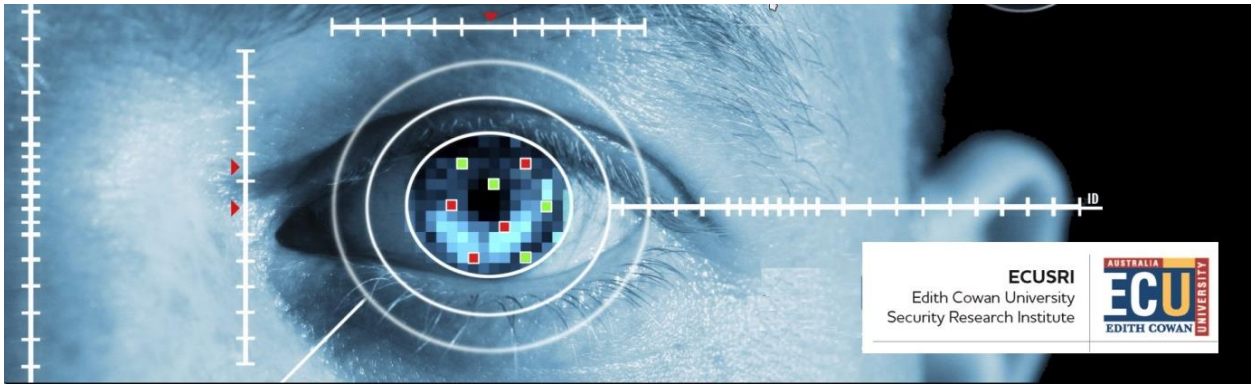
2017

The Proceedings of 15th Australian Information Security Management Conference, 5-6 December, 2017, Edith Cowan University, Perth, Australia

Craig Valli (Ed.)
Edith Cowan University

DOI: 10.4225/75/5a84fe5695b55



Proceedings of the

15th Australian Information Security Management Conference

5-6 December 2017

Edith Cowan University, Perth, Australia

AUSTRALIA ECU EDITH COWAN UNIVERSITY

Proceedings of the
15th Australian Information Security Management Conference

Published By
Security Research Institute
Edith Cowan University

Edited By
Professor Craig Valli
Security Research Institute
Edith Cowan University

ISBN 978-0-6481270-8-6

CRICOS Institution Provider Code 00279B

Sponsors

ECUSRI
Edith Cowan University
Security Research Institute

AUSTRALIA ECU EDITH COWAN UNIVERSITY

SC8

CISCO

Supporters

AUSTRALIAN SECURITY MAGAZINE

acs AUSTRALIAN COMPUTER SOCIETY

Australian and New Zealand FORENSIC SCIENCE SOCIETY

Conference Foreword

The annual Security Congress, run by the Security Research Institute at Edith Cowan University, includes the Australian Information Security and Management Conference. Now in its fifteenth year, the conference remains popular for its diverse content and mixture of technical research and discussion papers. The area of information security and management continues to be varied, as is reflected by the wide variety of subject matter covered by the papers this year. The papers cover topics from vulnerabilities in "Internet of Things" protocols through to improvements in biometric identification algorithms and surveillance camera weaknesses.

The conference has drawn interest and papers from within Australia and internationally. All submitted papers were subject to a double blind peer review process. Twenty two papers were submitted from Australia and overseas, of which eighteen were accepted for final presentation and publication.

We wish to thank the reviewers for kindly volunteering their time and expertise in support of this event. We would also like to thank the conference committee who have organised yet another successful congress. Events such as this are impossible without the tireless efforts of such people in reviewing and editing the conference papers, and assisting with the planning, organisation and execution of the conference.

To our sponsors, also a vote of thanks for both the financial and moral support provided to the conference. Finally, thank you to the administrative and technical staff, and students of the ECU Security Research Institute for their contributions to the running of the conference.

Yours sincerely

Conference Chair
Associate Professor Mike Johnstone, Security Research Institute, Edith Cowan University

Congress Organising Committee

Congress Chair:

Professor Craig Valli

Committee Members:

Professor Gary Kessler – Embry Riddle University, Florida, USA
Professor Glenn Dardick – Embry Riddle University, Florida, USA
Professor Ali Babar – University of Adelaide, Australia
Dr Jason Smith – CERT Australia, Australia
Associate Professor Mike Johnstone – Edith Cowan University, Australia
Professor Joseph A. Cannataci – University of Malta, Malta
Professor Nathan Clarke – University of Plymouth, Plymouth UK
Professor Steven Furnell – University of Plymouth, Plymouth UK
Professor Bill Hutchinson – Edith Cowan University, Perth, Australia
Professor Andrew Jones – Khalifa University, Abu Dhabi, UAE
Professor Iain Sutherland – Glamorgan University, Wales, UK
Professor Matthew Warren – Deakin University, Melbourne, Australia

Congress Coordinator:

Ms Emma Burke

Table of Contents

FINANCIAL FRAUD RISK MANAGEMENT AND CORPORATE GOVERNANCE..... **5**
Raymond Lutui, Tau'aho 'Ahokovi

SECURITY READINESS EVALUATION FRAMEWORK FOR TONGA E-GOVERNMENT INITIATIVES.... **14**
Raymond Lutui, Semisi Hopoi, Siaosi Maeakafa

EVALUATING IP SURVEILANCE CAMERA VULNERABILITIES..... **25**
Brian Cusack, Zhuang Tian

TONGA'S ORGANISATIONAL VULNERABILITY TO SOCIAL ENGINEERING..... **33**
Raymond Lutui, Viliami Fe'aomoeata

ASSESSMENT OF SECURITY VULNERABILITIES IN WEARABLE DEVICES..... **42**
Brian Cusack, Bryce Antony, Gerard Ward, Shaunak Mody

NEUROSECURITY FOR BRAINWARE DEVICES..... **49**
Brian Cusack, Kaushik Sundararajan, Reza Khaleghparast

INTELLIGENT FEATURE SELECTION FOR DETECTING HTTP/2 DENIAL OF SERVICE ATTACKS..... **57**
Erwin Adi, Zubair Baig

A SRI LANKAN HACKING CASE STUDY..... **64**
Ishan Senarathna, Matthew Warren

SECURITY VULNERABILITIES AND CYBER THREAT ANALYSIS OF THE AMQP PROTOCOL FOR THE INTERNET OF THINGS..... **68**
Ian Noel McAteer, Muhammad Imran Malik, Zubair Baig, Peter Hannay

A CRITICAL ANALYSIS OF SECURITY VULNERABILITIES AND COUNTERMEASURES IN A SMART SHIP SYSTEM..... **81**
Dennis Bothur, Guanglou Zheng, Craig Valli

THE 2017 HOMOGRAPH BROWSER ATTACK MITIGATION SURVEY..... **88**
Tyson McElroy, Peter Hannay, Greg Baatard

A REVIEW OF DATA BREACHES AND LOSSES THAT OCCURRED FROM LAPTOPS THAT WERE STOLEN OR OTHERWISE MISPLACED IN 2015 AND 2016..... **97**
Samuel Griffith Wakeling, Peter Hannay, Zubair Baig

A COMPARISON OF 2D AND 3D DELAUNAY TRIANGULATIONS FOR FINGERPRINT AUTHENTICATION..... **108**
Marcelo Jose Macedo, Wencheng Yang, Guanglou Zheng, Michael N Johnstone

LITERATURE-BASED ANALYSIS OF THE INFLUENCES OF THE NEW FORCES ON ISMS: A CONCEPTUAL FRAMEWORK…… **116**
Zahir Al-Rashdi, Dr Martin Dick, Dr Ian Storey

CORE ELEMENTS IN INFORMATION SECURITY ACCOUNTABILITY IN THE CLOUD…… **125**
Zahir Al-Rashdi, Dr Martin Dick, Dr Ian Storey

AN INVESTIGATION INTO SOME SECURITY ISSUES IN THE DDS MESSAGING PROTOCOL……. **132**
Thomas White, Michael N. Johnstone, Matthew Peacock

DECEPTIVE SECURITY BASED ON AUTHENTICATION PROFILING…… **140**
Andrew Nicholson, Helge Janicke, Andrew Jones, Adeeb Alnajaar

THE CONVERGENCE OF IT AND OT IN CRITICAL INFRASTRUCTURE…… **149**
Glenn Murray, Michael N. Johnstone and Craig Valli

FINANCIAL FRAUD RISK MANAGEMENT AND CORPORATE GOVERNANCE

Raymond Lutui[1], Tau'aho 'Ahokovi[2]
Auckland University of Technology[1], Christ's University in Pacific[2]
School of Engineering, Computer & Mathematical Sciences[1], School of Business & Law[2]
rlutui@aut.ac.nz[1], tahokovi@bigpond.com

Abstract

Risk management is important so that risk is assessed, understood and appropriately managed. This is important both for conformance and performance. It is essential that strategic planning and management decisions are made appropriately in the context of the risk appetite of the corporation and its various stakeholders – especially its shareholders. If a company does not have a good understanding of risk, the likelihood of conformance and performance failure is high, this implies good internal and external corporate intelligence. Large global corporations have a significant impact on economies around the world. These entities are subject to intense competition and require investor and customer confidence to underpin their activities. Poor governance adversely affects customers and investors, and makes corporation uncompetitive. This can also affect entire economies. In the context of the Global Financial Crisis (GFC), the collapse of the US investment bank Lehman brothers demonstrates that corporate failure can hurt economies globally. The failure of Lehman Brothers to properly manage and understand risk is a clear example of the failure of good governance.

Keywords*:* Fraud Risk Management, Corporate Governance, Good Governance, Fraud,

INTRODUCTION

The upsurge of financial scandals in the era of the 21st century raised awareness of deep-seated fraudulent activities (Kerr and Murthy 2013). Financial statement fraud has cast an increasingly adverse impact on the individual investors and the stability of global economies (Zhou and Kapoor 2011). The failure of Enron has caused about a $70 billion lost in the capital market. The Computer Security Institute reported a significant increase in financial fraud cases recently (Reddy et al. 2012). The rise of many fraudulent occurrences is a serious inhibitor for potential investors because fraudulent financial reports have created a substantial negative impact on company reputations and market value (Hogan et al., 2008).

Financial statements are basic documents to reflect a company's financial status (Beaver 1966). Fraudulent financial reports are perpetrated to increase stock prices or to get loans from banks (Ravisankar et al., 2011). Financial statement fraud detection is vital because of the devastating consequences of financial statement frauds (Ngai et al. 2011). Fraud behaviours are often subtle in the beginning (Chivers et al., 2013), therefore, it is difficult to detect them. Regulations play an important role to emphasize the responsibility of auditors to assess the risk of fraudulent financial reporting adequately (Srivastava et al. 2009). However, detecting frauds remains difficult because of the lack of a commonly accepted definition of reasonable assurance, limitations of audit methods and the cost constraints (Spathis, 2002; Hogan et al., 2008).

The board of directors is the body that oversees the activities of an organisation. The board has a wide range of roles and functions that address both performance and conformance. It is preferable that the roles and responsibilities of the board be explicitly set out in a written chatter or constitution. The board must ensure appropriate procedures are in place for risk management and internal controls, and it must also ensure that it is informed of anything untoward or inappropriate in the operation of those procedures. Any major operation issues will also be brought to the attention of the board for appropriate consideration and decision making.

Despite these expectations, in many high-profile corporate collapses it is apparent that the board was informed about key business decisions or simply chose to comply with management. For example, in the case of a former prominent Australian company, HIH Insurance, it was apparent that the major takeover of another company, FAI

Insurance, was undertaken without rigorous debate at board level or due diligence being carried out before the transaction was finalised (CPA, 2016).

Companies that can demonstrate good corporate governance practices have advantages. With the increasing globalisation of business and competition for capital, companies that can provided assurances that the company is being appropriately managed the cost of capital. Furthermore, the expansion of company shareholdings to a broader base (in many countries, small shareholders are becoming increasingly common, either by direct investment or indirectly through their superannuation plans), combined with more organised and active shareholders lobby groups, is placing more scrutiny on company management (Drever et al., 2007, p. 153). In this paper, a review of the risk management and corporate governance principles is made in order to identify weaknesses in corporate governance and also to identify how to improve. It then proposes a focus on improvement for risk management and the corporate governance benefit.

RISK MANAGEMENT

Risk management is defined as the "process of understanding and managing risks that the entity is inevitably subject to in attempting to achieve its corporate objectives (CIMA, 2005). For an organisation risks are potential events that could influence the achievement of the organisation's objectives. Risk management is about understanding the nature of such events and where they represent threats, making positive plans to mitigate them. Fraud is a major risk that threatens the business, not only in terms of financial health but also its image and reputation.

Risk management is also an increasing important process in many businesses and the process fits in well with the precepts of good corporate governance. In recent years, the issue of corporate governance has been a major area for concern in many countries. In the UK, the first corporate governance report and code of best practice is considered to be the Cadbury Report in 1992, which was produced in response to a string of corporate collapses. There have been a number of reports since, covering provisions around area such as executive remuneration, non-executive directors, and audit committees. The principles of these various report have been brought together to form the Combined Code on Corporate Governance (Combined Code).

The Need for Governance

Governance describes the overall guidance of organisations and focuses on achieving strong performance while ensuring compliance with obligations. Effective governance is very important and poor governance has often led financial disasters for individual companies, and even whole economies. Governance is the system by which companies are directed and controlled, and accountability is assured. While the concept is usually associated with corporate governance, that is the governance of large listed corporations, similar governance principles should apply to all enterprises. Governance relates to the responsibilities of the board of directors towards investors and other stakeholders, and involves setting the objectives and direction of the company and is distinguished from the management of the enterprise on a daily basis, which is the job of full-time executives.

The governance of enterprises is broadly structured by the law, not just corporate law but also employment law and so forth. It is the first duty of directors to ensure that the enterprise operates with in the law. However, beyond requiring a board of director's exercise certain duties such as the duty of care and diligence, corporation law gives considerable scope for directors to exercise decision-making in the best interests of the company. It is here where the skills of governance become critical: the capacity to understand and interpret the strengths and weaknesses of the enterprise, and how to direct the enterprise towards business success while maintaining accountability and good relationships with all stakeholders. Good governance is a hallmark of enterprises that achieve improving and sustainable performance even in changing and unpredictable environments.

Good governance aims to ensure that organisations are properly run in the best interests of their shareholders, including the optimal performance of national and international economies. At an organisational level, the behavioural styles and business management practices of managers (and other employees) or directors can result in outcomes that are not in the best interests of shareholders and other stakeholders. These situations can range from relatively minor technical breaches of policies or practices, to more serious cases where excessive risk-taking or poor controls place the ongoing survival of the organisation at risk (CPA, 2016)

Many other countries have also produced reports on corporate governance, usually accompanied by codes of best practices. For example, South Africa has had the King Report (version I and now II) since 1994, Malaysia has had its Code of Corporate Governance in place since 2000 and Sri Lanka issued the Rules on Corporate

Governance as part of its Listing Rules in January 2007 (CPA, 2016)

Corporate governance requirements in the US are now largely set out within the Sarbox legislation, as previously mentioned (US Congress 2002, Sarbanes-Oxley Act 2002); these requirements extend beyond the US, capturing any company that is SEC listed and its subsidiaries. Some other countries have also introduced a statutory approach to corporate governance, such as that in the US, although none are currently as comprehensive. A number of international organisations have also launched guidelines and initiatives on corporate governance, including the Organisation for Economic Co-operation and Development (OECD) and the European Commission.

In extreme cases, public organisations may be run more as personal fiefdoms where personal greed is put ahead of the interests of shareholders and other stakeholders. To reduce undesirable consequences for shareholders and other stakeholders and to ensure personal accountability, organisations need an appropriate system of checks and balances in the form of corporate governance framework. This framework emphasises both conformation and performance as vital elements of the way the companies are run.

The Role of the Board

As corporations grow in size, there is also a separation of the ownership and management. Over time, the legal duties and responsibilities of directors have evolved to protect the interests of the owners, who are not able to observe closely the daily occurrences within a corporation. In most jurisdictions, there is a core group of director's duties and responsibilities that have arisen from either statue or case law. The key duties are to:

- Avoid conflict of interest and where these exist, ensure they are appropriately declared and as required by law, otherwise mange correctly
- Act in best interests of the corporation
- Exercise powers with proper purposes
- Retain discretionary powers and avoid delegating the director's responsibility
- Act with care, skill and diligence
- Be informed about the corporation's operations, and
- Prevent insolvent trading

Consequently, the board of directors should have implemented a strategy settings design to identify potential events that may affect the entity (Gelinas, Dull & Wheeler, 2012). These strategy settings reflect in a framework which is called "Enterprise Risk Management" (ERM).

In formal corporate governance principles, managers are the agents of the board responsible for pursuing the vision of the company as developed by the board, and fulfilling the strategic direction determined by the board. The CEO in most companies is also a director and a member of the board (and there are often other executive directors such as the CFO of the company). These executives' directors have a full role working with the board to advance strategic direction and establish the policies and value of the company. Once these are decided, it is the manager's duty to actively pursue these, and the board's role is to monitor the results for the business.

Of course, in reality the interface of governance and management is more complex. Often boards and management respect and understand the different roles and have a commitment to make the relationship work. However, sometimes tensions do emerge, for example, in the choice of strategy. Because of rapidly changing markets and technology, boards often have to be continuously engaged in strategic decisions, unlike in the past. At times, managers may feel that the board is becoming too involved in the implementation of strategy when it is the management team who have operational experience required to guide strategy to success. On other occasions, the board may feel that managers are making significant strategic decisions without properly securing the approval of the board (CPA, 2016).

Skeet (2015) examines this issue from the perspective of both the board of directors and the management team. When CEOs are asked what issues contribute to the board and management being at cross purposes, they point to two main factors: directors acting 'out of position' and attempting to play a management role; or a conflict of interest where, even if disclosed, directors are not able to place the interests of the organisation above their own or those of the group they are representing.

Often what boards interpret as arrogance of the CEO and the management team can be, in reality, a lack of experience, strategic direction differences or deceit. These can all lead to the management team withholding information from the board. Board members should consider what information they do not currently have and then request this additional information if they feel the CEO and the management team may be concealing

something. This is a legal right of the board, and the management team is not permitted to suppress this information, once requested. The board is able to draw on multiple points of view when making decisions, which is strength of shared governance (Skeet, 2015). For example; there was a tension occurred some years ago at BHP Billiton when a newly appointed CEO began negotiating for major acquisitions without fully consulting the board. The board became concerned about the serious risk implication of the CEO's actions, and the contract of the CEO was terminated. With the appointment of another CEO, the BHP board was careful to agree on a series of protocols regarding the scope for independent decision-making by the CEO on financial and other matters, and the issues that always needed to be brought to the board for consideration. These protocols appear to have worked well, and in other large corporations, similar, clear understanding exists between board and executive management on their respective roles and powers (CPA, 2016).

This tension, occurred some years ago at BHP Billiton when a newly appointed CEO began negotiating for major acquisitions without fully consulting the board. The Board concerned about the serious risk implication of the CEO's action, and the contract of the CEO was terminated. With the appointment of another CEO, the BHP board was careful to agree of protocols regarding the scope for independent decision-making by the CEO on the financial and other matters, and the issues that always needed to be to the board for consideration. These protocols appear to have worked well, and in other corporations, similar, clear understanding exist between board and executive management on their respective roles and powers.

It is certainly the case that it is management at the sharp end of delivering the aspirations of the board for the company. Boards of directors are often highly skilled at financial analysis, strategic thinking and policy development, but it is the managers who have to implement all of these, which requires considerable intellectual, operational and intellectual skills. It is the management who must inspire employees with the goals of the enterprise, delight customers with the quality of the product or service, convince suppliers and distributors that the company deserves their full support, and keep stakeholders onside (CPA, 2016).

Ensuring that there is the energetic commitment of managers to their task of realising the vision of the board and making the success of the company is ultimately the role of the CEO, who is the essential link between the governance mechanisms and the operational mechanisms of the company.

INTERNATIONAL PERSPECTIVES ON CORPORATE GOVERNANCE

Globalisation has caused major changes in the way incorporations are run. Inevitable changes in the size and the structure of companies, including their ownership structures, have had a substantial effect on the way corporations are controlled. For example, many traditional Australian companies, some listed on ASX, are now effectively controlled by owners in diverse locations such as the United States, China, Singapore, India, the United Kingdom and Germany. These owners are subject to governance standard that differ from those in Australia. Even so, listing in Australia means that they must comply with Australian governance standards in addition to those of their own country (CPA, 2016)

United Kingdom

In 1991, following a series of high profile corporate collapses, the London Stock Exchange, together with industry and accounting and finance professionals, established the Cadbury Committee. The Cadbury report, *Financial Aspects of Corporate Governance* (CFACG, 1992), gave recommendations to companies that have been adopted in varying degrees by the European Union, the United States, the World Bank and many other countries and regions. The recommendations on governance had an important feature that is still used today – the concept of 'comply or explain'. This approach meant that if a company chose not to comply with a governance recommendation, the company had to identify the non-compliance and then explain it to shareholders. This may also be described as 'if not, why not' reporting.

United States

The Committee of Sponsoring Organisation of the Treadway Commission (COSO) was formed in 1985 to sponsor the National Commission on Fraudulent Financial Reporting. Its 1994 report, Internal Control-Integrated Framework (COSO, 1994), provided a detailed definition and discussion of internal control. In 1999, it reported on fraudulent financial reporting (COSO, 1999). Important findings included the frequent involvement of the CEO and CFO in frauds, captured boards that were dominated by insiders, and unqualified opinions by auditors despite the fraud.

In response to a loss of investor confidence following corporate scandals in the United States, the US Congress passed the Sarbanes-Oxley Act in 2002. The purpose of the Act was to protect investors and provide guidelines for financial reporting.

Australia

The Ramsay Report (Ramsay 2001) examined the adequacy of Australian legislative and professional requirements regarding the independence of external auditors and made recommendations for changes. Some parts of the report were concerned directly with audit independence and others were designed generally to enhance audit independence; for example; establishing audit committees and board to oversee audit independence issues. In 2002, the Australian Stock Exchange (since renamed Australian Securities Exchange) responded to calls for it to play a greater role in corporate governance through the establishment of the Corporate Governance Council. The Council released the first edition of its *Principles of Good Corporate Governance and Best Practice Recommendations* (ASX CGC 2003). These were revised in 2007 and titled Corporate Governance Principles and Recommendations. The 2007 revision was amended in 2010. The third edition was released in 2014.

Again, the Australian government released a discussion paper (CLERP, 2004) in the aftermath of the collapses of, among others, Enron in the United States and HIH Insurance in Australia. This paper known as Corporate Law Economic Reform Program (CLERP) outlined proposals for audit and financial reporting reform, as well as other legislative proposals, to improve corporate governance practices in Australian companies. This report was passed by the Australian Government, coming into effect on 1 July 2004. There are many other international organisations that focus on improved corporate governance. Many of them, such as the Business Roundtable, an association of chief executives of leading US companies, and The International Corporate Governance Network (ICGN), a not-for-profit body founded in 1995, have produced their own recommended codes and guidelines.

