2. Rule 23).
Lawful use of force. (T1. Rule 11) (T2. Rule 68).
Lawful self-defence. (T1. Rule 13) (T2. Rule 19, 71).
Necessity and proportionality as requirements for lawfulness. (T1. Rule 14) (T2. Rule 26, 72).
Imminence and immediacy as requirements for lawfulness. (T1. Rule 15) (T2. Rule 73).
Consequences of participation civilians forfeit protection. (T1. Rule 29).
Distinction between civilian and military targets. (T1. Rule 31) (T2. Rule 96).
Prohibition on attacking civilian objects. (T1. Rule 37) (T2. Rule 99).
Means and methods apply equally to international and non-international conflict. (T1. Rule 41).
Prohibition against unnecessary injury and suffering. (T1. Rule 42) (T2. Rule 104).
No booby-trap objects. (T1. Rule 44) (T2. Rule 106).
Additional protocol 1 prohibits reprisals against civilians, civilian objects, dams, dykes, and nuclear electrical generating stations. (T1. Rule 47, 80, 81) (T2. Rule 109,140,141).
Cyber attacks not directed at lawful targets are prohibited. (T1. Rule 49).
Prohibition against excessiveness/proportionality. (T1. Rule 51).
Constant care to be taken for civilians and civilian objects. (T1. Rule 52) (T2. Rule 114).
Precautions to choose means and methods to cause minimal incidental injury to civilians and destruction of civilian property. (T1. Rule 54) (T2. Rule 117).
Precautions us to proportionality. (T1. Rule 55) (T2. Rule 113).
Effective advanced warnings. (T2. Rule 120).
Cancellation or suspension of attacks that cause incidental loss to civilian life and civilian objects. (T1. Rule 57) (T2. Rule 119).
Effective warnings where attacks affect civilian population. (T1. Rule 58) (T2. Rule 102).
Ruses are permitted. (T1. Rule 61.) (T2. Rule 123).
Cyber espionage and information gathering do not violate law of armed conflict. (T1. Rule 60, 66) (T2. Rule 32).
Protected persons, objects and activities. (T1. Rule 50,75,76, 112) (T2. Rule 123)
How are these principles reflected in the national laws of the ten jurisdictions considered?

Scenario: Bank (private sector)

Current DDoS attack against online banking systems, exploiting banking client devices. Attacker seeks to obtain client credentials to commit theft. Bank knows location of the C&C server of the attacker.

Scenario: Hospital (private sector)

Imminent ransomware attack on hospital system, will exploit vulnerability in third party POS system in hospital. Attacker seeks to extort ransom from hospital.

Issues arising (some examples)

Timing is critical to lawfulness of the response (imminent, current, post attack).
Is a crime perpetrated against the bank or hospital, it's customers and/or third parties?
Does tort law apply (wrongfulness in civil, as opposed to criminal law)?
Does failure by the bank or hospital (its directors and officers) to act (omission) constitute a failure in the exercise of due care and diligence, or negligence?
Does the bank or hospital have a right or obligation to protect/defend the bank or hospital, its assets (money, property, infrastructure), it's customers or third parties (and their assets)?
Does it matter who the attacker adversary is (organised crime, nation state attack, nation state proxy, individual, vigilante)?
If so, does the right/obligation extend to self-defence?
If so, does the right/obligation extend to pre-emptive action before the attack?
Does the bank or hospital have a right or obligation to gather intelligence (obtain evidence) on the C&C infrastructure of the attacker?
If not, was an offence committed in obtaining intelligence?
Will the evidence be admissible in a court of law?
What if the bank or hospital attack the C&C infrastructure of the attacker and this results in physical damage to property belonging to the perpetrators?
What if the bank or hospital attacks the C&C infrastructure of the attacker, and in the process takes down the revenue generating online businesses of innocent third parties compromised by the perpetrators in carrying out the attack against the bank or hospital?
What if the bank or hospital attack the C&C infrastructure of the attacker, and in the process causes personal injury or death to the attacker or innocent third party?
What if the bank or hospital attack the critical infrastructure of the attacker, and in the process causes damage or destruction the infrastructure of a foreign state?
What if the bank or hospital contract the services of an offshore organisation to provide defensive/offensive services, including employing targeted malware to cripple C&C infrastructure of an innocent third party?

What if the bank or hospital employ the services of an offshore organisation to provide offensive services including deploying remote exploits to compromise the services/devices of the attacker (i.e. hack the hacker)?
Does the state have a duty or obligation to act?
Where/how does jurisdiction vest?

Definition of Critical Infrastructure - Australian Government Department of Home Affairs

When considering the impact of cyberwar on infrastructure it is useful to refer to specific definitions within national laws. In Australia, critical infrastructure provides services that are essential for everyday life such as energy, food, water, transport, communications, health and banking and finance.[xxvi] Typically critical infrastructures demand higher levels of protection and defence and attacks against them fall strictly under traditional forms of warfare.