OECD PRINCIPLES OF CORPORATE GOVERNANCE

The OECD (Organisation for Economic Co-operation and Development), with members and funding sources from countries with major market-oriented economies, has developed international best practice principles of governance which was first published in 1999 (OECD, 1999) and were updated in 2004 (OECD, 2004) with a new first principle giving a broad view of governance including performance. A review of these principles started in 2014 and following an extensive consultation, the updated principles were released in September 2015, entitled G20/OECD Principles of Corporate Governance (QECD, 2015). These principles were considered as the international best practice by referring to specific guidance, codes and recommendations on corporate governance produced by the OECD and the Financial Reporting Council of the United Kingdom, who were become global leaders in the development of corporate governance principles. It also considers the ASX Principles, as they also provide leadership in corporate governance.

The OECD Principles specify six principles:

- Ensuring the basis for an effective corporate governance framework
- The rights and equitable treatment of shareholders and key ownership functions
- Institutional investors, stock markets, and other intermediaries
- The role of stakeholders in corporate governance
- Disclosure and transparency; and
- The responsibilities of the board

DISCUSSION

Although corporate governance is usually linked to management, there is a strong bond between corporate governance and ethics and/or social responsibility of the business. Corporate governance encourages a trustworthy, moral, as well as ethical environment. From this point of view governance takes into account the transparency of the internal and external audit, the sincerity of the managers regarding the company's financial results and financial statements, the manager actions towards the small stakeholders and many more (Panfilli 2012). Organization for Economic Co-operation and Development (OECD) considers that corporate governance has the role to specify the distribution of rights and of responsibilities between different categories of people involved in the company like: board of directors, executives, shareholders and others, establishing rules and procedures for making decisions on the activity of a certain company. OECD also mentions that corporate governance is at the same time, both a set of relations between management, board of directors, shareholders and

other interested groups and the structure through which company sets the objectives and the necessary means to reach those, but also the system of incentives offered to the board of directors and management in order to increase the objectives in the interests of shareholders and society.

Poor ethical leadership, lack of personal integrity, mismanagement, fraud, corruption, and violation of corporate governance rules are the main contributors towards bankruptcy and financial failures in large organizations. Most of these organizations have comprehensive corporate governance codes in place, implemented by the left brain Big Four accountancy firms (PwC, KPMG, Ernst & Young and Deloitte), McKinsey, America's Top Corporate Governance Law Firms, which apparently are not working at all. They made things worse and created a stable basis for more corruption (Ramperad & Fawumi, 2015). Current approaches to corporate governance are extremely formal, bureaucratic, cosmetic, not holistic and non-authentic, and therefore provide no protection from potentially catastrophic ethical failures. A sustainable and innovative solution to this global epidemic is needed urgently. It is time that awareness that corporate governance cannot be controlled effectively with formal and exhaustive rules, regulations, guidelines, and procedures only. It is about decency and personal integrity and this must be cultivated from within. Personal integrity has no need of rules and laws.

It must be a way of life: To quote what Plato said in 340 BC: "Good people do not need laws to tell them to act responsibly, while bad people will find a way around the laws". Research shows that a large percentage of the world's population is bad (Rampersad & Fawumi, 2015) For example, America has around 5% of the world's population, and 25% of its prisoners. Roughly one in every 107 American adults is behind bars. Among them are also many Executives, leaders and professionals. Most corporate governance programs make things worse by creating a stable basis for more corruption and are doomed to fail. Why is the lesson from Plato not learned and focus on creating a culture of good people, in which personal values are aligned with the laws and embedded in the mind of the people, instead of focusing on laws (corporate governance) only? An innovative methodology was launched for creating a culture of good Chairman's, Presidents, CEOs, CFOs, managers and employees, in which high ethical values are aligned with their corporate governance rules, regulations and guidelines and embedded in their mind (Ramperad & Fawumi, 2015).

Many law, accounting and business management professors at the US top schools are blamed for most of the corporate governance failures. They lack both emotional and spiritual intelligence. This inner process starts with self-knowledge, or knowing, which leads to wisdom. Between knowing and wisdom lies an enormous distance which can be reduced by systematic application of the authentic governance system. This will help them to create balance between the left and right sides of their brain. The left half of their brain has mainly an analytical, logical and quantitative function, while the right half of their brain has an intuitive and holistic function. They do not have a proper balance between the left and right sides of their brain. These professors and most of their graduates use the left side of their brain only; because of this, they miss opportunities that allow them to become more adept at using the right hemisphere of the brain and to deal with complex corporate governance problems in an integrated and authentic way. This is also the main reason why Harvard Business School professor Kaplan's balanced scorecard implementations fail and lack sustainability (Ramperad & Fawumi, 2015).
Many public company shareholders have been unpleasantly surprised by major accounting charges resulting from previously undisclosed enterprise risks. These charges typically come without warning in prior audited financial statements. Public shareholders have a right to wonder why they have not received warnings of these risks in prior audited financial statements and question the effectiveness of the board of directors in performing its oversight role (Lipman, 2012). Most directors of public companies focus on the tone at the top of the organization. However, these same directors do not necessarily know whether that "tone" reaches throughout the organization and may fail to assess the culture of the organization. Independent directors cannot adequately perform their oversight role without receiving enterprise risk information from lower-level employees.

Independent directors who rely on independent auditors to disclose unasserted claims arising out of enterprise risks should be aware of the very limited duties of independent auditors in investigating such risks. Typically, this issue is handled by the independent auditors by having top management of the company represent in writing to the independent auditor that they are not aware of unasserted material contingent liabilities and having the company's counsel agree to advise management of any such contingent liabilities. However, top management may not themselves be aware of such unasserted enterprise risks because of silos within the organization or because of the reluctance of employees to use hotlines. Even if management knows of possible contingent liabilities they might incorrectly determine that these liabilities are immaterial. Independent auditors currently are not required to check on the effectiveness of these employee hotlines or the employee culture, including the willingness of employees to use hotlines to report enterprise risk (Lipman, 2012**).**

RECOMMENDATION

After considering risk management and corporate governance principles employed by a different country, it is clear that all the policies and principles adopted were mainly focussing on how to avoid and minimise risk and also to maintain good corporate governance. It can also take into account the self-interest characteristics of individual. This study recommends various contributions in order to improve and effectively enforce the principles and policies stated.

1. Establishing of a Forensic investigation team
Whenever an allegation arises, it is recommended that the Forensic Investigation Team head by a qualified Forensic Accountant should handle the investigation. This is to ensure the board and the management are not acting in any favour or bias.

2. Formulate a response

- The objectives of the investigation should be clearly identified along with resources required, the scope of the investigation and the timescale.
- The objectives will be driven by the organisation's attitude to fraud and the preferred outcome for dealing with fraud.
- An action plan should be prepared and roles and responsibilities should be delegated in accordance with the skills and experience of the individuals involved.
- The individual in overall control of the investigation should be clearly identified, as should the powers available to team members.
- Reporting procedures as well as protocols for handling and recording evidence should be clearly understood by everybody.

3. Follow up action

- There are lessons to be learned from every identified incident of fraud.
- The organisation's willingness to learn from experience is as important as any other response.
- Large organisations may consider establishing a special review to examine the fraud with a view to recommending improvements to systems and procedures.
- Smaller organisations may consider discussing the issues with some of its more experienced people, with the same objectives in mind.
- It is important that recommended changes are implemented promptly.

CONCLUSION

The secret of a successful company is the ability of its board and senior management to assess its principles and policies in order to make decisions that achieve the correct balance over time. While the best corporations do this well, poorer corporations do it less effectively and those that do it worst almost inevitably cease to exist. The many rules and expectations confronting corporations, along with the relationships must be understood and managed. It has been identified that both conformance and performance are central components of corporate governance. Both aspects of corporate governance must be satisfied so that diverse international societies achieve effective utilization of the capital resources employed in their enterprises. The United Kingdom is one of the world's most important investment locations. This is due to the fact that their corporate governance practices was deemed to be the best practice which other nations like New Zealand and Australia were willing to follow and incorporate it to their governance practice. The rules relating to investment in and through the London Stock Exchange have provided leading-edge practical approaches that have been followed successfully in many jurisdictions. Without vigilance, good governance is often forgotten in strong economic times, only to be remembered when financial trouble arises.

REFERENCES

ASX Corporate Governance Council (ASX CGC) 2014, Corporate Governance Principles and Recommendations, 3rd edn, Australian Securities Exchange, Sydney, accessed 12 September 2017, http://www.asx.com.au/documents/asx-compliance/cgc-principles-and-recommendations-3rd-edn.pdf

Australian Securities Exchange (ASX) 2014, *ASX Listing Rules, Chapter 1* 'Admission', Sydney, accessed 10 September 2017, http://www.asx.com.au/regulation/rules/asx-listing-rules.htm.

Beaver, W. H. (1966). Financial ratios as predictors of failure. *Journal of Accounting* Research, 4, 71–111.

Cadbury, S. A. (2000). *Family firms and their governance: Creating tomorrow's company from today's* (p. 5). London: Egon Zehnder International.; hhttp://www.ecgi.org/codes/documents/Cadbury.pdf .

Chivers, H., Clark, J. A., Nobles, P., Shaikh, S. A., & Chen, H. (2013). Knowing who to watch: Identifying attackers whose actions are hidden within false alarms and background noise. *Information Systems Frontiers*, *15*(1), 17-34.

Committee of Corporate Governance (CGC) 1998, *Final Report (Hample Report*), Gee Publishing, London, accessed 08 September 1017

Committee of Sponsoring Organisations of the Tredway Commission (COSO) 1994, *Internal Control-Integrated Framework*, American Institute of Certified Public Accountants, New York.

Committee of Sponsoring Organisations of the Tredway Commission (COSO) 1999, *Fraudulent Financial Reporting* 1987-197; *An Analysis of US Public Companies*, accessed 05 September 1017, http://www,coso.org./publications/ffr_1987_1997.pdf.

Corporate Law Economic Reform Program (*CLERP) 2004*), *Corporate Law Economic Reform Program (Audit Reform and Corporate Disclosure) Act 2004* (CLERP 9), Australian Federal Parliament, Canberra

CPA Program, *Ethics and Governance*, 2016, published by Deakin University, Geelong, Victoria, Australia

Drever, M, Stanton, P & McGowan, S. 2007 Contemporary Issues in Accounting, *John Wiley & Sons Australia, Ltd.*

Financial Reporting Council (FRC) 2003, *The Combine Code on Corporate Governance*, London.

Financial Reporting Council (FRC) 2012, *The UK Stewardship Code* (FRC Code), accessed 02 September 2017, http://www.frc.org.uk/Our-Work/Codes-Standard/Corporate-Governance.aspx.

Financial Reporting Council (FRC) 2014, *The UK Corporate Governance Code* (FRC Code), accessed 04 September 2017, http://www.frc.org.uk/Our-Work/Codes-Standard/Corporate-Governance.aspx.

Hogan, C. E., Rezaee, Z., Riley, R. A., & Velury, U. K. (2008). Financial statement fraud: insights from the academic literature. Auditing: *A Journal of Practice & Theory*, 27(2), 231–252.

Kerr, D. S., & Murthy, U. S. (2013). The importance of the CobiT frame- work IT processes for effective internal control over financial reporting in organizations: an international survey. *Information & Management*, 50(7), 590–597.

Knowing who to watch: identifying attackers whose actions are hidden within false alarms and background noise. *Information Systems Frontiers*, 15(1), 17–34.

Lipman, F, D, (2012), *Why corporate governance failures continue*, accessed 09 September 2017, http://tcbblogs.org/governance/2017/04/17/why-corporate-governance-still-continues-to-fail/

Ngai, E. W. T. (2003). Selection of web sites for online advertising using the AHP. Information and Management, 40(4), 233–242.

Ngai E, Hu Y, Wong Y, Chen Y, Sun X. The Application of Data Mining Techniques in Financial Fraud Detection: *A Classification Framework and an Academic Review of Literature. Decision Support Systems 2011*; 50: 559–569. Accessed 10 September 2017.

Organisation for Economic Co-operation and Development (OECD) 1997, OECD *Economics Survey: France*, Paris

Organisation for Economic Co-operation and Development (OECD) 1999, OECD *Principles of Corporate Governance; Paris*

Organisation for Economic Co-operation and Development (OECD) 2003, *White Paper on Corporate Governance in Asia*, Paris

Organisation for Economic Co-operation and Development (OECD) 2004, OECD *Principles of Corporate Governance* ("OECD Principles"), accessed 11 September 2017, http://www.oecd.org/corporate/ca/corporategovernanceprinciples/31557724.pdf.

Organisation for Economic Co-operation and Development (OECD) 2010, *Corporate Governance and the Financial Crisis: Conclusions and Emerging Good Practices to Enhance Implementation of the Principles*, accessed 12 September 2017,

http://www.oecd.org/daf/ca/corporategovernanceprinciples/44679170.pdf

Organisation for Economic Co-operation and Development (OECD) 2011, *Reform Priorities in Asia: Taking Corporate Governance to a Higher Level*, Paris, accessed 11 September 2017, http://www.oecd.org/corporate/ca/49801431.pdf.

Organisation for Economic Co-operation and Development (OECD) 2015, G20/*OECD Principles of Corporate Governance* ("OECD Principles"), accessed 13 September 2017, http://www.oecd.org/corporate/principles-corporate-governance.htm.

Panfilii, A 2012, *Failure of corporate governance – intention or negligence*, accesses 13 September 2017, http://www.actionamresponsabil.ro/failure-of-corporate-governance-intention-or-negligence/17037

Rampersad, H & Fawumi, A 2015, Why Corporate Governance Fails and Lacks Sustainability, accessed 12 September 2017, https://guardian.ng/features/executive-briefs/why-corporate-governance-fails-and-lacks-sustainability/

Ramsay, I, Independent of Australian Company Auditors, *Review of Current Australian Requirements and Proposals for Reform*, Commonwealth of Australia, Canberra, October. Accessed 12 September 2017.

Ravisankar, P., Ravi, V., Rao, G. R., & Bose, I. (2011*). Detection of financial statement fraud and feature selection using data mining techniques. Decision Support Systems,* 50(2), 491–500

Reddy, K., Venter, H. S., & Olivier, M. S. (2012). Using time-driven activity-based costing to manage digital forensic readiness in large organizations. *Information Systems Frontiers,* 14(5), 1061–1077.

Skeet, A. (2015). 'When boards and management conflict', Markkula Centre for Applied Ethics, Santa Clara University, May 2015, accessed September 2017, http://www.scu.edu/ethics/practicing/focusareas/business/board-management-conflict.html.

Spathis, C. (2002). Detecting false financial statements using published data: some evidence from Greece. Managerial Auditing Journal, 17(4), 179–191.

Spathis, C., Doumpos, M., & Zopounidis, C. (2003). Using client performance measures to identify pre-engagement factors associated with qualified audit reports in Greece. *The International Journal of Accounting*, 38(3), 267–284.

Srivastava, R. P., Mock, T. J., & Turner, J. L. (2009). *Bayesian fraud risk formula for financial statement audits. Abacus*, 45(1), 66–87.

US Congress 2002, Sarbanes-Oxley Act 2002, Washington DC.

Zhou, W., & Kapoor, G. (2011). *Detecting evolutionary financial state- ment fraud. Decision Support Systems,* 50(3), 570–575.

SECURITY READINESS EVALUATION FRAMEWORK FOR TONGA E-GOVERNMENT INITIATIVES

Raymond Lutui[1], Semisi Hopoi[2], Siaosi Maeakafa[3]
[1]School of Engineering, Computer & Mathematical Sciences, Auckland University of Technology
Auckland, New Zealand
[2, 3]School of Computer Science, Christ's University in Pacific, Nuku'alofa, Tonga
rlutui@aut.ac.nz, halahalalahi@gmail.com, gmaeakafa@yahoo.com

Abstract

The rapid expansion of the Information and Communication Technologies (ICTs) in the Pacific have reached the Kingdom of Tonga. The submarine fibre-optic cable which connects Tonga to Fiji and onward to a hub in Sydney went live 2013. Now the people of Tonga experience the high-speed impact of digital communication, fast international access, and social changes such as the government is implementing a digital society through e-government services. This study focuses on identifying the factors that will later become a vulnerability and a risk to the security of Tonga government e-government initiatives. Data was collected through interviews with three government officials, document analysis, and critical reflection on the theory context. Consequently, a security-readiness evaluation framework has been designed from the data analysis to inform the e-government initiatives. This study contributes a security-readiness evaluation framework for use in developing countries to guide the implementation of e-government initiatives.

Keywords: E-Government, information security, information systems, cyber security, security threats, security risks, socio-technical

INTRODUCTION

The immense growth in the communication technologies' domain over the last two decades, have changed the way we live our lives and conduct day to day businesses. Information and Communication Technology (ICT) is one of the most vital traits of every new development (Alshehri & Drew, 2010, p.35). This trend of technological advancements has reached the Kingdom of Tonga. In 2010, Tonga Communications Corporation (TCC) – the country's leading provider of complete end-to-end telephony and Internet services – launched the new mobile-broadband-enabled GSM to replace the previous macro-GSM network (Grealish, 2010, P.1). On the 21st August 2013, the first submarine cable (fibre optic) that connects Tonga to the outside world went live (Matangi Tonga, 2013, P.1).

The Kingdom of Tonga is one of the developing countries in the South Pacific with a population of just over 107 thousand. In 2009, the Government of Tonga has identified Information and Communication Technologies (ICT) as an engine for growth in a national ICT vision and strategy. This focuses on Education, Health, Environment Sustainability, and Industry Growth. The National ICT policy for Tonga consists of six main components - Provision of ICT in Homes and Communities, Education and Skill Development, E-Government, Industry Growth and Economic Development, An enabling technical infrastructure and the ICT related legislation (Ma'u, 2015, p.1). The rapid growth of ICT technologies in Tonga is evident in the literature. In 2010, the Tonga Communications Corporation (TCC) – the country's leading provider of complete end-to-end telephony and Internet services – launched the new mobile-broadband-enabled GSM to replace the previous macro-GSM network. Before the introduction of the fibre optic cable, only 20% subscribers across the country for Internet. At the time of writing this paper, over 75% of the country subscribes for an Internet connection.

The Government of Tonga is now talking about implementing e-Government services as part of the National ICT policy for Tonga. The aim is to generate a more efficient way for its government to deliver information and services to its citizens and the business community over the Internet (Cullen & Hassall, 2017, P.4). The completion of the fibre optic cable project, boosted the experience of the people of Tonga in cyber space. The e-government initiative is in its planning stage but need to consider all factors that might affect its successful completion. The aim of this study is to develop a framework to evaluate the readiness of Tonga e-government with regards to cyber security.

According to Grönlund and Horan (2005), the e-government field appeared in the late 1990s. E-government is also known by other names such as e-Gov, Electronic Government, Digital Government, Electronic Governance, and so on (p.39). E-government is defined as, the government owned or operated systems of information and communication technologies that transform relations with citizens, the private sector and/or other government agencies so as to promote citizens' empowerment, improve service delivery, strengthen accountability, increase transparency, or improve government efficiency (Ndou, 2004, p.18). Ndou (2004) rightly argued that, e-governance is more than just a government website. The strategic objective is to support and simplify governance for government, citizens and businesses. The use of ICTs can connect all three parties and provide support for processes and activities (p.18).

The potential of what ICT technologies can provide towards public administration and governance procedures is apparent. However, technology trends such as mobile computing, social media, etc., have introduced new challenges in e-government service design and implementation (Layne & Lee, 2001, p.122). On the other hand, challenges introduced by e-government to information security and privacy is significant. Taking into account the importance and criticality of systems involved in a comprehensive e-government services framework such as e-health, e-tax services, e-education, e-ID, e-procurement, etc., a security breach is immense due to the amount of personal information collected (Layne & Lee, 2001, p.125). Figure 1 shows the security priorities of e-government.

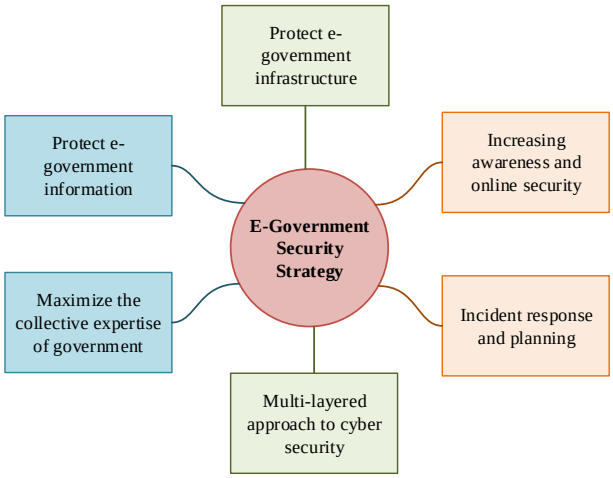


Figure 1: E-Government security priorities

As a result, with e-government, information security issues are becoming ever more prominent. An attack on e-government information system such as hacker attacks, malicious software intrusions, computer crimes and privacy breaches constitute great threats to information security. In addition, developments in both science and technology have posed new challenges to information security (Zhang et al., 2015, p.3). This study can contribute to both theory development for security system integration and practical guidance for operational system that protect information. Taking a socio-technical approach to e-government will show that the balance is between the system and the users. This has to be sensitively and economically managed in order to gain the optimal performance.

The biggest threat to information security today is social engineering and the socio-technical interface between systems and its users. Theoretically, most information system can be strongly protected from attacks but the vulnerability remains of human factors who either by mistake, by trickery, or by intention, knowingly or unknowingly compromise information security (Algarni, et al., 2013, p.510). The literature suggests that it is possible to design e-governance system that interlocks management systems and operational control systems. However, such relationship is dynamic along the lines of interaction so that concepts such as defence in depth are no longer relevant. The security of an information system has both technical and social dimensions.

SOCIO-TECHNICAL APPROACH

The increasing availability of ICT technologies quantified the complexity of socio-technical systems (Vespignani, 2012, p.32). Sociotechnical system (STS) is defined as an approach to complex organizational work design that recognizes the interaction between people and technology. The term also refers to the interaction between society's complex infrastructures and human behaviour (Salnitri, Paja, & Giorgini, 2014, p.50). Therefore, STSs are complex systems where social (human and organizational) and technical components interact with each other to

achieve common objectives. For instance, healthcare systems, smart cities, air traffic management, etc. Based on the statement, e-government systems are complex systems and only make sense to employ socio-technical approach in designing such a system. This study is a social technical system design study which is designed to deal with the governance, management and control of security risk.

In a smart city, citizens will be constantly accessing e-government services such as tax-payment, e-visa application, e-electricity, etc. The amount of information exchanged in such system is substantial, and such information is sensitive and should be secured. In the modern holistic view, the sociotechnical system (STS) is the whole system, not one of two side-by-side systems (Whitworth, 2009, p.395). For instance, a pilot plus a plane are two side-by-side systems with different needs, one mechanical (plane) and one human (pilot). Human Computer Interaction (HCI) suggests these systems should interact positively to succeed (Issa & Isaias, 2015, p.20). However, the plane and the pilot can be seen as a single system. On the mechanical level, the body of the plane and the body of the pilot both have weight, volume, and so forth. However, the pilot adds the human thought level which is above the plane's mechanical level. This allows the plane and pilot system to strategize and analyse.

The socio-technical concept that will be developed changes the priorities, for example, if a social system sits next to a technical one, it is usually secondary. Then, when a social system sits above a technical one, it guides the entire system and that is the primary factor in system performance (Miller, 2004, p.31). Online communities such as e-government, are social-technical systems (STS), built upon social requirements as well as technical ones like bandwidth. As technical problems are increasingly solved, social problems like spam rises. If software can do almost anything in cyberspace, there is still the challenge of what should it do? (Whitworth, de Moor, & Liu, 2006, p.249).

Social Element of Information Security

Securing information assets whether in storage or transmitting around the system is critical and challenging for businesses who rely on ICT to support day to day processes (Dhillon & Backhouse, 2000, p.125). However, advancements of information security technologies do not always guarantee security of information assets. The human factors of information security will always be a challenge and an issue in terms of managing of security (Siponen et al., 2014, p.217). As a result, a security framework is required to combine systems, operation, and internal controls to ensure confidentiality, integrity, availability of the critical information assets. Due to the seriousness of threats of unauthorized access over the Internet, effective information security management is one major concerns. A malicious insider is defined as a current or former employee, contractor, or business partner who - has or had authorized access to an organization's network, system, or data, has intentionally exceeded or intentionally used that access in a manner that negatively affected the confidentiality, integrity, or availability of the organization's information or information systems (Silowash et al., 2012, p.9).

Security mechanisms such as firewalls, intrusion detection systems (IDS), and electronic access systems to enter the building or rooms. These are implemented particularly as a protection against external threats. However, that still does not address the issues of insider threats. Insiders not only aware of their organization's policies, procedures, and technology but, they also aware of the known vulnerabilities in the system. The ISACA's Cybersecurity Snapshots of the issues facing organizations revealed that, the top three cyber threat are social engineering, insider threats and advanced persistent threats. The 2017 Cyber Security Survey reported that 65% of the IT security professional respondents are not confident in their organizations security posture (Meyer, 2017, p.2). According to (CERT, 2011), the threat of attack from insiders is real and substantial. Another study conducted by (CSO Magazine, 2011, p.1) explained that 46% of the respondents thought that damage caused by insider attacks was more severe than damage from outsider attacks. The study found that the most common insider crimes were - unauthorized access to or use of corporate information; unintentional exposure of private or sensitive data; viruses, worms, or other malicious code; theft of intellectual property (IP).

(Silowash et al., 2012, p.2) discussed that, insider threats are influenced by a combination of technical, behavioural, and organizational issues and must be addressed by policies, procedures, and technologies. (Silowash et al., 2012, p.9) also added that their current analysis recognizes that - intellectual property (IP) theft, IT sabotage, fraud, espionage, and accidental insider threats seems to be the unique patterns of insider threat behaviour. Siponen et al., (2014) pointed out that, the key threat to information security comes from employees who do not comply with information security policies (p.220).

Technical Element of Information Security

In von Solms (2010), the author discussed the five waves of information security. The first wave is technical wave, second wave is management wave, third wave is institutional wave, fourth wave is information security governance wave, and the fifth wave is the cybersecurity wave (p.2). In the era of the technical wave, Information

Security main concern was a form of Identification and Authentication for logging onto the mainframe system, and Logical Access Control. Most of these functions were handled by technical people.

Whitman and Mattord (2016) explained that, there are six critical components of information system, the hardware, software, networks, people, procedures, and data enable information to be input, processed, output, and stored (p.114). Each of these IS components has its own strengths, weaknesses, characteristics and security requirements. The hardware component is dealing with technical and physical technology that is responsible executing, storing and transmitting of the data. Settanni et al., (2017) stated that, today's information systems are increasingly complex. Their interconnected nature exposes them to advanced cyber threats (p.167). Anderson and Fuloria (2010) pointed out that, a failure of critical infrastructure can cause significant damage within a short period of time (p.55). There are a number of various threats to critical infrastructure such as the most obvious, natural disasters such as flooding, fire, equipment malfunction or also human error. However, recently, targeted attacks by hackers on the information system infrastructures has become significant (Miller & Rowe, 2012, p.51).

There is a lot involved in physical/technical elements of information security. However, physical/technical security system such as a 24x7 security guard and surveillance cameras. Access to areas where confidential work is done usually require electronic access cards or biometric access system. Security cameras is in place and uses of personal memory devices or CD/DVD ROMs are not allowed. All network/Information System traffics are configured to go through a firewall and a proxy server so web access and activities can be monitored and controlled. Access rights are controlled so no one can have the rights to execute .exe files. All files are redirected to store on the server not locally so they can be backed up regularly and stored offsite.

DESIGN OF THE STUDY

Exploratory research, on the other hand, is employed in this type of study as it allows the researcher to gain a deeper understanding of an issue or problem. Jebb et al. (2017) describe that, exploratory perspective is designed ready to find patterns that are different than expected. Such patterns may provide theoretical insights or provide information to guide further analyses (p.266). As a result, Exploratory research is the chosen methodology for this study.

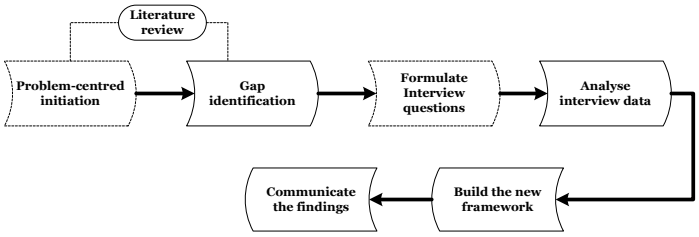


Figure 2: Design of the study

The study is designed to first identify the problem – the gap for this study based on the literature survey. The next task is to define the problem and extract its significance. This involves analysing the existing and relevant literature; identify the problem, interview experts in the field. This presents an opportunity to employ a gap analysis strategy by completing a component-by-component analysis. Therefore, influencing how the interview questions should be designed to create awareness of factors that will affect the security readiness of Tonga e-government initiative. The next stage is designed to deal with the formulation of the interview questions for data collection. The next stage deals with data analysis, in this phase, the factors that will affect the security readiness of the Tonga e-government will be identified. The data gathered from this phase will be used to construct the security evaluation framework for the Tonga e-government initiative. The final phase is designed to allow the researcher to employ various scholarly electronic databases to communicate the outcome of the study. This communication might include the problem and its importance, the artefact and its effectiveness to other researchers and practitioners in the field including the Government of Tonga. It is also suggested that researchers should conclude the study with communicating the implications of the result for the practical field.

Data Analysis and Discussions

An interview was conducted with three senior government officials of the Tonga government. At the moment, there only two government ministries that was formed to work on the Tonga e-government initiative together with a consultant from the Asian Development Bank (ADB) based on the National ICT policy for Tonga. This is why this study only interview the three top officials that are working on the Tonga e-government. The three officials are the consultant from the ADB, the CEO of Tonga Computer Emergency Response Team (CERT) and the Senior

Engineer from the Ministry of Meteorology, Energy, Information, Disaster Management, Climate Change and Communications (MEIDECC). Table 1 summaries the answers received and has been categorised into four categories but grouped into only two main groups.

Table 1: Security requirement categories interview data

Questions	Technological	Policies & Regulations
• What are the benefits of employing ICT? • What are the challenges faces by the Government in implementing ICT? • How well are e-Government initiatives aligned with Tonga frameworks and requirement? • Who is driving these initiatives? • How are they going to be funded?	• Access to information • Decision making • Centralized DataBase • Share Information • Easy communication • Build each Ministry's ICT capacity • Unsecured software application • Unsecured ICT infrastructure • Inexperience staffs • Unskilled IT staffs	• Lack of proper tools for the job • Lack of funding • Lack of inter-ministry communication • Lack of top management support • Too much politics • Dependent on overseas donations • Out-of-date computer crimes act

The next phase is to prepare and get the data organized. The next phase is designed to provide a description of the case considered. During this phase, data coding is done also, a number of questions is asked such as, what is this incident about? what category does this incident indicate? what property of what category does this incident define? what is the 'main concern' of the participants? The coding of the data was based on their properties and security requirements. Properties considered in this study are conditions, causes, consequences, hierarchies and contexts. However, this study further grouped these categories into two groups as it shown in table 1. The purpose is not only to summarize the interview data but to identify their inter-relationships and establish how they help to explain the phenomenon under study.

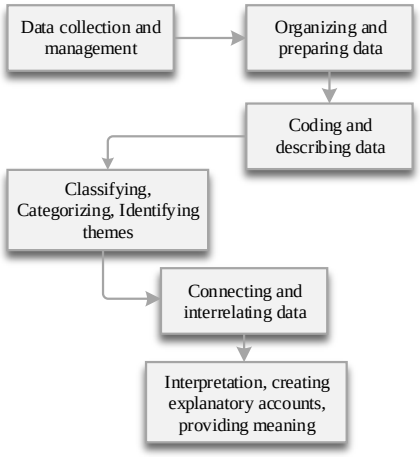


Figure 2: Data analysis technique

The final phase is designed to allow for interpretation, developing explanations and provide meaning for the data. As mentioned earlier, this study is taking a socio-technical approach to cyber security of e-government, more specifically, Tonga e-government. Looking at the interaction between social and technical dimensions of modern life as dynamic, constantly changing that both emerge from and shape modern society. As the technological age advanced, however, it presented an ever-growing array of new issues and unknowns. In response, practitioners began to focus on more specific areas for action, and most researchers began to select more narrowly defined subjects for study. The security needs are expressed through the processes.

SECURITY READINESS EVALUATION FRAMEWORK

Based on the coding done of the interview data, following are the security challenges to Tonga e-government can be identified. However, due to the complexity of e-government systems, these security challenges might affect implementation and management. According to Evans (2011), by 2020 there will be over 50 billion connected

objects against a population of 7 billion. An object can be anything embedded with computation, storage and communication capabilities with different capacities (sensor, actuator, mobile phone, desktop, laptop, printer, car, fridge, oven, etc.) (p.7). Therefore, security in such environment is paramount. The main objectives of e-government initiatives are to provide one-stop quality public services and value-added information to citizens and businesses. At the same time, to enable government agencies to work together and achieve internal efficiency and effectiveness of operations (Lee et al., 2005, p.99).

The UN survey reported that efforts are being made to ensure privacy and security of personal data yet challenges remain. Some related to the technical difficulties associated with ensuring interoperability of systems. However, proliferation of technologies makes it difficult to provide integrated services such as e-health, etc. Major security services required for e-government consists of three aspects that should be considered:

- **Confidentiality:** refers to protection of information from unauthorized disclosure e.g. to the press or to release through improper disposal techniques, or to those who are not entitled to have the same.
- **Integrity:** is about protecting information from unauthorized modification, and ensuring that information, such as a beneficiary list, can be relied upon and is accurate and complete.
- **Availability:** is to ensure that the information is available when it is required.

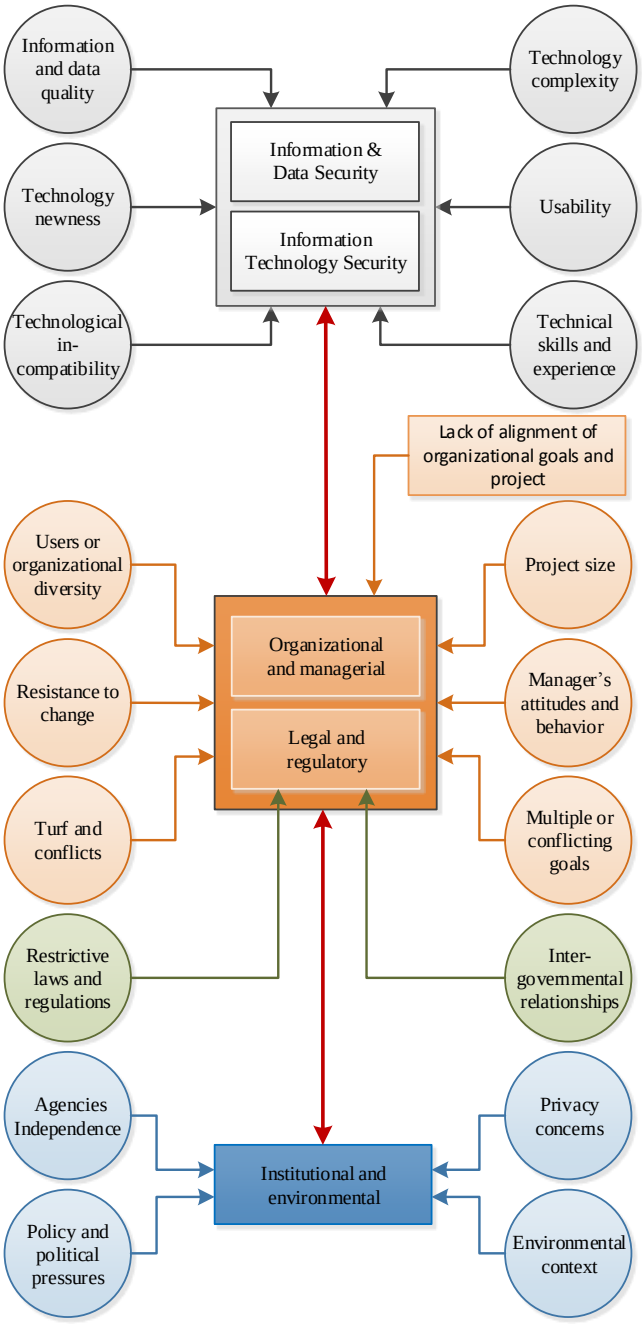


Figure 4: Security e-readiness evaluation framework.

Information and Data Security

E-government initiatives are goals set up by governments and accomplished through the use of ICTs. E-government initiatives are focused on capturing, management, use, dissemination, and sharing, of information. A number of the challenges relate to the information at the core of E-government initiatives. Researchers in the e-government domain principally focused on the issues relating to quality and accuracy of the data. However, legal experts expressed their concerns regarding the fact that every data processing duplicates the risk of abuse (intended or not). On the other hand, poor data quality increases operational cost because time and other resources are spent detecting and correcting errors. Information is power and organizations tend to be sensitive about giving out information as it reflects on the quality of their operations. As a result, information has to be used differently and its management in order to satisfy the digital requirements. System usability and ease of use are an important factor to consider when designing E-government initiatives. Complexity and newness of technology are also constraints that can potentially affect the results. The lack of relevant technical skills within the ministries has been found to be an important factor as well as the shortages of qualified technical personnel.

Organizational and Managerial

The size of the implementation and the diversity of users and organizations involved are two of the main challenges for E-government initiatives. There are at least two other problems related to the goals and objectives of the initiative. First is the lack of alignment between organizational goals and the E-government project. This alignment may be understood as a certain type of balance that needs to be in place to achieve one or more goals. Contextual factors are often addressed in information systems (IS) research as situational, organizational, environmental, task, and technology characteristics with influence for the outcome of an E-government project.

E-government initiatives are required to be value-driven and not technology driven. The promised benefits of E-government cannot be based on digitizing information and putting it on the web alone. Technology needs to be fully realized and the transformation brought about to better serve the citizens. E-government is to facilitate an information society that can influence every aspect of daily life. Yet, more detached observers maintain that there is no post-industrial society therefore, individual interests and associated behaviours lead to internal conflicts about change and resistance to innovation is such as E-government.

Legal and Regulatory

Government ministries operate according to a specific and formal sets of rules. Li (2003) pointed out that, e-government is not a technical issue, but an organizational issue (p.45). As a result, to successfully implement e-government principles and functions, will require a new set of rules, policies, and laws. There will also be changes to satisfy the requirements for electronic activities such as electronic archiving, signatures, transmission of information, data protection, computer crime, intellectual property rights and copyright issues and so on. Dealing with e-government means signing a contract or a digital agreement, which has to be protected and recognized by a formalized law (Hwang et al., 2004, p.10). In Tonga, e-business and e-government laws are not yet available.

Institutional and environmental challenges

The availability of resources is an economic challenge associated with emerging technologies that is a great challenge for developing countries. Obstacles with funding, regulations, and patents that can derail technology development and adoption (Woodson, 2016, p.1410). However, in this environment, institutions are not only laws and regulations, but also norms, actions, or behaviours that people accept as good or take for granted. On the other hand, deployment of security measures can lead in to an intrusion on users' privacy. Privacy and related security issues are challenges that must be adequately addressed in E-government IT initiatives (Arroyo et al., 2015, p.455). In this digital age, it is not possible to imagine conducting day to day work without the use of technology.

External pressures such as policy agendas and politics may affect the results of IT initiatives. The discussion highlights the range of complex and various challenges public managers must face as they work in the e-government area (Walker, 2015, p.297). Success is not only about choosing the right technology, but also managing organizational capabilities, regulatory constraints, and environmental pressures. For E-government managers to be successful in their initiatives they must be aware of these challenges and use appropriate strategies to overcome the institutional and environmental challenges.

Digital Divide

The digital divide refers to the gap in opportunity between those who have access to the Internet and those who do not. Those who do not have access to the Internet will be unable to benefit from e-government services. Digital Divide is a very serious matter in Tonga, not all citizens currently have equal access to computers, it can an issue of affordability or lack of necessary skills (Lu, 2001, p.1). Government should look providing Internet-enabled computers in schools and public libraries. Yet, few types of challenges that government may face are-affordability, elderly, language barrier, and the inexperienced or computer illiterate or not so well educated (Hassani, 2006, p.250). Government should provide basic skills training to both its employees and citizens in order to let them participate in e-government development applications.

CONCLUSION

The result of the literature survey highlighted the fact E-government is a complex system and Socio-technical system (STS) is defined as an approach to complex organizational work design that recognizes the interaction between people and technology. As a result, this study decided to take a Socio-Technical approach to e-government security because, in the modern holistic view, the sociotechnical system (STS) is the whole system,

not one of two side-by-side systems that is – social and technical. Human Computer Interaction (HCI) suggests these systems should interact positively to succeed.

Interview was used as the method of data collection. The growth of ICT technologies is very fast however, the country is very small and lack of expert personnel in the field seems a real issue. The interview data was collected from three experts in Tonga that are currently involve with the e-government initiative. The purpose of the study is to identify the factors that may become a risk and a vulnerability of Tonga e-government services. E-government presents several technical, economic and social challenges that will surface as the E-government development moves forward. Looking at various reasons to why security-readiness is so important to both governmental and non-governmental organizations, the ease of using these measures is most prominent. Having an easily quantifiable set of indicators is vital. This will provide the government of Tonga with an overview of their situations. It can be also used as a basis for comparison and future planning. This advantage arises from the fact that security-readiness measures have the ability to summarize a broad set of characteristics of the Kingdom of Tonga.

There are a number of existing security-readiness evaluation frameworks with various objectives, methodologies and results. However, there is no one fits all evaluation framework yet, and this study has contributed a security-readiness framework that is both relevant and practical for implementation in a developing country.

REFERENCES

Algarni, A., Xu, Y., Taizan, C., & Yu-Chu, T. (2013). Social engineering in social networking sites: Affect-based model, *Proceedings of the 8th International Conference for Internet Technology and Secured Transactions (ICITST-2013)* (pp. 508-515). UK: IEEE.

Alshehri, M., & Drew, S. (2010). E-Government Fundamentals, *Proceedings of the IADIS International Conference ICT, Society and Human Beings* (pp. 34-42). Freiburg: MCCSIS.

Anderson, R., & Fuloria, S. (2010). Security Economics and Critical National Infrastructure. In T. Moore, D. Pym, & C. Ioannidis (Eds.), *Economics of Information Security and Privacy* (pp. 55-66). Boston: Springer.

Arroyo, D., Diaz, J., & Gayoso, V. (2015). On the Difficult Tradeoff Between Security and Privacy: Challenges for the Management of Digital Identities. In Á. Herrero, B. Baruque, J. Sedano, H. Quintián, & E. Corchado (Eds.), *International Joint Conference: CISIS'15 and ICEUTE'15* (pp. 455-462). Cham: Springer.

Basu, S. (2004). E- government and developing countries: an overview. *International Review of Law, Computers & Technology, 18*(1), 109-132.

Cullen, R., & Hassall, G. (2017). E-Government in Pacific Island Countries. In R. Cullen & G. Hassall (Eds.), *Achieving Sustainable E-Government in Pacific Island States* (pp. 3-32). Cham: Springer

DeLisi, P. S. (1990). Lessons from the steel axe: culture, technology, and organizational change. *MIT Sloan Management Review, 32*(1), 83.

Dhillon, G., & Backhouse, J. (2000). Technical opinion: Information system security management in the new millennium. *Commun. ACM, 43*(7), 125-128.

Evans, D. (2011). The internet of things: How the next evolution of the internet is changing everything. *CISCO white paper, 1*(2011), 1-11.

Grealish, A. (2010). *Altobridge wireless network goes live in tonga*. Retrieved July 25, 2017, from https://www.realwire.com/releases/Altobridge-Wireless-Network-goes-Live-in-Tonga

Grimsley, M., & Meehan, A. (2007). e-Government information systems: Evaluation-led design for public value and client trust [journal article]. *European Journal of Information Systems, 16*(2), 134-148.

Grönlund, Å., & Horan, T. A. (2005). Introducing e-gov: history, definitions, and issues. *Communications of the association for information systems, 15*(1), 39.

Hassani, S. N. (2006). Locating digital divides at home, work, and everywhere else. *Poetics, 34*(4–5), 250-272.

Hwang, M.-S., Li, C.-T., Shen, J.-J., & Chu, Y.-P. (2004). Challenges in e-government and security of information. *Information & Security, 15*(1), 9-20.

Issa, T., & Isaias, P. (2015). Usability and Human Computer Interaction (HCI). In *Sustainable Design: HCI, Usability and Environmental Concerns* (pp. 19-36). London: Springer.

Jebb, A. T., Parrigon, S., & Woo, S. E. (2017). Exploratory data analysis as a foundation of inductive research. *Human Resource Management Review, 27*(2), 265-276.

Jones, S., Hackney, R., & Irani, Z. (2007). Towards e-government transformation: conceptualising "citizen engagement" A research note. *Transforming Government: People, Process and Policy, 1*(2), 145-152.

Layne, K., & Lee, J. (2001). Developing fully functional E-government: A four stage model. *Government Information Quarterly, 18*(2), 122-136.

Lee, S. M., Tan, X., & Trimi, S. (2005). Current practices of leading e-government countries. *Commun. ACM, 48*(10), 99-104.

Li, F. (2003). Implementing E-Government Strategy in Scotland: Current Situation and Emerging Issues. *Journal of Electronic Commerce in Organizations, 1*(2), 44-65.

Lu, M.-t. (2001). Digital Divide in Developing Countries. *Journal of Global Information Technology Management, 4*(3), 1-4.

CSO Magazine. (2011). 2011 cybersecurity watch survey: How Bad Is the Insider Threat? *CSO Magazine, January, 1*(1), 1-8.

Matangi Tonga. (2013). *Tonga's high-speed internet goes live august 21*. Retrieved July 25, 2017, from http://matangitonga.to/2013/08/14/tonga%E2%80%99s-high-speed-internet-goes-live-august-21

Ma'u, P. (2015). E-Government in Tonga. *Asia-Pacific Regional Forum on e-Government, 1*(1), 1-19.

Miller, C. A. (2004). Human-computer etiquette: Managing expectations with intentional agents. *Communications of the ACM, 47*(4), 31-34.

Miller, B., & Rowe, D. (2012). A survey SCADA of and critical infrastructure incidents. *Proceedings of the 1st Annual conference on Research in information technology* (pp.51-56). Canada: ACM.

Meyer, D. (2017). Check Point's 2017 Cyber Security Survey Shows Key Concerns and Opportunities among IT Professionals. *2017 Check Point Software Technologies, 1*(1), 1-3.

Ndou, V. (2004). E-government for developing countries: opportunities and challenges. *The Electronic Journal of Information Systems in Developing Countries, 18*.