Conclusion and Future Work

Scenario 1

Current DDoS attack against online banking systems, exploiting banking client devices. Attacker seeks to obtain client credentials to commit theft. Bank knows location of the C&C server of the attacker.

Issue	Australia	NZ	UK	USA	China	Singapore	India
Self-defence	✓	✓	✓	✓	✓ Qualified: State relationship.	✓	✓
Conspiracy	✓ Qualified: Number of people.	✓	✓	✓	✓	✓	✓
Corporate responsibility	✓	✓	✓	✓	✓ Qualified: State relationship.	✓	✓

Table 1: **National Laws Scenario 1**

Scenario 2

Imminent ransomware attack on hospital system, will exploit vulnerability in third party POS system in hospital. Attacker seeks to extort ransom from hospital.

Issue	Australia	NZ	UK	USA	China	Singapore	India
Self-defence	✓	✓	✓	✓	✓ Qualified: State relationship.	✓	✓
Conspiracy	✓ Qualified: Number of people.	✓	✓	✓	✓	✓	✓
Corporate responsibility	✓	✓	✓	✓	✓ Qualified: State relationship.	✓	✓

Table 2: **National Laws Scenario 2**

Singapore Penal Code

Chapter IV — GENERAL EXCEPTIONS	Right of private defence	96. Nothing is an offence which is done in the exercise of the right of private defence.
	Right of private defence of the body and of property	97. Every person has a right, subject to the restrictions contained in section 99, to defend — (a) his own body, and the body of any other person, against any offence affecting the human body; (b) the property, whether movable or immovable, of himself or of any other person, against any act which is an offence falling under the definition of theft, robbery, mischief or criminal trespass, or which is an attempt to commit theft, robbery, mischief or criminal trespass.
	Acts against which there is no right of private defence	99.(3) There is no right of private defence in cases in which there is time to have recourse to the protection of the public authorities.
	Extent to which the right may be exercised	(4) The right of private defence in no case extends to the inflicting of more harm than it is necessary to inflict for the purpose of defence.
	When such right extends to causing any harm other than death	101. If the offence is not of any of the descriptions enumerated in section 100 (*largely wrt threats of death and severe physical attacks*) physical, the right of private defence of the body does not extend to the voluntary causing of death to the assailant, but does extend, under the restrictions mentioned in section 99, to the voluntary causing to the assailant of any harm other than death. [Indian PC 1860, s. 101].
	Commencement and continuance of the right of private defence of the body	102. The right of private defence of the body commences as soon as a reasonable apprehension of danger to the body arises from an attempt or a threat to commit the offence, though the offence may not have been committed; and it continues as long as such apprehension of danger to the body continues. [Indian PC 1860, s. 102].

	When the right of private defence of property extends to causing death	103. The right of private defence of property extends, under the restrictions mentioned in section 99, to the voluntary causing of death or of any other harm to the wrongdoer, if the offence, the committing of which, or the attempting to commit which, occasions the exercise of the right, is an offence of any of the following descriptions: (a) robbery; (b) house-breaking by night; (c) mischief by fire committed on any building, tent or vessel, which building, tent or vessel is used as a human dwelling, or as a place for the custody of property; (d) theft, mischief or house-trespass, under such circumstances as may reasonably cause apprehension that death or grievous hurt will be the consequence, if such right of private defence is not exercised. [Indian PC 1860, s. 103].
	When such right extends to causing any harm other than death	104. If the offence, the committing of which, or the attempting to commit which, occasions the exercise of the right of private defence, is theft, mischief, or criminal trespass, not of any of the descriptions enumerated in section 103, that right does not extend to the voluntary causing of death, but does extend, subject to the restrictions mentioned in section 99, to the voluntary causing to the wrongdoer of any harm other than death. [Indian PC 1860, s. 104].
	Commencement and continuance of the right of private defence of property	105.—(1) The right of private defence of property commences when a reasonable apprehension of danger to the property commences. (2) The right of private defence of property against theft continues till the offender has effected his retreat with the property, or till the assistance of the public authorities is obtained, or till the property has been recovered. (3) The right of private defence of property against robbery continues as long as the offender causes or attempts to cause to any person death or hurt or wrongful restraint, or as long as the fear

		of instant death or of instant hurt or of instant personal restraint continues. (4) The right of private defence of property against criminal trespass or mischief, continues as long as the offender continues in the commission of criminal trespass or mischief. (5) The right of private defence of property against house-breaking by night continues as long as house-trespass which has been begun by such house-breaking continues. [Indian PC 1860, s. 105].

Table 3: **Singapore Penal Code Right of Private Defence**

Findings

I have worked with international law in the field of cyberwar for well over a decade. I am fascinated at the general resistance of leaders, and I cite specifically those in policy and law, to consider the wealth of human history and knowledge that exists in legal systems across the world. People are fundamentally the same. Our instinct for survival is primal – be it as individual civilians, as corporate citizens or as nation states.
I believe that at this critical time in the history of the world, we need to draw upon the wealth of our survival tactics, by resorting to the law and behaviours that are proven to have worked and agree to adapt and apply them to cyberwar scenarios and to new age societies.