Nkwe, N. (2012). E-Government: Challenges and Opportunities in Botswana. *International Journal of Humanities and Social Science, 2*(17), 39-48.

Salnitri, M., Paja, E., & Giorgini, P. (2014). Preserving Compliance with Security Requirements in Socio-Technical Systems. In F. Cleary & M. Felici (Eds.), *Cyber Security and Privacy: Third Cyber Security and Privacy EU Forum, CSP Forum 2014, Athens, Greece, May 21-22, 2014, Revised Selected Papers* (pp. 49-61). Cham: Springer.

Seifert, J. W. (2003). A primer on e-government: Sectors, stages, opportunities, and challenges of online governance *Library of Congress Washington DC Congressional Research Service, 1*(1), 1-25.

Settanni, G., Skopik, F., Shovgenya, Y., Fiedler, R., Carolan, M., Conroy, D., . . . Olli, P. (2017). A collaborative cyber incident management system for European interconnected critical infrastructures. *Journal of Information Security and Applications, 34*, 166-182.

Sharma, S. K., & Gupta, J. N. (2003). Building Blocks of an E-Government—A Framework. *Journal of Electronic Commerce in Organizations, 1*(4), 34-48.

Silowash, G., Cappelli, D., Moore, A., Trzeciak, R., Shimeall, T., & Flynn, L. (2012). Common Sense Guide to Mitigating Insider Threats. 4th edn. CERT Carnegie Mellon Software Engineering Institute: Carnegie Mellon University.

Siponen, M., Adam Mahmood, M., & Pahnila, S. (2014). Employees' adherence to information security policies: An exploratory field study. *Information & Management, 51*(2), 217-224.

United Nations. (2008). *UN E-Government Survey 2008: From E-Government to Connected Governance*. New York: United Nations publication.

United Nations. (2016). *United Nations E-Government Survey 2016: E-government in Support of Sustainable Development.* New York: United Nations publication.

Vespignani, A. (2012). Modelling dynamical processes in complex socio-technical systems. *Nature physics, 8*(1), 32.

Vinod Kumar, T. M. (2017). E-Democracy for Smart Cities: Conclusion and Path Ahead [Vinod Kumar2017]. In T. M. Vinod Kumar (Ed.), *E-Democracy for Smart Cities* (pp. 523-551). Singapore: Springer.

von Solms, S. H. (2010). The 5 Waves of Information Security – From Kristian Beckman to the Present. In K. Rannenberg, V. Varadharajan, & C. Weber (Eds.), *Proceedings of the 25th IFIP TC-11 International Information Security Conference. Security and Privacy – Silver Linings in the Cloud: Held as Part of WCC 2010* (pp. 1-8). Brisbane: Springer.

Walker, E. T. (2015). The politics of information: Problem definition and the course of public policy in America. *Interest Groups & Advocacy, 4*(3), 297-301.

Whitman, M. E., & Mattord, H. J. (2016). *Principles of Information Security* (5ed.). USA: Cengage.

Whitworth, B. (2009). A brief introduction to sociotechnical systems. In *Encyclopedia of Information Science and Technology, Second Edition* (pp. 394-400): IGI Global.

Whitworth, B., de Moor, A., & Liu, T. (2006). Towards a Theory of Online Social Rights. In R. Meersman, Z. Tari, & P. Herrero (Eds.), *On the Move to Meaningful Internet Systems 2006: OTM 2006 Workshops: OTM Confederated International Workshops and Posters, AWeSOMe, CAMS, COMINF, IS, KSinBIT, MIOS-CIAO, MONET, OnToContent, ORM, PerSys, OTM Academy Doctoral Consortium, RDDS, SWWS, and SeBGIS 2006, Montpellier, France, October 29 - November 3, 2006. Proceedings, Part I* (pp. 247-256). Berlin: Springer.

Willoughby, M., Gómez, H. G., & Lozano, M. Á. F. (2010). Making e-government attractive [journal article]. *Service Business, 4*(1), 49-62

Woodson, T. S. (2016). Public private partnerships and emerging technologies: A look at nanomedicine for diseases of poverty. *Research Policy, 45*(7), 1410-1418.

Zhang, H., Han, W., Lai, X., Lin, D., Ma, J., & Li, J. (2015). Survey on cyberspace security. *Science China Information Sciences, 58*(11), 1-43.

EVALUATING IP SURVEILANCE CAMERA VULNERABILITIES

Brian Cusack, Zhuang Tian
Cyber Forensic Research Centre, Auckland University of Technology, Auckland, New Zealand
brian.cusack@aut.ac.nz, zhuang_tain@hotmail.com

Abstract

Hacking of IP surveillance camera systems came to public attention in 2016 when the high bandwidth and resources were exploited for a massive DDoS attack that affected one third of all US Internet services. A review of previous studies show that a vast number of IP cameras have been hacked because the default usernames and passwords have not been changed from the factory defaults. In this research we asked, What are the vulnerabilities of an IP surveillance camera? The purpose of the study was to provide identification of vulnerabilities and guidance for the protection of surveillance camera systems. The research shows that the tested surveillance camera had many vulnerabilities and that there is urgency for distributing alerts and best practice guidelines.

Keywords: Hacking, CCTV vulnerability, Evaluation, Security

INTRODUCTION

Closed Circuit Television (CCTV) systems have proliferated in businesses and for private use. The surveillance systems are relatively inexpensive and provide multiple sensors that feed information back to a centralised processing station and monitoring screens. The application is for monitoring assets and human behaviour for risk management. The sensors provide different data types that include visual, audio, infrared, and other spectrum data. Monitoring may proceed by human observation, automation, archival mapping, or a combination of these. Many systems have software to assist human decision-making, and resource management systems to optimise the cost of surveillance against the benefits it may deliver. Research has shown that these CCTV surveillance systems have critical points of failure (Costin, 2016). In addition, Ozkan (2016) shows that over 100,000 wireless Internet Protocol (IP) cameras in the research sample had little or no information security protection. Others show that surveillance cameras from 79 vendors are vulnerable to Remote Code Execution (RCE) (Kirk, 2016; Costin, 2016). The security problem is increased by vendors are selling IP cameras using the "white labelling" business model with the same firmware developed by the same company across the product range and with unprotected RCE. The vulnerability allows an attacker to seize control of the camera for manipulation. Manipulation can have several features, such as, data seizure, mechanical manipulation, anti-forensic data planting, exploitation of the bandwidth resource, end-user deception, and zombie exploitation (McKee, et al., 2017). A significant weakness is that most IP cameras only log authenticated requests and have no traces on the camera of user activity or unique identification. Hence, an attacker can be anonymous while acquiring real-time video streams, archived footage; email, FTP, other credentials, and access to the system resource controls. The significant vulnerability grants an attacker invisibility and the ability to host malware; run arbitrary software such as botnets, proxies and scanners; and create backdoors for future access. Consequently, a CCTV system is generally available to unauthorised control, and the system itself, can sponsor attacks on other systems (Coole, et al., 2012; Cuputo, 2014; Costin, 2016). In this paper, we test an out-of-the box camera to identify security vulnerabilities.

BACKGROUND

On 21 October 2016, a massive DDoS attack against Dyn, a domain name system (DNS) provider, broke a large portion of the Internet, causing a significant outage to hundreds of websites and services (CCTV, 2017). Although, Dyn did not disclose the actual size of the attack, but it has been speculated that the DDoS attack could be much bigger than the one that hit French Internet service and hosting provider OVH that peaked at 1.1 Terabytes per second (TBps), which is the largest DDoS attack known to date (Smith, 2013). The attack was caused by a botnet that consisted of 100,000 devices infected by malware named Mirai. The Mirai malware targeted Internet of Things (IoT) devices such as IP cameras and digital video recorders (DVR) that have weak default passwords, making them easy to infect (Wu, et al., 2010; Zanella, 2014; Kirk, 2016). A similar study by Minin (2015) found that a malicious attacker took control of the cameras remotely and controlled movement, redirected the video feeds, and worked out the password for the wireless network the device was connected. The owners of the surveillance camera systems were not aware of the system compromise and the use for a massive attack. A similar study analysed Motorola's Focus 73 (Minin, 2015) outdoor security camera. Images and video taken by the camera can be delivered to a mobile phone application. One attack showed how it is possible to scan for cameras connected to

the Internet and then to get a reverse root shell to forge control credentials. Additionally, by tampering with DNS settings, the attacker can intercept the alerts that the camera sends to its owner, as well as to see video clips that would be sent to a cloud storage service. Further analysis showed that the firmware is a generic and used in other kinds of IP cameras. The code is not encrypted or digitally signed leaving open a backdoor for malware to be uploaded to the camera (Gotham Digital Science, 2012).

The argument for protection when the surveillance system is on a dedicated network without access to other client systems, is shown to be false (Tekeoglu et al., 2015). The experiment was performed on MicroDigital, HIVISION, CTRing and a substantial number of other rebranded devices. The result shows the tested surveillance systems transmit the user name and password in clear text over port 5920 transmission control protocol (TCP) during authentication stage. The experiment also used a Metasploit framework to perform brute-force and dictionary attacks on the tested devices. The sample showed that 70% of the instances utilised had the default vendor passwords that had not been changed.

The list of known CCTV vulnerabilities have been published in a database (CCTV Calculator, 2017). They list vulnerabilities existing in the vendors product range including Siemans, ZoneMinder, Zhuhai RaySharp, Samsung, Grandstream, WESPMonitor, WebGate, D-link, Panasonic, Cisco, Hikvision, FOSCAM, Y-Cam, TRENDnet, CIPCAMPTIWL, Dahua, TVT, AVTECH, Brickcom, TP-LINK, AirLive, Axis, Sony, QNAP, Arecont Vision, GeoVision, March Networks, Canon, FlexWATCH, Mobotix and Linksys. The vulnerability discovered in GeoVision DVR systems allows a remote attacker to execute arbitrary code by calling the GetAudioPlayingTime method with arguments. Tian (2014) shows more detailed vulnerabilities in GeoVision include directory traversal in geohttpserver and SanpShotToFile in GeoVision LiveX. Weak encryption schemes for passwords allows attackers to obtain the password via sniffing (Wu, et al., 2010). The sysinfo script in GeoHttpServer allows remote attackers to cause a DoS via a long password, and triggering a buffer overflow. When GeoHttpServer is configured to authenticate users, it allows attackers to bypass authentication and access unauthorised files via a URL that contains %0a%0a – code injection (Bruschi, et al., 2003; Bojinov, et al., 2009). These examples indicate the GeoHttpServer has several vulnerabilities that gives access for an attacker to perform unauthorised activities within the surveillance system. Nonetheless, these vulnerabilities were discovered in the period between 2004 and 2011 and no information is provided regarding whether or not these vulnerabilities have been fixed by the manufacturers since. Further research (Gotham Digital Science, 2012; Kyaw, et al., 2016) shows a remote file disclosure vulnerability in GeoHttpServer. The code has no authentication requirement and hence an attacker can exploit this vulnerability to retrieve and download stored files on the server such as ‘boot.ini’ and ‘win.ini’.

RESEARCH METHODOLOGY

The aim of this research was to answer the research question: What are the vulnerabilities of an IP surveillance camera? To answer the question, the research has six phases (Figure 1). These phases include literature review, system setup, pilot testing, data collection, data analysis and its comparison with results of previous research. Different research phases employ different research methods. The literature review section, for example, provided understanding for the work of different authors and their recommendations for future research. This phase constitutes the qualitative part of the study. The data collection, on the other hand, included a pilot study and experiment conducted by testing the camera by trying different exploits. The system setup phase set up the equipment for the field trials. These rational phases constitute the quantitative part of the study. The final phase compares the results obtained from both parts of the study in a mixed methods approach (Bryman, 2012).

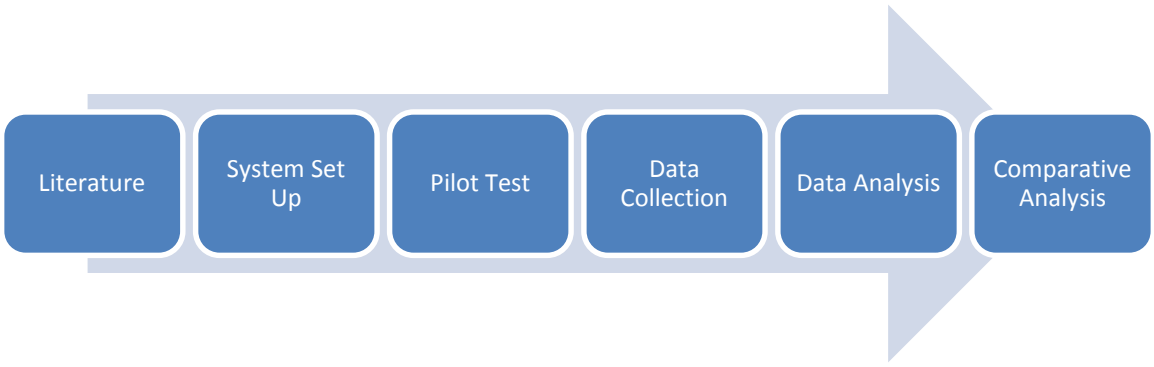


Figure 1. Research Phases

System Design

The following devices are used in the research, and the system design is in Figure 2:

- *Target IP surveillance camera (10.0.0.2): GeoVision GV-FD220D 2MP H.264 IR fixed IP Dome camera*
- *Network switch: Thomson TG585 v8 ADSL2+ wireless gateway*
- *Client (10.0.0.5): Lenovo laptop Thinkpad X200 Table with Intel Core 2 Duo CPU L9600 2.13GHz ×2, 242.9 GB HDD and Windows 7 32-bit*
- *Attack device 1(10.0.0.6): Lenovo laptop Thinkpad X200 Table with Intel Core 2 Duo CPU L9600 2.13GHz ×2, 242.9 GB HDD and Kali Linux Rolling 2016.2 32-bit*
- *Attack device 2 (10.0.0.3): Acer laptop Aspire V3-371-501P with Intel Core i5-4210U 1.7GHz, 4GB DDR3, 500 GB HDD and Windows 8.1 64-bit*

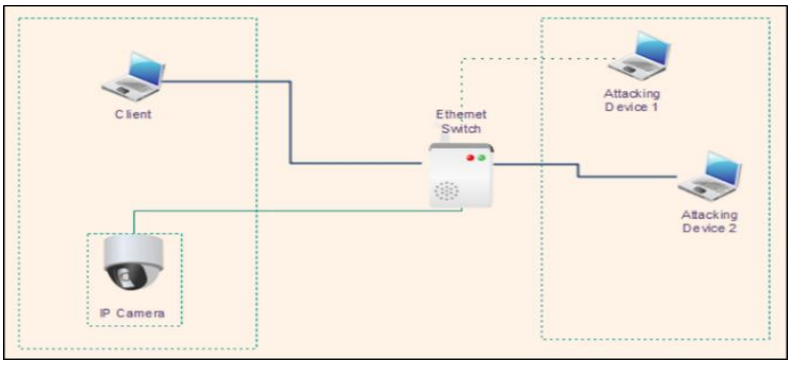


Figure 2. System Design

Pilot Test

After setting up the IP surveillance system network, a pilot run was made to configure and test the camera functionalities as well as network connections amongst all the devices. The user can connect to the IP camera either through *Windows Explorer* by entering its IP address in the URL field; or use GeoVision *DMMultiView* client software to connect the camera's DVR by selecting the host IP address and type of device. A User can use *GvIP Device Utility* to find the IP camera IP address. *GV IP Device Utility* is an application software to help the user to manage IP cameras, update their firmware, identify them by their IP addresses within a local area network (LAN) or backup and restore their settings (Figure 3).

Figure 3. GeoVision DMMultiview User Authentication and GeoVision DMMultiview Live Capture

The attacking device ran *Kali Linux*, so we also needed to test whether it can connect to the IP camera in the pilot study, and to ensure a penetration test is possible using preinstalled tools from the attacking device.

Data Management

The tools *Angry IP Scanner, WireShark, ophcrack, Burpsuite and Cain & Abel*, were tested in the pilot study for performance and functionality. Each has their own built-in data processing ability as specified by the distinct

features and functions of the tool. Others tested, such as *Nmap Hydra, Nikto and Metasploit,* are command-line based and are relevant for data collection from IP cameras. The collected data are automatically processed and analysed by these tools. The results can be saved to a file; analysis performed, and a report generated. Data collection was undertaken with website and IP camera DVR penetration testing tools and techniques. For result accuracy, the same tools are used 3 times and then the collected data compared to identify any variations. Data dump files are created for each penetration tool used, and the collected data analysed.

RESEARCH FINDINGS

Angry IP Scanner and Nmap were used to collect information about the target system, such as its IP address, media access control address (MAC), manufacturer and server information. The *Angry IP Scanner* is a fast lightweight cross-platform IP address and port scanner; used to scan IP addresses in any range. It includes information on any of the ports by simply pinging each IP address to check if it is alive, then optionally resolving its hostname, determining MAC addresses and the vendor (Figure 4).

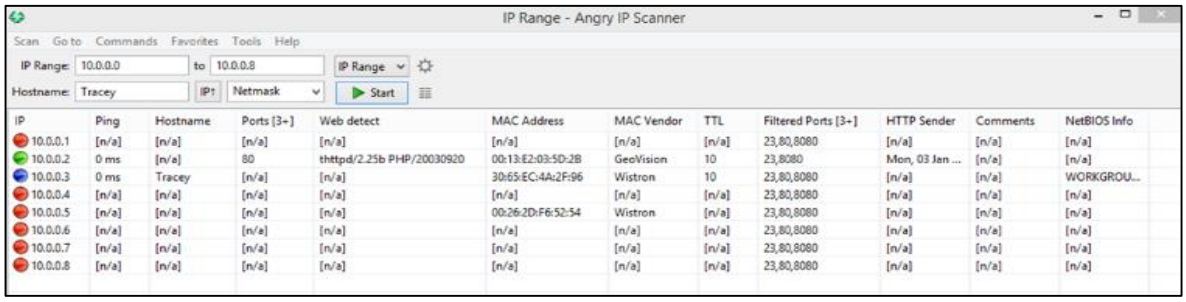


Figure 4. Angry IP Scanner Network Scanning Result

The result shows that our target IP surveillance system's IP and MAC address, its manufacturer information as well as its active ports. Once, we obtain the target IP address, we used *Nmap* for further reconnaissance. Nmap is a free and open source utility for network discovery and security auditing. It uses raw IP packets to determine available hosts on the network, services offered, operating system (OS) they are running, type of pocket filters and firewalls in use as well as other user characteristics. From the Nmap scanning results, the target IP camera has TCP port 80, 111 and 10000 open. Hence, it is shown again that a user can login to the target IP surveillance system through Windows Explorer via port 80, and port 10000 is the virtual switch system (VSS) port for video streaming. Thus, to further the research IP packets were collected, and packet sniffing and spoofing performed to identify any vulnerabilities in the system. Packet sniffing and spoofing are methods that identify the weak points of network system, particularly on a layer 2 switched network. A LAN uses address resolution protocol (ARP) with holes enabling the attacker to sniff packets and lodge ARP spoofing attacks.

WireShark was put into monitoring and capturing mode to authenticate to the target surveillance system website application, in order to capture the user name and password either in clear text or in hash values. The captured packets were then analysed and by following the TCP packet streams, other matters for further investigation were discovered. Firstly, we were able to find the user name and password; and the two MD5 hash values. There were also two groups of 50 bits assigned to two variables, namely *gUserName* and *gPassword*. Finally, we saved both MD5 hash values to be decrypted. Similarly, *Cain & Abel* was used to recover passwords by sniffing the network, cracking encrypted passwords using dictionary, brute-force and cryptanalysis attacks, recovering wireless network keys, revealing password boxes, uncovering cached passwords and analysing routing protocols. *Cain & Abel* sniffs the network for the target device, and then launches attacks. We used ARP poisoning to perform a man-in-the-middle (MITM) attack. During a MITM attack, the attacking device secretly intercepts, replays and potentially alters the communication between two parties who believe they are directly communicating with each other – in this case the camera and its control. The ARP poisoning feature caught the username and password when a client computer authenticated with the target IP surveillance system. There were 9,652 packets transmitted between the target IP surveillance system and its client were captured. The two MD5 hash values captured when logging into the target IP surveillance system from Windows Explorer browser, were sent to a hash value cracker – *ophcrack* to decrypt the hash values. *ophcrack* is a free open source program that cracks hash values, and Windows log-in passwords by using Lan Manager hash (LM) through a rainbow table. After entering both captured MD5 hash values into Wireshark, it returned the results as "*empty*". Thus, *WireShar*k did not capture any packets related to the user name and password in either clear text or hash values. Thus, we required the alternative software for hash value cracking and to gain the user name and password for the target IP surveillance system. Two cracking techniques were used, namely: brute-force and dictionary.

To identify the range of vulnerabilities a IP camera may have we used many cracking tools including Hydra. It is also called THC-Hydra, and is a command-line-based network logon cracker that can use a dictionary attack to decrypt passwords from many protocols and applications. Before using Hydra to run a dictionary attack on the target IP surveillance system, we needed to generate a word list. Based on the previous research reviewed, we formed a dictionary of possible default passwords, including *admin* as a common user name. Previous studies show that *GeoHttpServer* have several vulnerabilities; and HTTP header contains much useful information. Thus, we used these clues to run *Hydra* with the *http-head* command (Figure 5).

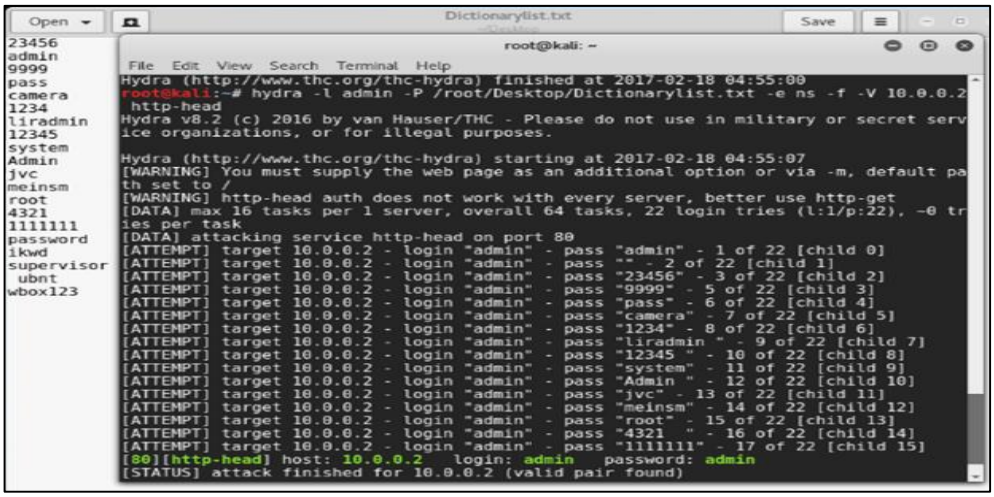


Figure 5. Hydra Http-Head Dictionary Attack Result

The results showed, 17 tries on 22 possible passwords; and one pair valid user name and password found. To confirm the result, we used the identified user name and password to login on the target surveillance system through Windows Explorer. The result confirmed they were correct. *Hydra* with *http-get* command was run to compare the results.

DVR is the heart of IP surveillance system network and has a weak default password. Therefore, we evaluated how well the target IP surveillance system can resist such an attack. *Metasploit* was chosen for the task of developing and executing exploit code against the remote target machine. The results showed the attacking computer was not able to establish connection with 10.0.0.2 on port 5920 - the port used by most IP surveillance systems. We also tried the ports 4550, 5550, 6550 and 10000, which are the system's data port, audio port and VSS port. Metasploit did not provide the option for a user to specify which port to exploit so we tried other tools. Nikto was used to perform web server scanning on the target IP surveillance system. *Nikto* is an open source web server vulnerability scanner, which performs comprehensive tests against web items, including over 6700 potentially dangerous files/programs, checks for outdated versions of over 1,250 servers, and version specific problems on over 270 servers. It also checks for server configuration items such as the presence of multiple index files, HTTP server options, and will attempt to identify installed web servers and software. The following vulnerabilities were identified:

- *The anti-clickjacking X-Frame-Options header is not present*
- *GET The X-XSS-Protection header is not defined. This header can hint to the user agent to protect against some forms of XSS*
- *GET The X-Content-Type-Options header is not set. This could allow the user agent to render the content of the site in a different fashion to the MIME type*
- *OSVDB-2119: GET/shopexd.asp?catakigud='42:VP-ASP Shopping Cart 5.0 contains multiple SQL injection vulnerabilities. CVE-2003-0560, BID-8159*
- *OSVDB-3092: GET /htpasswd: This might be interesting...*
- *OSVDB-3268: GET /tmp/: Directory indexing found.*
- *OSVDB-3092: GET /tmp/: This might be interesting...*
- *OSVDB-3268: GET /images/: Directory indexing found*
- *OSVDB-3268:GET /images/?pattern=/etc/*&sort=name: Directory indexing found*

Another tool used was the *Burp* suite, which is a Java based software platform of tools for performing security testing of web application. The suite of products combines automated and manual testing techniques and consists of a number of different tools, such as a proxy server, web spider, scanner, intruder, repeater, sequencer, decoder, collaborator, extender, and to brute force a login page. After installing the attacking device with *Burp*, Internet

Explorer is then configured to work with *Burp*. It can operate as MITM between the web browser and the target IP surveillance system web server, and it intercepts the traffic exchanged between the browser and the server. *Internet Explorer* (*IE*) was used to connect to the server and enter the correct user name and password. The interception and capture of the POST request gave the username and password that is supplied to the server. This can occasionally be a GET request also. The result shows that both the username and password are MD5 hash values (Figure 6).

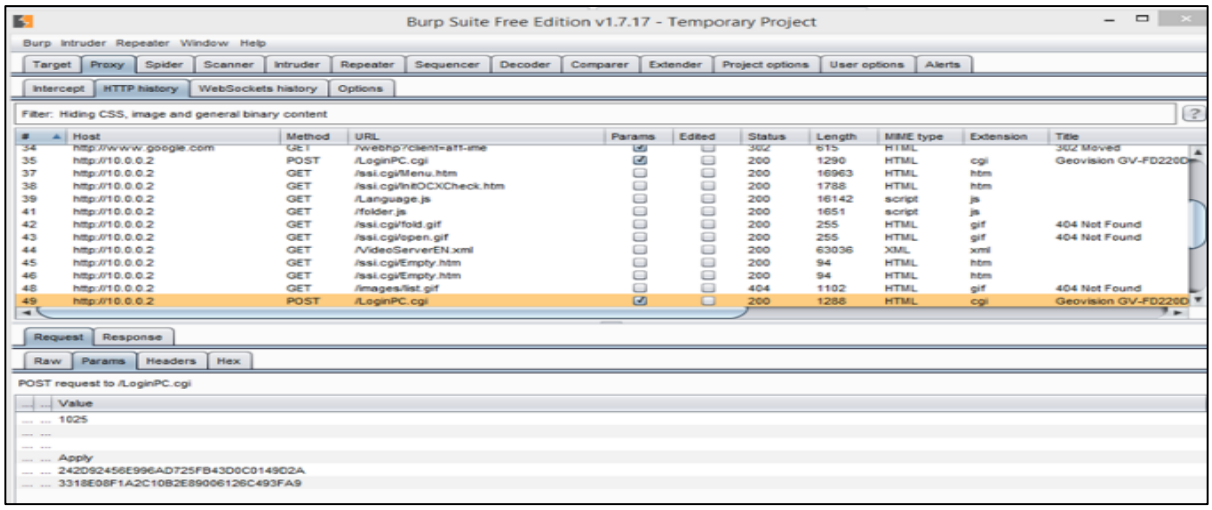


Figure 6. Burp Network Packet Capture Result

We then used *Intruder* and *Sniper* to perform dictionary attacks on the MD5 hash values captured. Instead of trying to decrypt the captured hash values, *Intruder* and *Sniper* allows the attacker to perform a dictionary attack precisely on the captured MD5 hash value fields. The dictionary attack is performed, and a valid user name and password will be shown. Based on the word list used for the attack, there are 27 words used, 54 requests tried and responses in total. Table 1 summarises the vulnerabilities identified in the research and the software used.

Table 1 Summary of vulnerabilities

Software	Functionality	Attack
Pilot Study		
Windows Explorer	Get Camera IP address	Vector
GeoVision DMMultiView	Connect the camera's DVR	Vector
GvIP Device Utility	Manages IP camera, firmware updates, IP addresses within a local area network (LAN), backup records and restore settings	Vector
Main Study		
Angry IP Scanner	Collects IP address, media access control address (MAC), manufacturer and server information	Reconnaissance
WireShark	Collects IP address, media access control address (MAC), manufacturer and server information; packet capture	Reconnaissance
ophcrack	Hash value cracker	Analysis
Burpsuite	Performs security testing of web applications	Reconnaissance
Cain & Abel	Recovers passwords, cracks encrypted passwords using dictionary, brute-force and cryptanalysis attacks, recovers wireless network keys, passwords and routing protocols	Analysis
Nmap	Collects IP address, media access control address (MAC), manufacturer and server information, and system characteristics	Reconnaissance
Hydra (THC-Hydra)	A network logon cracker that can use a dictionary attack to decrypt passwords for many protocols and applications	Analysis
Nikto	Web server scanning for the target IP surveillance system	Reconnaissance
Metasploit	Develops and executes exploit code against a remote camera	Active agency

CONCLUSION

In this research, we tested an out of the box GeoVision GV-FD220D 2MP H.264 IR fixed IP Dome camera for security vulnerabilities. Although the code injection and directory traverse exploitation techniques were rebuffed, many other points of vulnerability were identified. The two points of entry to the camera system were openly accessible through Windows Explorer or the GeoVision DMMultiView client. The password to the system was easily cracked (the factory default) and the GvIP Device Utility entry gained to control the IP camera. A fuller exploration of the whole surveillance system demonstrated the scope of a number of tools and the ability to gain control of critical information. Countermeasures are required to protect the IP camera from hacking and exploitation of the communication resources. Strong advice is to change the access password from the default, and then to change the password regularly. Detection of surveillance activity is required on a moment-by-moment basis and layers of protection are required to satisfy an attacker but also to maintain system integrity. Similarly, critical information requires encryption, protection by tunnelling, and cryptographic complexity to confuse analysis. The defeat of active agency can come by change management controls, benchmark auditing on a moment-by-moment basis, and the regular updating of IP Camera anti-virus software. Our research suggests that IP cameras are vulnerable to exploitation and we advocate for a greater urgency in distributing countermeasures.

REFERENCES

Bojinov, H., Bursztein, E. & Boneh, D. (2009). XCS:Cross channel scripting and its impact on web applications. The 16 ACM Conference on Computer and Communication Security (pp. 420-431). Chicago, IL, USA.

Bruschi, D., Ornaghi, A. & Rosti, E. (2003). S-ARP: a secure address resolution protocol . The 19th IEEE Annual Computer Security Applications Conference (pp. 66-74).

Bryman, A. (2012). Social research methods. Oxford: Oxford University Press.

Caputo, A. (2014). Digital video surveillance and Security second edition. London: Elsevier.

CCTV Calculator. (2017). Vulnerability database. Retrieved from CCTV Calculator: https://www.cctvcalculator.net/en/knowledges/vulnerability-database/

Coole, M., Woodward, A. & Valli, C. (2012). Understanding the vulnerabilities in Wi-Fi and the impact on its use in CCTV systems. The 5th Australian Security and Intelligence Conference (pp. 36-43). Perth, WA, Austrailia : Edith Cowan University.

Costin, A. (2016). Security of CCTV and video surveillance systems; Threats, vulnerabilities, attacks, and mitigations. The 6th International Workshop on Trustworthy Embedded Devices (pp. 45-54). Vienna, Austria: ACM.

Gotham Digital Science. (2012). Using metasploit to access standalone CCTV video surveillance systems. Retrieved from Gotham Digital Science: https://blog.gdssecurity.com/labs/2012/5/15/using-metasploit-to-access-standalone-cctv-video-surveillanc.html

Kirk, J. (2016). Security camera riddled with flaws that let attackers hack your video and your network. Retrieved from PC World: https://www.pcworld.com/article/3030014/security/study-of-another-ip-camera-reveals-serious-problems.html

Kyaw, A., Tian, Z. & Cusack, B. (2016). Wi-Pi: a study of WLAN security in Auckland City. International Journal of Computer Science and Network Security, 16(8) 68-80.

McKee, D., Clement, S., Almutairi, J. & Xu, J. (2017). Massive-scale automation in cyber-physical systems: Vision & challenges. IEEE 13th International Symposium on Autonomous Decentralized System (pp. 5-11). Bangkok, Thailand.

Minin, V. (2015, June 10). GeoVision (GeoHttpServer) webcams - Remote file disclosure. Retrieved from Exploit Database: https://www.exploit-db.com/exploits/37258/

Őzkan, S. (2016). Geovision: Security vulnerabilities. Retrieved from CVE Details: https://www.cvedetails.com/vulnerability-list/vendor_id-2385/Geovision.html

Smith. (2013). Hacks to turn your wireless IP surveillance cameras against you. Retrieved from CSO Online: http://www.networkworld.com/article/2224469/microsoft-subnet/hacks-to-turn-your-wireless-ip-surveillance-cameras-against-you.html

Tekeoğlu, A. & Tosun, A. (2015). Investigating security and privacy of a Cloud-based wireless IP camera: NetCam. The 24th IEEE International Conference on Computer Communication and Networks (pp. 1-6), Las Vegas, NV, USA.

Tian, Z. (2014). Digital forensics in the cloud: Encrypted data evidence tracking. Auckland, New Zealand: Auckland University of Technology.

Wu, H., Ding, Y., Winter, C. & Yao, L. (2010). Network security for virtual machine in cloud computing. In Proceedings of The 5th IEEE International Conference on Computer Science and Convergence Information Technology (pp. 18-21).

Zanella, A., Bui, N., Castellani, A., Vangelista, L. & Zorzi, M. (2014). Internet of things for smart cities. IEEE Internet of Things, (1), 22-23.

TONGA'S ORGANISATIONAL VULNERABILITY TO SOCIAL ENGINEERING

Raymond Lutui[1], Viliami Fe'aomoeata[2]
[1]Digital Forensic Research Laboratories, Auckland University of Technology, Auckland, New Zealand
[2]CUP Research Institute, Christ's University in Pacific, Nuku'alofa, Tonga
rlutui@aut.ac.nz, vfeaomoeata@gmail.com

Abstract

Tonga is a small developing island in the south pacific and ICT is still in its early stages. In this paper we ask the questions, what is social engineering and who is this social engineer, what are the threats to Tonga, how can these threats be identified and which countermeasures can be taken to mitigate the risk of social engineering? The answers to these questions will lead to a social engineering risk management framework to make the risks of social engineering more transparent and help organisations implement mitigating controls against social engineering. The study was performed in four chosen organisations in Tonga, who were involved with Information Communications, Finance, and Cyber Security in order to model threats and countermeasures and develop a risk management framework.

Keywords: Risk management; Social engineering; Information security; Cyber security; Organisational Vulnerability; Security threats; Threat assessments.

INTRODUCTION

The technical aspects of information security have been in the spotlight for several years (Solomon and Chapple, 2005, p.56), and has made much progress. In general, large improvements in security can no longer be attained by upgrades in hardware or software. It is therefore difficult for attackers to achieve their goal through technical attacks alone and their focus shifts (even more) to the organisations employees (Richards, 2008, p.41). As a result, organisations need to direct increased attention toward the undertreated human factor of information security to guard and stay in control of their critical information. For many organisations, the weakest link in information security is now human (Mahfuth et al., 2017, p.1). Organisations need to raise the security on this human factor to an even par with the technical security Legg et al., 2015, p.1). In response, information risk management the top training priority for Information Technology security professionals (Luiijf, 2012, p.57). Organisations are looking to develop flexible frameworks that give insight to the risks involved and help them adapt to changing environmental factors.

Although there have been studies conducted on the human factor of Information Technology, it is still a relatively unexplored field of scientific research. In most cases, the literature does not have a scientific foundation and does not give a clear overview but merely discuss case descriptions (Tsohou et al., 2010, p.227). However, all of the previous mentioned studies show that the human factor can cause great damage to organisations, not only financial but, also to the organisation's image, which in turn influences the organisations goals and continuity in the long run (Drevin et al., 2006, p.448).

Ironically, employees are not only important assets, but also pose a great threat. Employees not only know where to look but have the advantage of obtained trust and accessibility to systems (Nurse et al., 2014, p.271). Attackers can misuse the employees or could even be one of them. There have been known cases of technical hack, and the most notorious human hacker, Chris Hadnagy, asserts that breaches start with a phishing email or vishing call, then they go to a technical hack (Shin, 2017, p.1). This study will primarily focus on the threats from external parties, however, also internal threats and culminate in a high level social engineering risk management model. This can be used to gain transparency on the subject, implement mitigating controls, and help organisations manage their social engineering risks.

This study focuses on *'social engineering'*, the manipulated compromise. Mitigating the threats of this manipulation will also reduce the intentional and unintentional compromising of systems and information therefore, lower overall risk. While this research hopes to provide incentives that may help to ensure business

continuity and give organisations in Tonga a clear view on social engineering, it also aims at finding out how to strengthen the weak link in information security, the human factor, by looking at:

"How social engineering occurs in organisations?"

The measures that can be used to stop social engineering from causing harm. How an organisation can measure the risks and their protection from social engineering threats and if necessary apply appropriate countermeasures to mitigate these risks and stay in control of their information, thus ensuring business continuity.

LITERATURE ANALYSIS

Social engineering has been defined as the unauthorised acquisition of sensitive information or inappropriate access privileges by a potential threat source, based upon the building of an inappropriate trust relationship with a legitimate user (Dudek, 2006, p.1). That is, pretending to be someone you are not, with the goal of misleading someone into giving out information they should not give. Social engineering is an aspect that involves both intellect and technical experience, but more importantly it is an evolving phenomenon that needs to be monitored constantly. If attackers are willing to be consistent in finding loopholes in the system, then security experts should balance and overcome that attempt. With that thought in mind, the basis of this literature analysis is to find out what is trending in not only the cyber world, but also to assess the status of organisations in Tonga with regards to social engineering security.

As a result, it is evident in the literature that a number of researchers have invested time and resources into exploring social engineering. In addition, they look for the latest techniques. However, despite the vastness of the exploration, this analysis will focus on a certain number of key elements that relates to the area of interest.

Hackers and Social Engineers

Hacking and social engineering are closely related. Social engineering tactics are applied to gather information in preparation of a hack and the motives and goals of both types of attacker are related (Ziccardi, 2013, p.75), as social engineers are also known as '*people hackers*'. It is therefore important to know who these (people) hackers are (Warren, & Leitch, 2010, p.427). In this section, a description of hacking and the hacker will be given along with the motives a social engineer may have.

Hackers, Crackers and Phreakers

There are hackers with good intentions. For instance, searching for vulnerabilities in the information system so they can be controlled. There are also hackers with bad intentions, using the identified vulnerabilities for personal gain. There are three types that all get the predicate 'hacker' in the media; hackers, crackers and phreakers (Milberry, 2012, p.112). The jargon dictionary defines a hacker as: "A person who enjoys exploring the details of programmable systems and how to stretch their capabilities, as opposed to most users, who prefer to learn only the minimum necessary" and "one who enjoys the intellectual challenge of creatively overcoming or circumventing limitations" (Watson, 2012, p.260). A hacker is therefore someone who seeks challenges and overcomes boundaries using his or her skills. Hackers follow an ethical code and do not act illegally, which differentiates them from the crackers (Long & Wiles, 2008, p.104)

A cracker is someone who breaks into the system with the goal of theft or vandalism and therefore does not act ethically (Rahalkar, 2016, p.90). Crackers form small groups within the hacking community and are seen as 'a lower form of life' by other hackers (Voiskounsky and Smyslova, 2003, p.178). Another name for these crackers is 'dark-side hackers' (Svensson, 2016, p.90). Finally, phreakers use information and social engineering skills to break into telephone systems and use these for various purposes such as, making long distance phone calls at another's expense, stealing phone card numbers or pretending to call from a secure location. People hackers - in contrast to technical hackers - focus on the weaknesses in the human, instead of the technology they use. The people hackers referred to in this study all have malicious intent and could therefore be classified as 'people crackers' according to previous classification (Barghuthi & Said, 2014, p.2). For the purposes of this paper, the term hacker will imply to those hackers, or crackers working with malicious intent (Richards, 2008, p.40).

SOCIAL ENGINEERING ATTACKS

Knowing why social engineers might attack is crucial for estimating the likelihood of a social engineering assault on a specific organisation, and to implement appropriate measures and controls to counter this assault (Lafrance, 2004, p.12). The motivation of different subcultures within the hacking community will now be discussed

followed by the motives of the social engineers. Bodhani (2013) identifies four subcultures within the hacker community, each with different motivation; casual hackers, political hackers, organised crime, and internal agents (p.65). There are also hackers that do not act as a member of a subculture. The Australian government performed research on the personal motives of a hacker (Madarie, 2017, p.80) such as monetary gain, intellectual challenge, power, and so on. The motives of the social engineer can be classified according to a variation on the results of this research (Krone, 2005, p.2). For each category, a general description of the motive is given, a classification in malicious or good intentions, and what role social engineering can play in an attack with this motive.
The way social engineering can be used in an attack is subject to the goal of the attack. If the goal is to acquire specific information, social engineering can play a great part in the attack. But the main challenges taken up by attackers are still technical; in most cases therefore social engineering will be used to gather information and prepare for the final attack. To stop the social engineer from succeeding, organisations need to apply measures to counter the social engineering attacks and tactics. They can change the environment of the asset, they can choose to act on occurring attacks or they can mitigate the social engineering risk by the structured implementation of countermeasures (Smith et al., 2013, p.250). This study focuses on transparency of social engineering and therefore on the structured implementation of countermeasures. Also, the information security controls, which encapsulate several measures to mitigate the social engineering risk will be classified and listed. After which the key elements pertaining to the human factor are discussed in more detail.

Information Security Controls

In order to secure organisation's data, certain controls must be in place. There are several proposed classifications found in the literature however, Harnesk & Lindström (2012) defines the three most cited dimensions - confidentiality, integrity and availability (CIA) (p.80). *Kind of measure*; physical, logical or organisational. *Moment of action*; corrective, repressive, preventive and detective. Most other models do not classify the reason of protection because, social engineering threatens the confidentiality, integrity and availability and the applied controls need to protect against all of these (Luiijf, 2012, p.56). The classification is based in part on a classification by the National Institute of Standards and Technology (NIST) and complemented with input from the IT Infrastructure Library (ITIL), they both classify the controls on two dimensions. The classification proposed here also consists of two axes, the first according to the *function of control*, the second according to the *level in the organisation* (Dempsey et al., 2011, p.5).

Function of Control

The function of a control is related to its place and effect in the security management process. (Tse, 2004, p.1507). The ITIL classification for security management is used to complement and add an extra level to the process (McPhee, 2008, p.5), defined by the NIST as this only discusses a limited number of functions. Security controls are safeguards or countermeasures employed in order to avoid, detect, or minimise security threats or risks to information, computer systems, or other assets (Tayouri, 2015, p.1098). These controls can be classified based on several criteria. For instance, the time they act, in relation to a security incident: Before the event, ***preventive controls*** are employed to prevent security incident from occurring such as, locking out unauthorized intruders; if a security incident occurs; during the event, ***detective controls*** are intended to identify and characterize an incident in progress such as, by sending out an alert; after the event, ***corrective controls*** are in place to limit the extent of any damage caused by the incident e.g. by recovering the organisation to normal working status as efficiently as possible (Whitman, 2004, p.52). To support the specific controls against social engineering some *general security controls* need to be implemented. These form a base for the more specific controls and will probably -in part- be implemented already to protect other assets from other forms of attack.

DESIGN OF THE STUDY

A guidance of a methodology is highly recommended to maintain the integrity of the findings. Information security experts are aware of social-engineering threat but to date have never seemed to focus their efforts on studying and understanding in depth how and why cyber criminals are using social-engineering method as a weapon (Alexander, 2016, p.2). Yildiz (2007) argued that, some research suffers from definitional vagueness of its concept (p.647). A researcher has to decide the type of research to be conducted in order to answer the pivotal research question that will disclose new knowledge. Exploratory research, on the other hand, is employed in this type of study as it allows the researcher to gain a deeper understanding of an issue or problem (Straub, et al., 2004, p.63).

Due to the fact that this topic has not been explored in depth and never in Tonga, Exploratory Research approach is employed to guide this study. An exploratory study is a valuable means of finding out 'what is happening; to

seek new insights; to ask questions and to assess phenomena in a new light (Saunders et al., 2012, p.139). Exploratory research design does not aim to provide the final and conclusive answers to the research questions,

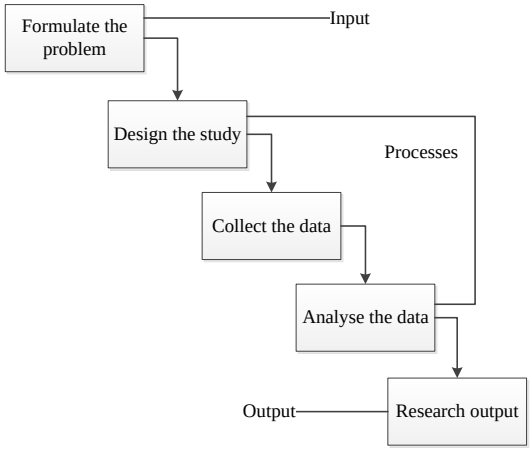


Figure 1: The design of the study

but merely explores the research topic with varying levels of depth but to help to give a better understanding of the problem (Singh, 2007, p.38). Unstructured interviews are the most popular primary data collection method with exploratory research (Sreejesh, et al., 2014, p.47). The interviews were held using a leading questionnaire of open questions. The questionnaire consists of the following stages as represented in figure 2:

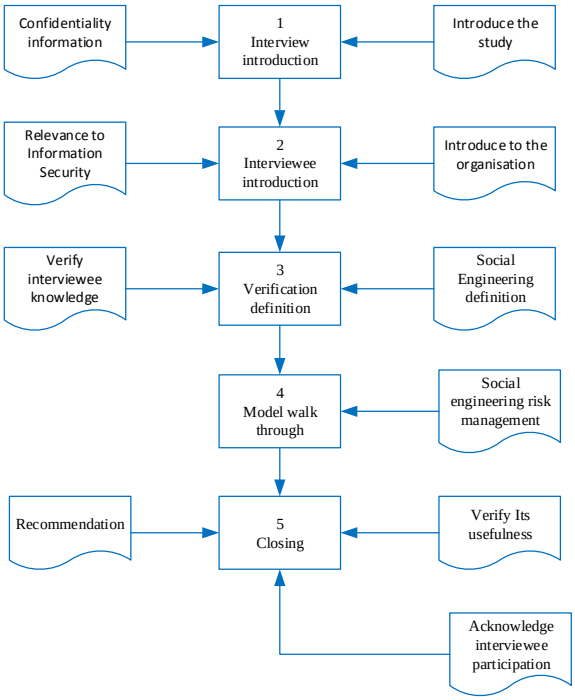


Figure 2: The questionnaire steps

The questionnaire structures the interview but still makes it possible to go deep enough to answer the research questions with a clear foundation.