Collectively, the private sector, with the cooperation of governments is positioned to raise the bar against mal-actors.

I invite you to join me in the conversation because there is more than enough evidence in the national laws that I have researched over many years to demonstrate sufficient basis to establish norms for acceptable behaviour at an international level in cyberspace.

We just need to agree to do so – as the merchants did all those years ago.

References

[i] Michael N Schmitt, *The Tallinn Manual* (Cambridge University Press, 2017) examines the international law governing cyber warfare, with cyber warfare is used in a purely descriptive, non-normative sense. Except when explicitly noted otherwise, the rules and commentaries in chapter 1 of the manual apply both in times of peace and in times of armed conflict (whether international or non-international in nature). During an international armed conflict, the law of neutrality also governs the rights and obligations of states in regard to cyber infrastructure and operations.
[ii] Charter of the United Nations.
[iii] Charter of the United Nations art 51. Nothing in the present Charter shall impair the inherent right of individual or collective self-defence if an armed attack occurs against a Member of the United Nations, until the Security Council has taken measures necessary to maintain international peace and security. Measures taken by Members in the exercise of this right of SELF-DEFENCE shall be immediately reported to the Security Council and shall not in any way affect the authority and responsibility of the Security Council under the present Charter to take at any time such action as it deems necessary in order to maintain or restore international peace and security.
[iv] Against a Member and until the Security Council takes measures to maintain international peace and security.
[v] Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I), of 8 June 1977: Part 1 General Provisions Article 1 — General principles and scope of application 1. The High Contracting Parties undertake to respect and to ensure respect for this Protocol in all circumstances. 2. In cases not covered by this Protocol or by other international agreements, civilians and combatants remain under the protection and authority of the principles of international law derived from established custom, from the principles of humanity and from the dictates of public conscience.
[vi] International humanitarian law distinguishes two types of armed conflicts, namely: international armed conflicts, opposing two or more States, and non-international armed conflicts, between governmental forces and non-governmental armed groups, or between such groups only. IHL treaty law also establishes a distinction between non-international armed conflicts in the meaning of common Article 3 of the Geneva Conventions of 1949 and non-international armed conflicts falling within the definition provided in Art. 1 of Additional Protocol II.
[vii] Joel R. Reidenberg, 'Lex Informatica: The formulation of Information Policy Rules Through Technology' (1998) Volume 76, Number 3, *Texas Law Review*.
[viii] Ibid.
[ix] United Nations Commission of International Trade Law <http://www.uncitral.org/>.
[x] The UNCITRAL Model law on Electronic Commerce adopted in June 1996, the Model Law on Electronic Signatures adopted in July 2001, and The Convention on the Use of Electronic Communications adopted in November 2005.
[xi] And Australia's promulgation of its first *Electronic Transactions Act 1999 (Cth).*
[xii]Department of Foreign Affairs and Trade. Available at https://dfat.gov.au/international-relations/themes/cyber-affairs/aices/pdf/DFAT%20AICES_AccPDF.pdf.
[xiii] Including: New Australian Government Data Sharing and Release Legislation Issues Paper; The Assistance and Access Bill 2018; Human Rights and Technology Issues Paper July 2018; My Health Records and My Health Records Amendment (Strengthening Privacy) Bill 2018; Treasury Laws Amendment (Consumer Data Right) Bill 2018.
[xiv] NATO. Available at https://www.nato.int/docu/review/2015/also-in-2015/hybrid-modern-future-warfare-russia-ukraine/EN/index.htm.
[xv] A malicious computer worm allegedly built by the USA and Israel which targeted the SCADA systems of the Iranian nuclear program.
[xvi] I refer to the allegations made in respect to HBGary and the National Intelligence Agency.
[xvii] I refer to tactics employed by President George W. Bush in Iraq, the 1999 Pentagon general counsel office guidelines for waging cyberwar, Stuxnet, the alleged North Korea attack on Sony Pictures, revelations made by Edward Snowden, and Kim Zetter in her book *Countdown to Zero Day: Stuxnet and the Launch of the World's First Digital Weapon*. Crown Publishing Group New York, NY, USA. 2014.
[xviii] I refer to allegations in relation to Russia and Fancybear.
[xix] A contention supported by criminal services offered on the darknet.
[xx] Zetter. Above n.
[xxi] I refer to the alleged Russian influence on the USA Elections.
[xxii] I refer to allegations made by the USA against China.

[xxiii] I refer to allegations that Cambridge Analytica used Facebook to influence the USA elections.
[xxiv] George Washington University Center for Cyber and Homeland security, “Into the Gray Zone,” (2016). Available at <https://cchs.gwu.edu/gray-zone-active-defense-private-sector-against-cyber-threats https://creativecommons.org/licenses/by/4.0/>.
[xxv] A comprehensive analysis of how existing international law applies to cyber operations. Available at https://bit.ly/2Rjag6h.
[xxvi] Available at https://www.homeaffairs.gov.au/about/national-security/critical-infrastructure-resilience).