Case Selection

The interviewed organisations were chosen based on the reliance of their business on information and IT and the level of risk of a social engineering attack. The interviews have therefore been held with an international IT service organisation due to its business focus on information processing and storage for external parties; a consulting organisation whose greatest assets are its personnel and knowledge; a regional governmental organisation due to its increased risk to a social engineering attack; and the Computer Emergency Response Team of the Tongan government (CERT) as the focus of this organisation is on the cyber security within the government by coordinating IT security incidents, informing and advising on these incidents and supporting the governmental

organisations in the prevention of, and response to security incidents. All these organisations have a different perspective on information, its value, the risks they run and possible counter measures. Within the visited organisations the interviews where held with security officers and/or other security responsible personnel. Together these interviews represent a valuable perspective on social engineering as these organisations and specific interviewees should be the ones at the forefront of information protection from for example social engineering. Next to these interviews, the opportunity presented itself to discuss this matter during a cyber security seminar held at Tonga National Centre followed by discussion between the security representatives of several governmental organisations, as well as organisations from the private sector.

RESEARCH FINDINGS

The findings from the interviews have been de-identified and have only in part been related to the organisations or market. The confidential use of interview findings was a precondition for cooperation of the organisations as the provided information could be used in identifying participating organisations and vulnerabilities within these, which is not the intention of this research. Therefore, diagrams and organisational descriptions cannot be made any more detailed. The following findings are related to the organisations activities and are structured according to the stages of the questionnaire followed by relevant comments not directly related to the questionnaire.

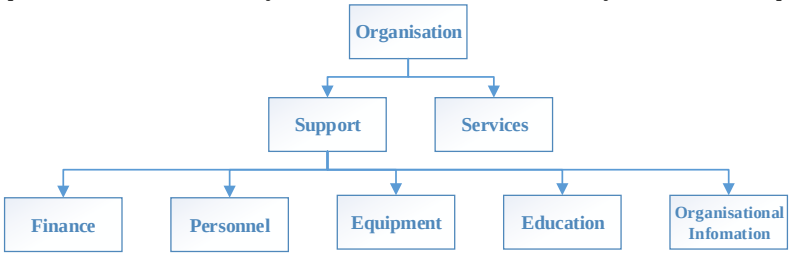


Figure 3: Organisation's activities

The organisation is divided in two parts, one executive in which the main activities are performed, the other supporting in which administrative activities are performed. The interview clearly focused on the organisational information division. The interview was held with the Chief Company bureau services, responsible for the organisational information.

Social engineering risk assessment model

The interviewee was not familiar with the term 'social engineering' but did recognise the description and examples. During the interview, the stated definition was used as reference.

Organisational Description

As the interview focused on the 'organisational information', the organisational description follows. The workplace and workstations are not related to the functions except for data mining. Access is not restricted to the local environment; however private use is restricted. Internet access is only available when necessary for the role or function an employee performs. The functions can be divided in three groups; primary, supporting and management. Authorisations are granted on a need to know basis and related to functional profiles. More authorizations may be provided on request. Segregation of duties is implemented within and between the functions.

Threat Identification

The organisation handles highly sensitive information, which is of great interest to criminal organisations as well as curious social engineers and hackers. But in general, all information in the organisation is of interest and can be of use to the social engineer. A short list of threats where identified, detailed threats more specific to the organisation have not been listed:

- Internal reports do not follow a workflow and can be anywhere on the work floor.
- Not all information is classified and can therefore be handled improperly.
- Access is logged. However, it happens that people log on to another's profile or use another's password.
- Some external parties need access to the system before they can be screened. However, these persons should be under supervision constantly.
- It is possible to intercept classified communications.
- People working at home create a threat.

Vulnerability Identification

Some vulnerability can be derived from the threats:

- It is not known where information is during processing; there is no accountability.
- Classification procedures are not followed.
- The password security policy is not followed.
- The authorisation process is not suitable for some activities.
- Some means of communication are not secure; however, they are necessary for operations.
- Procedures for media usage are not followed.
- People do not follow the information security procedures outside the office.

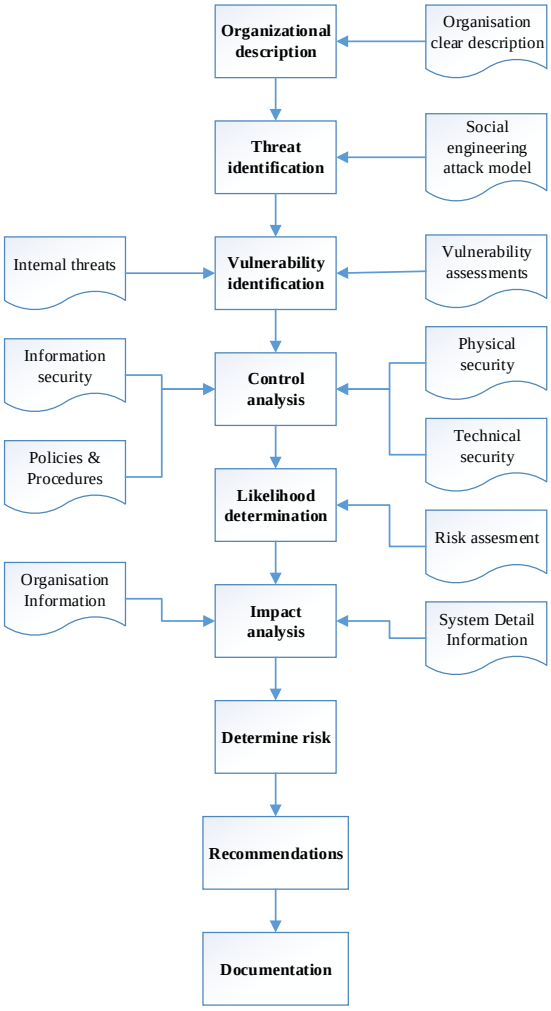


Figure 4: Social engineering risk assessment model

Prior to the interview, this organisation was faced with bad media coverage due to leakage of information after careless handling. The organisation thereafter performed a specific assessment on the information crossing the organisational boundary, this lead to the implementation of specific controls to counter this vulnerability.

Control Analysis

Some examples of controls are - Policy is implemented to stop information from crossing the organisational boundary. USB ports are generally disabled, Information that does need to leave the organisation on a memory stick or over the internet is secured through encryption, There is an awareness project that relates to the awareness of the information you use and training in how to use this so it stays secure, Leakage through personal contacts is traced and measures are taken if necessary, There are heavy penalties on deliberate leakage, Physical access is restricted through specific measures and the last one is, Penetration tests are performed, focusing on technical hacking through for example WIFI connections and Smartphones. But also, social engineering is tested through physical penetration testing and desk sniffing. The findings from this are used to confront people during the awareness trainings.

Likelihood Determination

There is a fair likelihood a social engineer can gather information from this organisation. However, more critical information will be less likely to leak due to the need to know basis on which it is spread through the organisation. In contrast, some threats on less critical information are simply accepted. So, the likelihood depends on the information and cannot be determined in general.

Impact Analysis

There are two general consequences of a successful social engineering attack;

- The image of the organisation can be harmed.
- The organisational processes and even people can be harmed.

Determine Risk

Even though awareness training and penetration tests are implemented there is still some social engineering risk. The organisational process sometimes prevails over the risk of leaking information. However, the risk is still present due to careless personnel.

Recommendations

Organisations need to follow a security management process consisting of a policy statement, followed by awareness, in turn followed by audits. In discussion with the interviewee, the following controls where identified which were already implemented in part - Authorisation management should be implemented, Physical access should be restricted through for example access gates, Data should always be classified, Physical pieces of information should be kept behind locked doors or in a vault, Server rooms should also be locked and hard disks with confidential information should be locked up and finally, Audits should be performed on the adherence to policy and procedures.

A general conclusion was that people see the world around them in which information is stolen, however they do not see the need to be careful with information they handle. Awareness training is required to remove this misconception.

DISCUSSION

To solve the research problem three main research questions where stated, the deliverables related to these questions will now be discussed to see if the research questions have been answered:

Which risks do organisations run as to social engineering?

To be able to identify social engineering risks the definition of social engineering is given in the introduction. Based on the knowledge gained on the findings and discussions, a risk assessment can be made. This *risk assessment* should be performed structurally this is also a component of the social engineering risk management model as in the discussion. When an organisation follows the steps in this model and more specifically the risk assessment this will help them to get a view on their specific social engineering risk. It however cannot give a general risk level, because of the great diversity in organisations.

The actual social engineering risk management model structures the risk management process and generates assurance for the management on their level of control over social engineering consisting of 10 steps - System and environment characterization, Objective setting, Threat & Vulnerability identification, Likelihood determination, Impact analysis, Risk Evaluation, determination & Response, Control analysis & Implementation, Supporting policy and procedures implementation, Information and communication management, Ongoing monitoring and evaluation. These steps can be related to the management process components of the Enterprise Risk Management Integrated Framework (ERM) of the Committee of Sponsoring Organisations of the Treadway commission (COSO) and therefore be implemented as part of this overall management process. The components are: internal environment, objective setting, event identification, risk assessment, risk response, control activities, information and communication and monitoring.

Therefore, this model is relatively elaborate and should be tailored to the organisation and/or incorporated in the organisations ERM process. Based on this social engineering risk management model and the observations from the research conclusions have been drawn and stated.

CONCLUSION

The paper discusses the social engineering risk management using a model in line with Enterprise Risk Management (ERM). The discussion started with the definition of social engineering risk management and its relevance and benefits to organisations; the limitation of social engineering risk in accordance with the organisations objectives. Also, the goal of implementing a social engineering risk management model based on existing risk management models is stated; to assist the organisation in managing the social engineering risk.

In conclusion, the social engineering risk management model could solve the research problem: The model is however still defined on a high-level and application in practice should show the actual usefulness. On this some recommendations for further research are stated.

REFERENCE

Alexander, M. (2016). Methods for Understanding and Reducing Social Engineering Attacks. *SANS Institute InfoSec, 1*(1), 1-34.

Barghuthi, N. B. A., & Said, H. (2014). Ethics behind Cyber Warfare: A study of Arab citizens awareness. *Proceedings of the 2014 IEEE International Conference on Ethics in Science, Technology and Engineering* (pp. 1-7). Chicago, IL: IEEE.

Bodhani, A. (2013). Bad: in a good way. *Institution of Engineering and Technology, 7*(12), 64-68.

Dempsey, K. L., Johnson, L. A., Scholl, M. A., Stine, K. M., Jones, A. C., Orebaugh, A., Chawla, N. S., Johnston, R. (2011). Information security continuous monitoring (ISCM) for federal information systems and organizations. *Special Publication (NIST SP)-800-137*.

Drevin, L., Kruger, H., & Steyn, T. (2006). Value-Focused Assessment of Information Communication and Technology Security Awareness in an Academic Environment. In S. Fischer-Hübner, K. Rannenberg, L. Yngström, & S. Lindskog (Eds.), *Security and Privacy in Dynamic Environments: Proceedings of the IFIP TC-11 21st International Information Security Conference (SEC 2006)* (pp. 448-453). MA: Springer.

Dudek, L. C. & Ruffin, L. K. (2006), *Social engineering & internal/external threats*. Washington: United States Department of the Interior.

Harnesk, D., & Lindström, J. (2012). Materializing Organizational Information Security. In C. Keller, M. Wiberg, P. J. Ågerfalk, & J. S. Z. Eriksson Lundström (Eds.), *Nordic Contributions in IS Research: Proceedings of the Third Scandinavian Conference on Information Systems, SCIS 2012, Sweden* (pp. 76-94). Berlin: Springer.

Krone, T. (2005). Hacking motives: High tech crime brief no. 6. *Australian Institute of Criminology, 6*(1), 1-2.

Lafrance, Y. (2004). Psychology: A precious security tool. *SANS Institute InfoSec Reading Room, 1*(1), 1-32.

Legg, P. A., Buckley, O., Goldsmith, M., & Creese, S. (2015). Caught in the act of an insider attack: detection and assessment of insider threat. *Proceedings* of *the 2015 IEEE International Symposium on Technologies for Homeland Security (HST)* (pp. 1-6). MA: IEEE.

Long, J., & Wiles, J. (2008). *No Tech Hacking: A Guide to Social Engineering, Dumpster Diving, and Shoulder Surfing*. MA: Syngress.

Luiijf, E. (2012). Understanding Cyber Threats and Vulnerabilities. In J. Lopez, R. Setola, & S. D. Wolthusen (Eds.), *Critical Infrastructure Protection: Information Infrastructure Models, Analysis, and Defense* (pp. 52-67). Heidelberg: Springer.

Madarie, R. (2017). Hackers' Motivations: Testing Schwartz's Theory of Motivational Types of Values in a Sample of Hackers. *International Journal of Cyber Criminology, 11*(1), 78-97.

Mahfuth, A., Yussof, S., Baker, A. A., & Ali, N. (2017). A systematic literature review: Information security culture *Proceedings of the 2017 International Conference on Research and Innovation in Information Systems (ICRIIS)* (pp. 1-6). Langkawi: IEEE.

McPhee, D. (2008). *Information Security Management Handbook* (6 ed., Vol. 2). NY: Auerbach Publications.

Milberry, K. (2012). Hacking for Social Justice. In A. Feenberg & N. Friesen (Eds.), *(Re) Inventing The Internet: Critical Case Studies* (pp. 109-130). Rotterdam: SensePublishers.

Nurse, J. R. C., Legg, P. A., Buckley, O., Agrafiotis, I., Wright, G., Whitty, M., Creese, S. (2014). A Critical Reflection on the Threat from Human Insiders – Its Nature, Industry Perceptions, and Detection Approaches. In T. Tryfonas & I. Askoxylakis (Eds.), *Proceedings of the Second International Conference, HAS 2014, Held as Part of HCI International 2014 Human Aspects of Information Security, Privacy, and Trust: Heraklion, Crete, Greece,* (pp.

270-281). Cham: Springer.*Rahalkar, S. A. (2016). Information Security Basics. In Certified Ethical Hacker (CEH) Foundation Guide (pp. 85-95). CA: Apress.*

Richards, G. (2008). Hackers vs slackers - control security. *Institution of Engineering and Technology, 3*(19), 40-43.

Saunders, M., Lewis, P., & Thornhill, A. (2012). *Research Methods for Business Students* (6 ed.). UK: Pearson.

Shanmugapriya, R. (2013). A study of network security using penetration testing. *Proceedings of the 2013 International Conference on Information Communication and Embedded Systems (ICICES)* (pp. 371-374). Chennai: IEEE.

Shin, L. (2017). *Be Prepared: The Top 'Social Engineering' Scams Of 2017*. Retrieved October 6, 2017, from https://www.forbes.com/sites/laurashin/2017/01/04/be-prepared-the-top-social-engineering-scams-of-2017/#4adc4fb7fec1

Singh, K. (2007). *Quantitative social research methods*. New Delhi: Sage Publications.

Smith, A., Papadaki, M., & Furnell, S. M. (2013). Improving Awareness of Social Engineering Attacks. In R. C. Dodge & L. Futcher (Eds.), *Information Assurance and Security Education and Training: Proceedings of the 8th IFIP WG 11.8 World Conference on Information Security Education, WISE 8, Auckland, New Zealand, July 8-10, 2013, WISE 7, Lucerne Switzerland, June 9-10, 2011, and WISE 6, Bento Gonçalves, RS, Brazil, July 27-31, 2009,* (pp. 249-256). Heidelberg: Springer.

Solomon, M. G., & Chapple, M. (2005). *Information Security Illuminated*. USA: Jones and Bartlett Publishers, Inc.

Straub, D., Boudreau, M.-C., & Gefen, D. (2004). Validation guidelines for IS positivist research. *The Communications of the Association for Information Systems, 13*(1), 63.

Sreejesh, S., Mohapatra, S., & Anusree, M. R. (2014). Business Research Design: Exploratory, Descriptive and Causal Designs. In *Business Research Methods: An Applied Orientation* (pp. 25-103). Cham: Springer.

Svensson, R. (2016). Exploiting Vulnerabilities. In *From Hacking to Report Writing: An Introduction to Security and Penetration Testing* (pp. 89-152). CA: Apress.

Tayouri, D. (2015). The Human Factor in the Social Media Security – Combining Education and Technology to Reduce Social Engineering Risks and Damages. *Procedia Manufacturing, 3*(Supplement C), 1096-1100.

Tse, D. (2004). Security in Modern Business: security assessment model for information security Practices. *Proceedings of the Eighth Pacific Asia Conference on Information Systems* (pp.1506-1517). Shanghai: AIS.

Tsohou, A., Karyda, M., Kokolakis, S., & Kiountouzis, E. (2010). Analyzing Information Security Awareness through Networks of Association. In S. Katsikas, J. Lopez, & M. Soriano (Eds.), *Proceedings of the 7th International Conference on Trust, Privacy and Security in Digital Business.* (pp. 227-237). Heidelberg: Springer.

Voiskounsky, A. E., & Smyslova, O. V. (2003). Flow-Based Model of Computer Hackers' Motivation. *Cyber Psychology and Behaviour, 6*(2), 171-180.

Warren, M., & Leitch, S. (2010). Hacker Taggers: A new type of hackers [journal article]. *Information Systems Frontiers, 12*(4), 425-431.

Watson, I. (2012). Digital Underworld. In *The Universal Machine: From the Dawn of Computing to Digital Consciousness* (pp. 259-283). Berlin: Springer.

Whitman, M. E. (2004). In defense of the realm: understanding the threats to information security. *International Journal of Information Management, 24*(1), 43-57.

Yildiz, M. (2007). E-government research: Reviewing the literature, limitations, and ways forward. *Government Information Quarterly, 24*(3), 646-665.

Ziccardi, G. (2013). Hacking and Digital Dissidence Activities. In *Resistance, Liberation Technology and Human Rights in the Digital Age* (pp. 73-123). Dordrecht: Springer.

ASSESSMENT OF SECURITY VULNERABILITIES IN WEARABLE DEVICES

Brian Cusack, Bryce Antony, Gerard Ward, Shaunak Mody
Cyber Forensic Research Centre, Auckland University of Technology, Auckland, New Zealand
brian.cusack@aut.ac.nz; bryceantony2@gmail.com; fx6207@autuni.ac.nz; shaunakmody14892@gmail.com

Abstract

Wearable devices have proliferated in usage and human experience, and they provide convenience for personal information requirements. These devices are both sensory and immersive for the diverse global network that is generally termed the Internet of things (IoT). The immediacy of the two-way communication created in the IoT has made vulnerable human behaviour and raised debate around information ownership and privacy expectations. The legitimacy of ownership of information and its reuse are prevalent problems. In this research, we tested four wearable devices that share 44% of the current market, for security vulnerabilities. We found serious weaknesses that could result in the unplanned disclosure of information and recommend further research into users expectations for safety.

Keywords: Wearable Devices, Vulnerabilities, Privacy, Hacking, Disclosure

INTRODUCTION

Wearable devices have been around for decades in the form of small electronic devices that compensate for human failure in for example, sight and hearing. Wearable hearing aids have been available in different forms for an extended period; however, the technology today is very different in flexibility and functionality from what was available 15 or 20 years ago. Today wearable devices are integrated into a body area network (BAN) for two-way communication and placed into a context that is generally termed the "Internet of Things" (IoT) (PwC, 2016). The purpose of these devices is not only to compensate for human physical and psychological failure, but more commonly to extend the reach of the human through interconnectivity with information sources (NIST, 2010). Consequently, today a human using a wireless Bluetooth earpiece may not be compensating for physical challenges but rather extending their sensory capability by connecting with personal or remote information sources. Developments that are more recent have included devices that monitor personal biological data, geolocation, and emotions. Some of these are used for health purposes, information exchange, and others for navigating around unfamiliar environments. The value of this technological opportunity is a global human experience of interconnectivity that has collapsed the barriers between internal and external environments and provided full personal immersion. In this fashion, a human may experience a fully augmented reality for the betterment of themselves and the systems in which they participate. The simplest functional architecture provides a connection between a wearable broadcast mechanism and a receiving station (Zhou, et al., 2014). Because most people have a smart phone on them most of the time, then the smart phone has become the receiving station for a multiplicity of different devices that the human may carry within their BAN. There are many examples of connectivity that both transmits information and receives information in the two-way relationship between the base station and the broadcast mechanism or sensors. The eyeglass that streams information to the user and directly to the eye has significant publicity. In addition, wearable watches and biometric monitoring equipment such as the Fitbit, provide personal information for decision-making (Burlacu, 2016; Stack, 2015).

The personal nature of the information managed by the BAN has raised the issue of information ownership (Schelleus, et al., 2014). Personal expectations to control the information from a wearable device, such as a Fitbit, may be a foregone conclusion of the user, and yet the patient owners of the device, the owners of the software, and the owners of the cloud services, and others involved in the brokerage and intermediation of services may all assert ownership of the data. The borderless interconnectivity of human and networks also presents inter-jurisdictional challenges regarding ownership of informational properties, and identification of who has the rights of disclosure and transaction. The fundamental principles of security design require the confidentiality, integrity, and accessibility to information. In situations where the user of a wearable device expects the exclusive ownership of the information that they produce using the device (Zhou, et al., 2014), then the confidentiality and the integrity of the information has to be preserved. Our research is concerned the vulnerability of wearable devices to attacks that can disclose the information the device produces. This research is an attempt to satisfy customer expectations for the confidentiality of their information, and the management of unwarranted disclosures. We tested four wearable devices that had 44% of the market share at the end of 2016 (IDC, 2016), for the presence of security

mechanisms that would preserve the confidentiality of information the device produced from the user actions. Each of these devices was presented on a watch strap to be attached to the wrist of the human. Each device had sensors that trapped a range of biometric data from the end user and also had interconnectivity to broadcast that information for processing, archiving, and providing feedback to the user. The majority of the information processing was done by cloud services in the form of historical logs that tracked and kept account of the user biometric data. All of the devices relied upon Bluetooth low energy as the wireless communication protocol for synchronisation back to the user smart phone (Great Scott, 2015). If the user did not wish to use their smart phone in for example, a gymnasium or a motor vehicle, then they could tether the device to the exercise machine, the motor vehicle network, or to any other local wireless network via the Bluetooth low energy connection for the same effect. The Bluetooth wireless and the tethering protocol is vulnerable to attacks that could violate the confidentiality of information, the integrity of information, and the accessibility to information when hijacked and subjected to service disruption (Cyr, et al., 2014).

BACKGROUND LITERATURE AND METHOD

A consistent theme in literature is the security vulnerability during the pairing of the wearable device with the base station. At the point of pairing the exchange of information is vulnerable. When pairing for the very first time Bluetooth employs one of three Secure Simple Pairing (SSP) strategies:

- Just works – pairs automatically as it requires with no user interaction. Convenient for IoT accessibility but the least secure.
- Numeric comparison - When pairing for the first time both the wearable and smartphone display an identical four to six-digit numerical key. If they match, the smartphone prompts the user to accept the connection.
- Passkey entry - both devices have a user interface for entering a four to six-digit code. One, or both devices must enter a passkey to successfully pair. The authors credit this approach as being the most secure (Lotfy & Hale, 2016; Pieterse & Olivier, 2014)

A review of three wearables by Lotfy and Hale (2016) found that the security of pairing strategies had significant gaps and potential vulnerabilities including:

- Man-in-the-middle attacks, eavesdropping, and packet injection. These kinds of attacks allow attackers to actively spy on wearable devices (user-correlation) and misuse the data.

The Pairing processes are defined as:

- Generic Access Profile (GAP) – the wearable defines a specific advertising protocol. This is important in our research design as this happens subsequent to initial pairing during recurring connection instances.
- Generic Attribute Profile (GATT) – service framework on top of the underlying transport protocol, called ATT (Attribute Protocol), which sets the mutually agreed data transfer standard.
- Both GAP and GATT operate in the 2.4 GHz bandwidth, transmitting at a speed of 1Mbit/sec.

While Bluetooth Low Energy (BLE) operates on the same frequencies as other Bluetooth technologies, it operates differently on the link and physical layer. BLE uses 40 total channels; three are used for advertising by unconnected wearables. The remaining 37 channels are used during GATT for data transmission after pairing. The sequential pairing process is shown in Figure 1 (Lotfy and Hale, 2016, p.3).

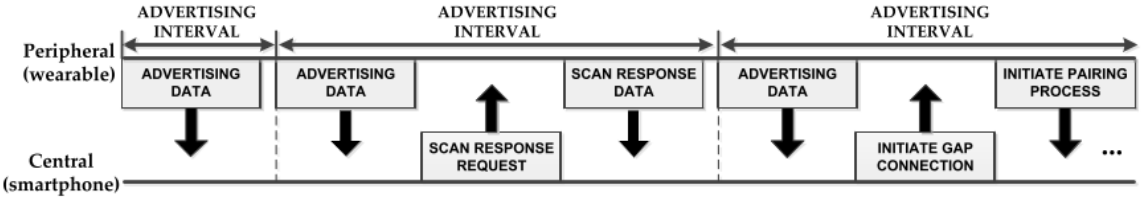


Figure1: GAP BluetoothLE advertising process showing advertising data, scan request, scan response and gap initiation packets

Personally identifiable information (PII) is information that is assignable to a particular human system factor and in some jurisdictions has legal protection (Boyle & Panko, 2014). Broadcasting a fixed MAC address, tied to an

individual's identity would fail the PII test by creating a unique user signature. When wearables create such a risk without user notification, then breaches the privacy has to be considered jurisdiction by jurisdiction. For example, the European Parliament's, Protection of Personal Data Directive, enacted on 5 May 2016 and requiring member states to have introduced into their national law by 6 May 2018, extends the definition of personal data to include that which can "*be identified, directly or indirectly*" (European Parliament, 2016, p. 3). A fixed MAC address, which risks unseen surveillance, breaches this requirement. The literature analysis identified that there are many attack classes (see Figure 2) in and around the use of Bluetooth connectivity. To focus our research and to make it feasible in the laboratory we selected the two attack classes and the five specific attacks highlighted in blue (Hassen, et al., 2017). In respect to Surveillance class, the risk created by a digital signature can be sub-categorised as:

i. **Blue-Printing** – MAC address spoofing for a man-in-the middle attack.

ii. **Blue-Stumbling** - Forced re-pairing attack.

iii. **Blue-Tracking** – a brute force attack designed to determine the data encryption key if one is used.

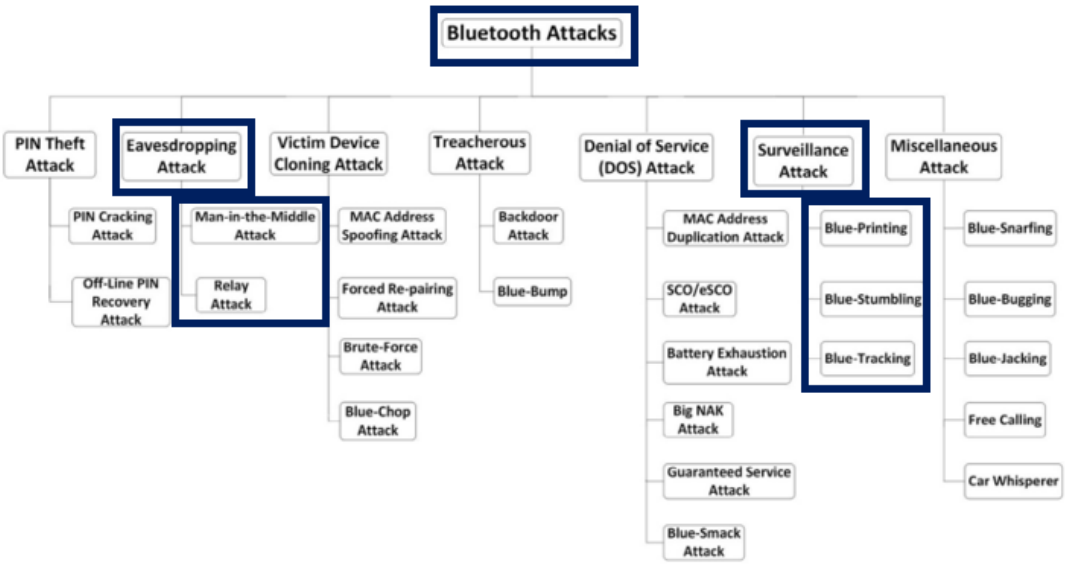


Figure 2 – Classification of Bluetooth Attacks (Hassan et al., 2017, p. 5)

In this research we tested four wearable devices that were selected from the top sales performers in the market and together they held a 44% share of the market at the end of 2016 (Cry, et al., 2014; Guo, 2015). The analysis was conducted in the research laboratory during the first two months of 2017. Figure 3 shows the release dates and devices selected (software versions were those at the time and do not account for any more recent updates).

Reference	Model	Release Date
1	Fitbit Charge	November-14
2	Firbit Surge	January-15
3	Xiaomi Huami	August-16
4	Samsung Gear3	November-16

Fitbit Charge HR	Firbit Surge	Samsung Gear3	Xiaomi Huami, or Amazfit
			
BLE version 4.0	BLE v. 4.0	BLE v. 4.2	BLE v. 4.0

Figure 3. The wearable devices tested

The research was structured to address concerns around device information visibility, pairing visibility, surveillance potential, and information disclosure. The major focus was on potential eavesdropping and surveillance attacks. All of the wearable devices relied upon Bluetooth low energy (BLE) as the wireless communication protocol for data synchronisation between the device and the user smart phone (or other paired network) (Grassi, 2014). The only variation was the Samsung Gear3 that was using version BLE 4.2 rather than

BLE 4.0. The testing with different smart phones was to confirm the consistence of the protocol on different devices. BLE is feasible for use across the Apple iOS, Android, Apple Mac OS, Linux and Microsoft Windows operating systems. In the testing we used a Samsung S6 edge, HTC 1M7, and an Apple iPhone 5S, but did not detect any variations in the BLE protocol execution that related to research concerns. For sniffing tools we selected Ubertooth, the HCI snoop log and the Adafruit sniffer (Lofty, et al., 2016).

The method used two Android phones, and the Adafruit for capture of the BLE packets, and TCP dump in to pcapng and pcap files respectively. The files acquired were then uploaded into Wireshark, an open source packet analyser with a graphical user interface and filtering capability. Using Wireshark the data was examined to determine whether it was transmitted in plain text, or in an encrypted format. Other analysis proceeded to locate any digital signatures, device identification, mappings for the wearer's movements and habits, internet access to device logs and databases, and geolocation correlation data. All the wearables include a back-end cloud service in which an individual's data is stored to ensure portability across devices, and we looked for any credentials providing access rights. The research design is shown in Figure 4, and the blue box indicates the target zone for the sniffing of pairing activity.

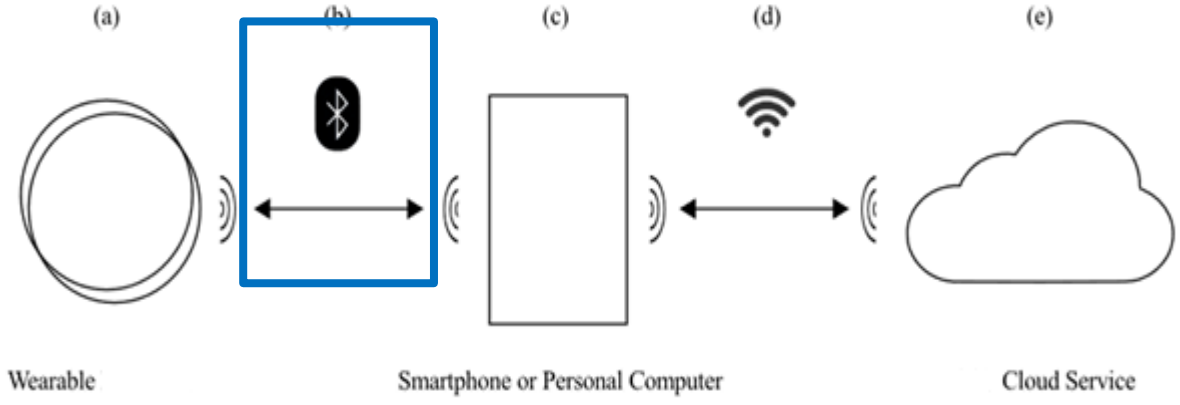


Figure 4. Research Design

FINDINGS

The HCI snoop log is an application that sits on smart phones. It was used in each instance to successfully capture the BLE log file which was then uploaded to Wireshark for analysis. Both encrypted and unencrypted information was found. The unencrypted packets were advertising packets that included the connection request and response. The remainder of the packets were encrypted suggesting that once the exchange protocols and keys have been agreed, all messages were encrypted. We found that the majority of the wearable devices packets were encrypted and the extent of the security mechanisms varied on a device-by-device basis, despite all of them relying upon the BLE protocol. The Adafruit sniffer could successfully follow a device once the connection had been established. The Ubertooth had a similar performance. Using the Adafruit for interception we found that when the wearable device and the base station paired the identity was disclosed in plain text. It may be good for efficiency purposes to have the brand and the watch identity publicly displayed in the wireless network, but for an attacker this is a bonus and makes an easy target. Each wearable device had a different performance and a different susceptibility to attack. Surprisingly the Amazon fit broadcasted the long-term encryption key in plain text during the initial setup as shown in Figure 5.

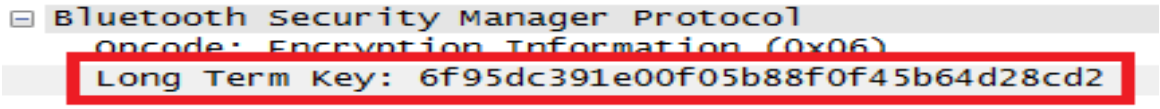


Figure 5. A Plain Text Encryption Key captured

Notably the other wearable devices had encrypted this information. However, the MAC address of the Charge and Surge did not change once connected, an identical result to the Amazon fit. The literature we had read stated that Fitbit had addressed the MAC vulnerability, but our findings suggested otherwise. A fixed MAC address for a session creates the risk of an attack vector based on user correlation or blue tracking. However, the research did

show that some of the previously reported vulnerabilities had been addressed by the end of 2016. For example, the Gear S3 was found to have advanced security mechanisms that were in keeping with the newer BLE version 4.2, and was the most secure. However, our research showed that the HCI snoop log functionality found in most smartphones captured the identity of the message senders, as well as the message in plain text. With this function enabled the smart phones owner's conversations using email, SMS and messaging applications such as Facebook are preserved until a log is deleted. In Figure 6 this vulnerability is illustrated from our experiments.

```
04-22 21:26:05.706  WearableManager.Make extender protocol for
04-22 21:26:05.707  WearableManager.Action Found = Reply to Mom
04-22 21:26:05.707  WearableManager.Action:: Put action for id : 1857
04-22 21:26:05.708  WearableManager.Empty page list
04-22 21:26:05.709  WearableManager.No Pages
04-22 21:26:05.711  ForwardManager.forwardNotification(com.whatsapp), Source = 5
04-22 21:26:05.713  NotificationServiceAPI.getAppNotificationLevel(packageName : com.whatsapp, sourceType : 2)
04-22 21:26:05.717  Config.com.whatsapp id: 2130840151 get resource: android.content.res.Resources@b6eeba1
04-22 21:26:05.778  ForwardManager.info list size = 1
04-22 21:26:05.781  DBMemory.checkDuplicationNotification(1, com.whatsapp, 1492843322000)
04-22 21:26:05.793  DBMemory.[color] createNotificationUnit : 0(0)
04-22 21:26:05.794  ForwardManager.Lock ReleasedDuplicate MEssage Discarded
04-22 21:26:05.794  ForwardManager.Lock released
04-22 21:26:05.823  WearableJsonBuilder.action id : 1857 getIcon :2130840151
04-22 21:26:05.856  [WearableManager][JSON_BUILDER] WExtender JSON : is privacy
04-22 21:26:05.857  Main.WearExtender action found
04-22 21:26:05.859  [Main]jsonObj :  is privacy
04-22 21:26:05.860  Main.Group:: pushSchedulerForParseAndForward() noti : com.whatsapp
04-22 21:26:05.861  ForwardScheduler.pushScheduler : Type: 6
04-22 21:26:05.862  Main.handleMessage()
04-22 21:26:05.863  ForwardScheduler.Got message in scheduler Handler: 6
04-22 21:26:05.864  Main.MSG_NOTIFICATION_43_FORWARD
04-22 21:26:05.865  Config.isSamsungDevice()
04-22 21:26:05.871  ForwardManager.forwardNotification(com.whatsapp), Source = 5
04-22 21:26:05.872  NotificationServiceAPI.getAppNotificationLevel(packageName : com.whatsapp, sourceType : 2)
04-22 21:26:05.877  ForwardManager.info list size = 1
```

Figure 6. Snoop log plaintext disclosures

Overall, the lab testing of these devices shows that manufacturers have made big steps to improve the security around wearable devices. The security improvements in BLE version 4.2 have shut down some of the previous attack vectors and undoubtedly, further improvements are evolving during 2017. In the wearable devices tested, the security vulnerabilities detected indicate the threat classes potentially faced by a user. It is also notable that the different wearable devices have different vulnerabilities but the most predominant issue to date is the disclosure of the MAC, which allows for user correlation and blue printing attacks. The Amazfit performed the poorest out of the four tested. It failed in each of the four threat classes, whereas the Gear S3 with the updated BLE version performed the best in our tests. The implication of these findings is for corporates, such as health insurers, who provide benefits to the customer when they are using wearable devices that have health control feedback loops (MLC, 2017). In the case of a wearable device that is vulnerable to manipulation, the sponsoring corporate may not have confidence that the information they are receiving, and the information on which they will make decisions regarding providing benefits to their customer, can be trusted. In the bigger picture, wearable devices may fail compliance criteria such as the requirements of the European Personal Data Directive. In these situations, the purchaser requires notification in the specifications of the device regarding the security precautions for information protection. There also needs to be independent testing so that the shrink-wrap claims may have some external validation. Figure 7. shows the results from the laboratory testing.

Threat Class	Charge	Surge	Gear S3	Amazfit
Public Name	x	x	x	User Correlation Blue-Printing
MAC	Blue-Printing User Correlation	Blue-Printing User Correlation	x	Blue-Printing User Correlation
Key	x	x	x	Blue-Tracking
Notification	x	x	Breach of PII	Blue-Printing

Figure 7. Summary of Vulnerability

CONCLUSION

Wearable devices are convenient technologies that extend human natural senses and capabilities. Our research shows that further consideration of information protection is required to avoid disclosure failures. The improvement of information security by adopting countermeasures for pairing vulnerabilities will allow the producer of the information choices regarding the control of its ownership. Further research topics arising from this research for future projects are:

The HCI Snoop log paired with the base station – Is the number of log files, and degree of information captured controlled by the base station, the device, or both?

The Decrypt packets of the Amazfit – Future research should confirm that having the long term key is sufficient to decrypt the information exchanged with the smartphone. Also, consumer testing should be broadened to include data synchronisation from the smartphone to the web application.

PII failings – Determine whether the security vulnerabilities breach consumer privacy laws in key markets.

A broader range of attacks – Extend the man-in-middle attacks to intercept and manipulate communications. This has far-reaching implications, not just limited to wearables but also to other IoT setups.

Framework – A compliance framework which brings visibility to data protections in wearables, addresses standards, and reduces industry wide variations.

Further research is required to establish baselines for wearable device user expectations. At present the technology is made functional and accessible to users but we argue that development is required to meet the full scope of socio-technical expectations. Current users want the advantages of the technology, they are using it in increasing numbers, but they also want assurances unwanted surprises of a personal nature will not be forthcoming. Such unwanted attention is unsolicited advertising, personal profiling, geolocation matching, and so on. Our research shows that the confidentiality and potential integrity of data produced by wearable devices tested were easily compromised. More than a regular patch-by-patch updating of software is required to assure users their information safety has been adequately addressed.

REFERENCES

Boyle, R. J., & Panko, R. R. (2014). *Corporate computer security*. Essex, England: Prentice Hall Press.

Burlacu, A. (2016). *Fitbit Tracker Likely Saved This Man's Life, Leading Doctors To Shock His Heart Back To Normal*. Retrieved 20 March 2017, from http://www.techtimes.com/articles/149164/20160411/fitbittrackerlikelysavedthismans lifeleadingdoctorstoshockhisheartbacktonormal.htm

Cyr, B., Horn, W., Miao, D., & Specter, M. (2014). Security Analysis of Wearable Fitness Devices (Fitbit). *Tech.rep.Massachusetts Institute of Technology, 2014, pp. 1-14.*

European Parliament. (2016). *Directive (EU) 2016/680*. Retrieved 22 May 2017, from http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2016.119.01.0089.01.ENG&toc=OJ:L:2016:119:TOC

Grassi, M. (2014). *How to capture Bluetooth packets on Android 4.4*. Retrieved 6 March 2017, from https://www.nowsecure.com/blog/2014/02/07/bluetooth-packet-capture-on- android-4-4/

Great Scott Gadgets. (2015). *Bluetooth Low Energy mode for Ubertooth*. Retrieved15 March 2017, from https://github.com/greatscottgadgets/ubertooth/blob/master/host/README.btle.md

Guo, F. (2015). Securing Wearable Devices. Retrieved 22 May 2017, from http://www.leiphone.com/news/201511/cMxCXDonsugGN892.html

Hassan, S. S., Bibon, S. D., Hossain, M. S., & Atiquzzaman, M. (2017). Security threats in bluetooth technology. *Computers & Security, ,*. doi: 10.1016/j.cose.2017.03.008.

IDC. (2016). *Fitness Trackers in the Lead as Wearables Market Grows 3.1% in the Third Quarter, According to IDC.* Retrieved from http://www.idc.com/getdoc.jsp?containerId=prUS41996116

Layton, J., & Franklin, C. (2016). *How Bluetooth Works*. Retrieved1 June 2017, from http://electronics.howstuffworks.com/bluetooth2.htm

Lotfy, K., & Hale, M. L. (2016). Assessing Pairing and Data Exchange Mechanism Security in the Wearable Internet of Things Symposium conducted at the meeting of the 2016 IEEE International Conference on Mobile Services (MS) doi:10.1109/MobServ.2016.15

MLC. (n.d.). *MLC Life Insurance On Track*. Retrieved 20 March 2017, from https://www.mlc.com.au/personal/importantupdates/ontrack

NIST. (2010). *Guide for Applying the Risk Management Framework to Federal Information Systems: a Security Life Cycle Approach*. Retrieved SP 800-37 Rev. 1, from http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-37r1.pdf

Pieterse, H., & Olivier, M. S. (2014). Bluetooth Command and Control channel. *Computers & Security, 45*, 75-83. doi:http://dx.doi.org/10.1016/j.cose.2014.05.007

PwC. (2015). The Internet of Things: The next growth engine for the semiconductor industry. Retrieved from https://www.pwc.de/de/technologie-medien-und- telekommunikation/assets/pwc-studie-prognostiziert-boom-in-der- halbleiterbranche.pdf

Schellevis, M., Jacobs, B., Meijer, C., & de Ruiter, J. (2016). Getting access to your own Fitbit data.

SIG. (n.d.). *Security, Bluetooth Low Energy*. Retrieved 2 June 2017, from https://www.bluetooth.com/~/media/files/specification/bluetooth-low-energy- security.ashx?la=en

Stack Overflow Community. (2015). *Analyzing Bluetooth Low Energy Traffic*. Retrieved from http://stackoverflow.com/questions/32640581/analyzing-bluetooth-low-energy-traffic

Stallings, W. (2014). *Cryptography and Network Security: Principles and Practice, 6th Edition*: Pearson.

Zhou, W., & Piramuthu, S. (2014, 18-21 June 2014). Security/privacy of wearable fitness tracking IoT devices Symposium conducted at the meeting of the 2014 9th Iberian Conference on Information Systems and Technologies (CISTI) STI.2014.6877073

NEUROSECURITY FOR BRAINWARE DEVICES

Brian Cusack, Kaushik Sundararajan, Reza Khaleghparast
Cyber Forensic Research Centre, Auckland University of Technology, Auckland, New Zealand
brian.cusack@aut.ac.nz, kaushik.sundararajan@gmail.com, khaleghparast@live.com

Abstract

Brainware has a long history of development down into the present day where very simple and usable devices are available to train for the control of games and services. One of the big areas of application has been in the health sciences to provide compensatory control to humans who may lack the usual capabilities. Our concern has been the protection of information in brainware so that a human intention may have confidentiality, integrity, and accessibility to the required implementation mechanisms for services. The research question was: What are the consequences of security failure in brainware? Our research tested a brainware device and found vulnerabilities. The most significant vulnerability was the ability to capture and inject communication packets so that a human intention could be hijacked. The consequences of this communication failure are for psychological harm to the human and unplanned for actions in the material environment.

Keywords: Security, Failure, Brainware, Hijacking, Harm

INTRODUCTION

Electro activity in the human brain has been studied for over a century and various applications devised that enhance human capabilities in areas where capability may be deficient. In particular thinking ability and motor control have been beneficiaries of brainware devices (Allison, et al., 2007; da Silva, 1996). Significant progress has been made from the times when invasive surgical operations were required to insert brainware devices inside a human brain to gain the benefits. Today brainware has become nonintrusive and the latest advancements have dry electrodes that sit on the human head collecting the electro activity of the brain (Wyecoff, et al., 2015). These are significant technological advancements that provide ease of use and ready access for research and learning. Some are woven inside baseball caps and other socially integrated headgear, and the device acts as an inconspicuous aid for enhanced human capability (Bonaci, et al., 2014; Kroeker, 2011). The headsets are also relatively inexpensive and available for purchase online or in gaming and electronics shops. They can be trained to control a wide variety of applications including, model cars, wheelchairs and games (Wolpaw et al., 2002). The simplest ones have a single electrode and minimal control functions such as up, down, left, and right; which is sufficient for a toy or a computer game. Other headsets have 14 and more electrodes and a greatly increased capacity to harness a wider variety of emotions in the human brain and to create a more refined control interface. The use of brainware is relatively simple once it has been trained (Jeunet, et al., 2016; Donoghue, 2002). The training of brainware software is similar to the training of voice activation and transcription software. The user in each situation has to go through a series of standardised algorithms that link the human variability to the standardised software processes. In brainware that is used for playing a game or controlling a wheelchair, the user has to think and not to move or speak. So for example, if I was training my brainware application to steer a remote control car, I would have to continue to think the word "left" until the electro activity in my brain mapped onto the preprogramed software for turning the car left. Sometimes the matching takes longer than others but providing the user is prepared to concentrate and put in the time to train the software, the effects are created by thinking. In the radio controlled car situation, once the brain-ware is trained, then it is possible to put on the headset, look at the remote control car (power on in car) and control its movement up to approximately 3 meters by using the correct thoughts. Similarly, for the training of the control of a wheelchair and other medical applications the user has to spend time synchronising their electro brain activity with the application they wish to use, but once completed the communication is relatively effective (Millan, et al., 2004).

A significant problem for human behaviour and human psychological stability arises once the user has trained the brainware to perform particular functions. If the application does not behave in the ways that it has been trained and the user expectation satisfied, the relationship is destabilised and the effectiveness of the technology undermined. There are several ways that this may occur but our specific research interest was in the situation were the brainware is hacked and unexpected responses to thoughts are presented to the user (Denning, et al., 2004). In this situation many unintended human behaviours may be demonstrated and the purpose of the technological advantage lost. Consequently our research took a brainware device and tested it for security vulnerabilities

(Martinovic, et al., 2012; Li, et al,. 2015). The results show that the communication between the headset and the computer interface or the device has vulnerabilities that disclose information regarding the intended control function and the brain to device mapping. We also performed test attacks to disconnect the thought from its intended action. In this research our objective was to demonstrate the vulnerabilities in the use of brainware, but anecdotally it was obvious the intervention had a negative impact on the user. Our concern is that suitable consideration is given to the securing of the communication between the headset and the devices so that the user intention is conveyed through to the effect. The implication of disruption in the communication channel is for unplanned actions, frustration and potential harm to the user. The consequences may be insignificant when a remote control car is being used for fun, but it is a much more serious case when humans are controlling prosthetic arms, wheelchairs, and sufficing control effects (Wolpaw, et al., 2002; Kroeker, 2011; Lauer, et al., 2000).

Other writers have defined Neurosecurity as the protection of neural devices from adversaries trying to exploit, block, eavesdrop, or generally disrupt neural signals (da Silver, 1996; Darvis, et al., 2004; Nijholt, et al., 2009). Confidentiality is critical in maintaining the privacy of information and it is for the developers to assure that the properties of the device cannot be exploited to disclose signals or any other protected information (Lauer, et al., 2000; Golub, et al., 2016; Li, et al., 2015). Similarly, an attacker should not be able to change device settings or initiate unauthorised operations that compromise the integrity of the device and its information. The availability of the device for clear and intended communication requires strong security measures. Neuro security is consequently the protection of confidentiality, integrity, and availability of the neural devices for the intended user, in such a way that the safety of a person's neural mechanisms, neural computation, and free will, are protected (Millan, et al., 2004). Our laboratory tests on devices suggest that neural security is lagging in some readily available brainware headsets on the market today.

DEVICE TESTING

The brain computer interface (BCI) consists of four components and a connecting signal (see figure 1). BCI requires a human user who has a sensing device that collects the electro chemical energy transmissions from the brain. The sensing device communicates to a signal processing module that puts the sensor signals into a manageable format for transmission either through wired or wireless media. The forth component is an application that will drive an effect, such as movement, decisions, control, and so on.

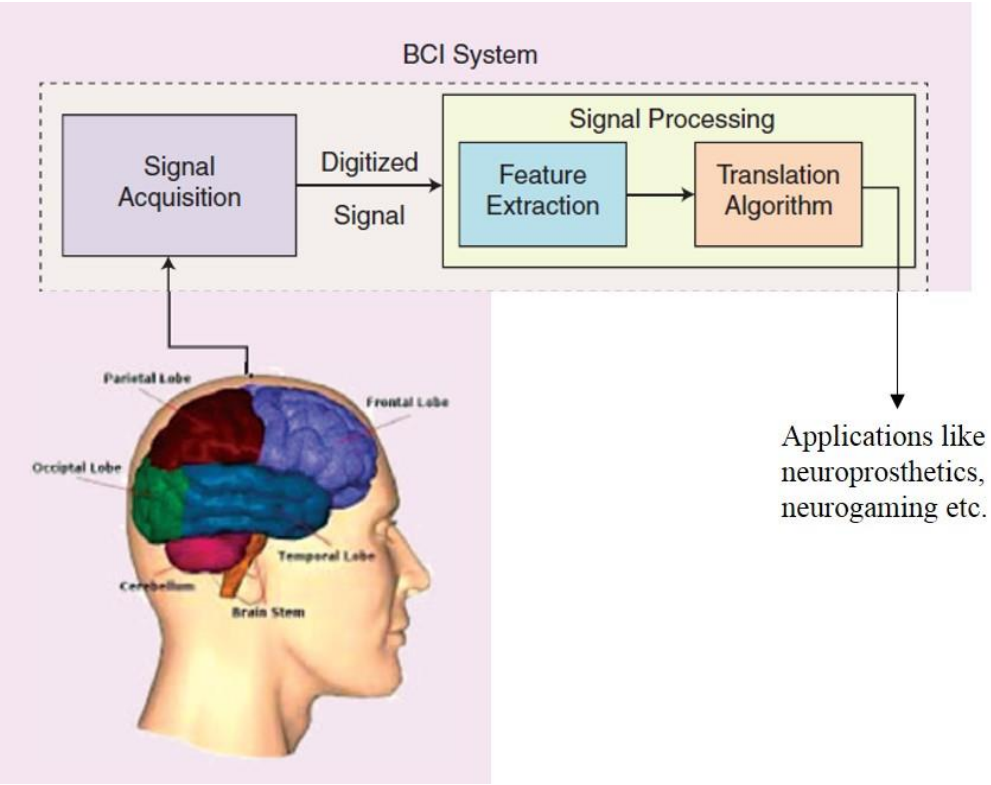


Figure 1. Brain Computer Interface Architecture

We chose the Emotiv Insight 5 channel EEG headset for testing. The Emotiv Insight is designed to detect performance levels of certain parameters that include the human attention level, focus level, engagement level, interest level, relaxation level and stress level. In addition to detecting performance, it also detects mental commands and facial expressions. These expressions include blink, wink, frown surprise, clench and smile. All these parameters are recorded using a computer based interface called Emotiv Xavier Control Panel. The control panel shows signal quality of the brainwear, mental commands, facial expressions, and the inertial sensors. In addition to these features, the control panel also provides connectivity to other platforms of Emotiv for information conditioning. Once all the five channels are green then training can begin, and once trained the headset can be used many times by the same user.

Figure 2. Emotiv Insight 5 channel EEG headset (Emotiv, 2016)

Previous research has established the vulnerability of devices to disclose critical information when tethering in Bluetooth wireless networks (Li, Ding, Conti, 2015). We assumed the vulnerability and only briefly checked the matter to confirm the problem and potential violation of confidentiality. However, this research was concerned with intervention in the communication between the headset and the device. Could we hijack the headset control and substitute alternative commands, unknown to the user; and hence, violate the integrity of the system?

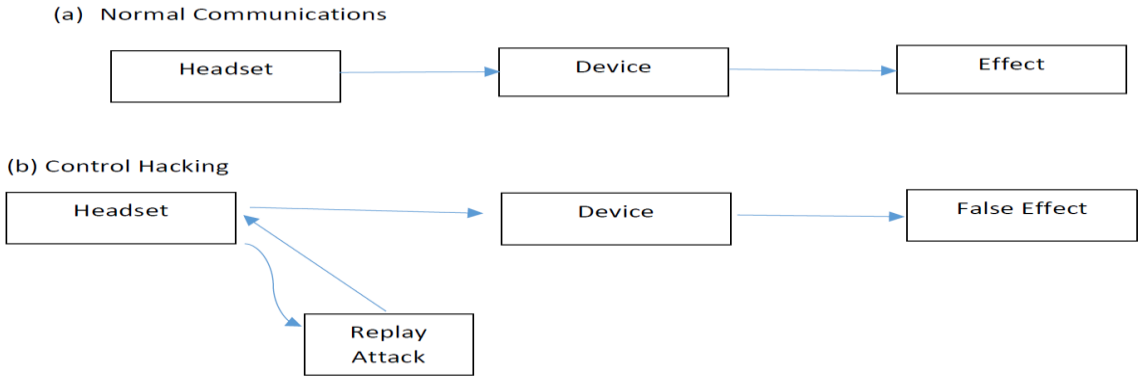


Figure 3. Research Design

Consequently, the Ubertooth-one and Adafruit sniffers were not capable of manipulating and resending changed packets, and hence a framework called 'Btlejuice' was deployed. The framework makes use of external Bluetooth dongles CSR 4.0, creates a clone of the target device, and intercepts the Bluetooth General Attribute Profile (GATT) from the top most layer of the Bluetooth protocol stack. The support software requires setting the target device (the headset), and then double clicking the headset icon so that the Bluetooth dongle proxies the services and characteristics of the Insight headset and pairs with the headset. The packets are then captured on the proxy for manipulation and the system tricked into accepting the proxy communications in a replay attack. Figure 3a

shows the normal information flow between the headset and the device and figure 3b how we hijack and replay fake messages to alter the device effect.

RESULTS

The implementation of the research design was challenging as we had to customise many of the tools used to fit the context. Similarly we had to access the Emotive code layer in order to audit the findings. Initially the standard Bluetooth sniffing tools functioned as expected and easily compromised the confidentiality of the communication between the headset and the device. However, the violation of communication integrity required the implementation of the Btlejuice system of hardware and software in order to create replay attacks that changed the intended device effects. The headset communication was connected to a proxy client which had the same characteristics and services as the Insight headset. On launching the application called Mental Commands, a dummy headset is displayed on the proxy, and can be manipulated to recreate any of the headset commands. The compromised commands were then sent back to the headset to broadcast to the device. As soon as the proxy application connects to the headset, the Btlejuice suite starts to capture all the data sent. The following set of screen shots (figures 4 to 9) show the results.

Figure 4. Screenshot of Btlejuice intercepting communication

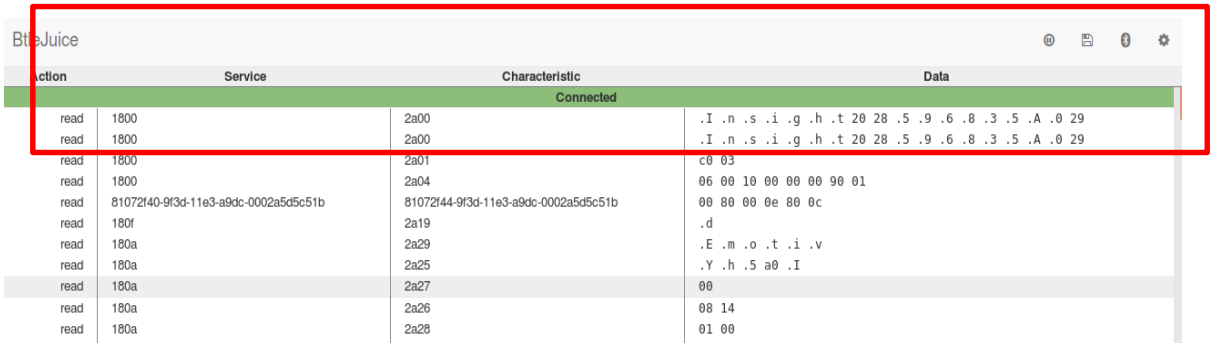


Figure 5. Screenshot of the application Mental Commands sending commands to the headset

All the data transmitted by the application on the proxy destined for the Insight headset could be intercepted and also had the option to modify the data. This feature known as 'on the fly modification' could be performed. Initially, any command sent by the headset will be sent to the proxy to confirm whether the data should be forwarded to the headset for sending to the device or not. Figure 6 presents a screenshot of the active data intercepted with an option to forward the data or simply devoid the headset of that specific data. This intercepted data could also be modified with a different command to the headset for a different or unintended function to perform.

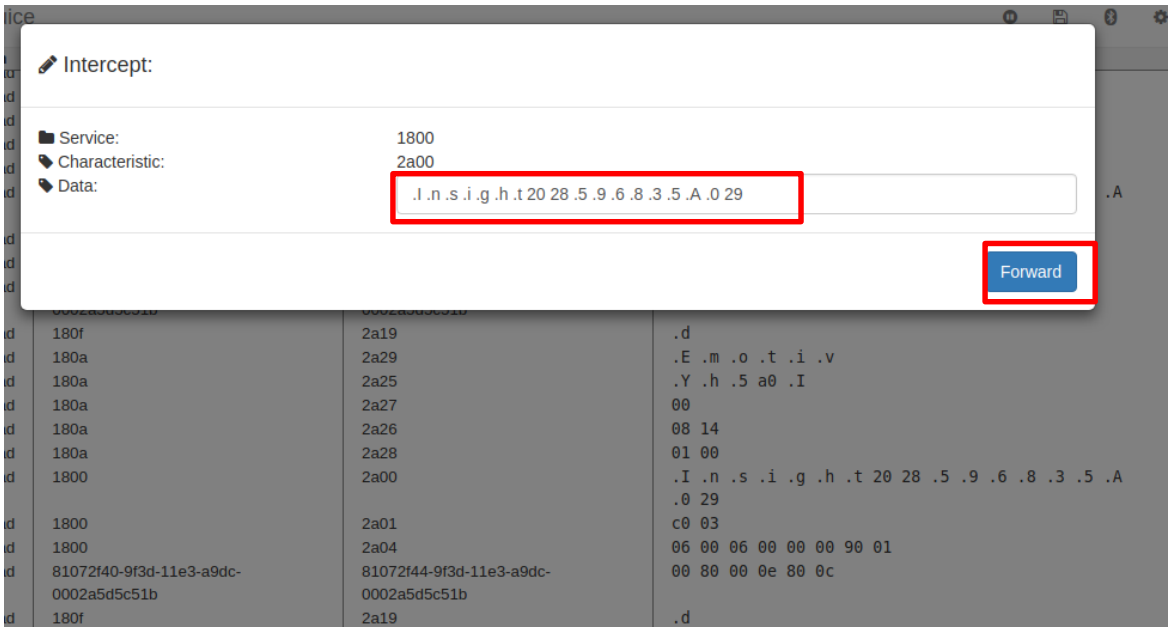


Figure 6. Screenshot of the active data sent from the proxy application to the headset

Figure 7 shows specific data captured for command 1800, and the way it may be modified to any other command. In this situation 1800 was associated with the device turning left effect. On this screen the turn right command 180f can be inserted to replace 1800, and the device effect subsequently changed. The lexicon of commands can be obtained from the headset or from the support literature.

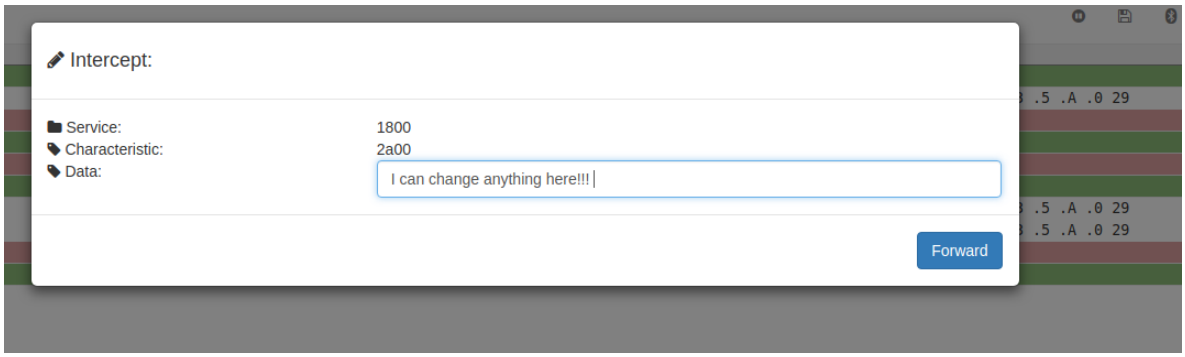


Figure 7. Screenshot of the modification of commands

The command change can also be seen in the Btlejuice terminal window in figure 8.

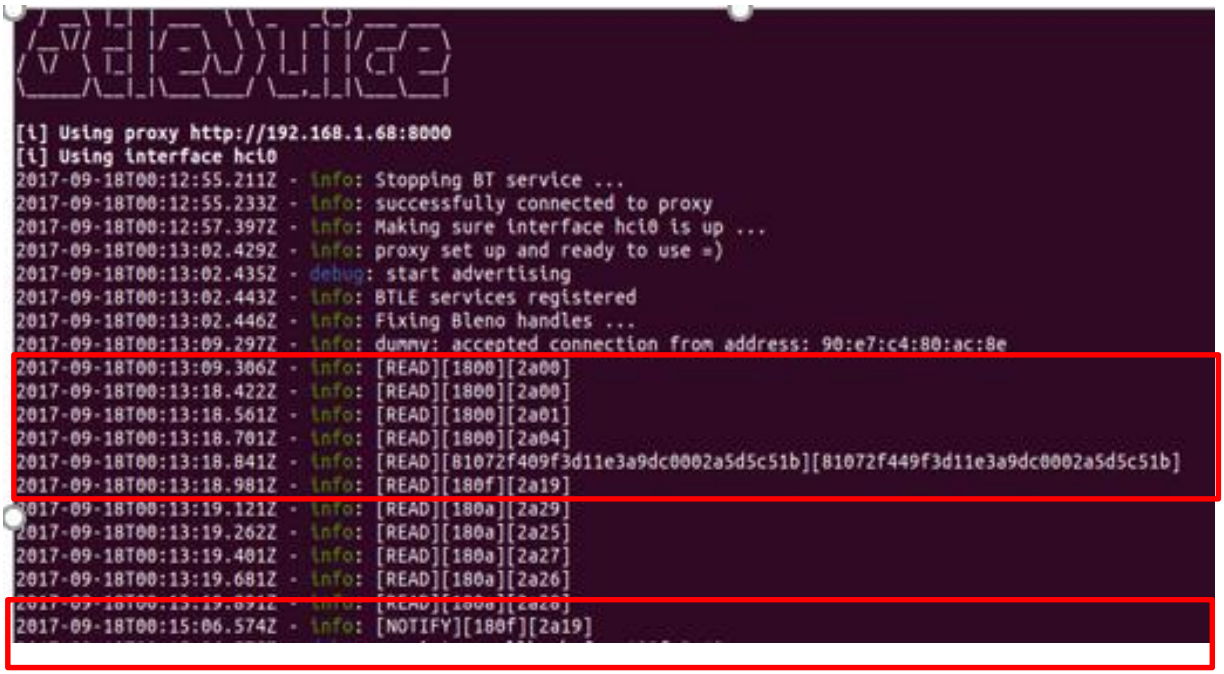


Figure 8. Screenshot of the modified data sent back to the headset

The images report the feasibility of a data modification attack and a replay attack; to violate the integrity of Brianware communication to a device. On further analysis, the nrf Bluetooth application could list both the devices, the real Insight headset with the mac address **f2:78:4a:15:77:bb** along with the fake Insight headset on the proxy with the mac address **00:1a:7d:da:71:14.**

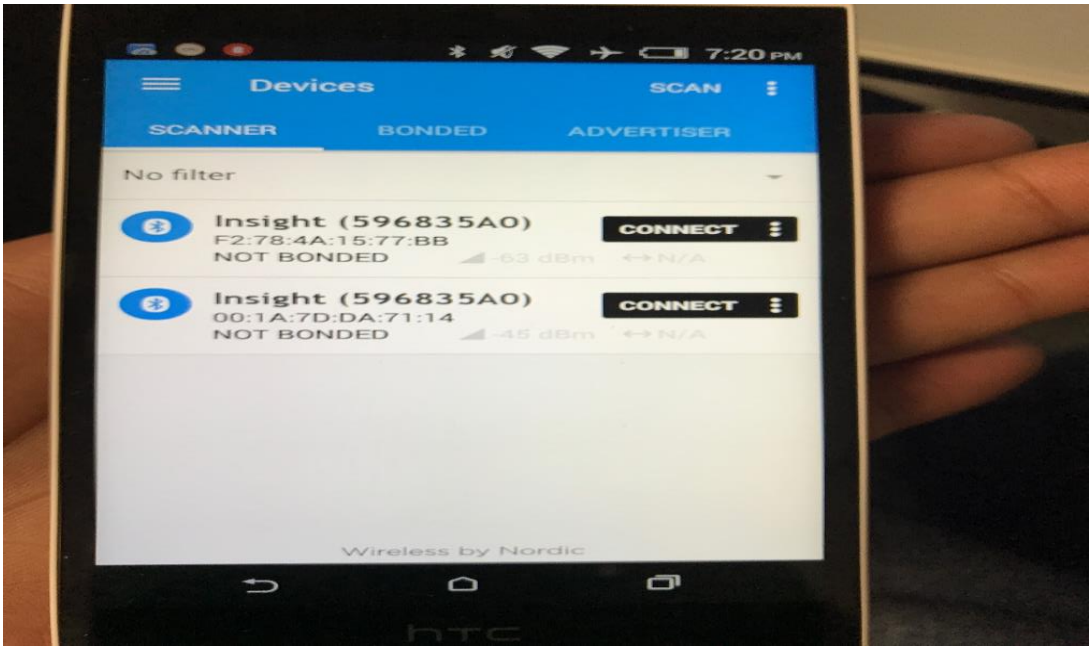


Figure 9. Screenshot of the nrf Bluetooth application showing the real and the fake control

The headset is connected to the proxy and the fake application on the smartphone. The fake device has been cloned with the same features and characteristics as the original device which fools the application to think that the fake device is the real device. Figure 9 presents a screenshot of two Insight headsets with different mac address that look identical, including the serial number of the headset.

CONCLUSION

The confidentiality of the headset was easily compromised but the violation of integrity was more difficult. We fundamentally structured a man in the middle attack where fake packets were substituted for the real ones. The attack progressed in two phases. In the first phase, the primary channel of communication sent the packets correctly but then the control was switched to a fake proxy. The proxy took the correct packets and substituted alternative ones back to the headset, and subsequently the communication stream to the device. This meant that the radio controlled car would get one signal that would tell it to turn right and then almost immediately another signal to turn left. The consequence was that the car would buzz but turn neither left nor right and remained frozen in the current state. The second phase of attack was to divert the primary communication channel and to substitute new control commands. This meant that if the primary channel had told the car to turn right then we removed the control packets and substituted a fake command to turn left. These attacks were successful and the remote control car became in control of the secondary information source. The effect demonstrated that it is possible to shift the primary control to a secondary source but there is still more research to design a sandbox that would quickly process the incoming raw signals and to substitute the fake commands. These findings are disturbing and indicate that the accessibility to the communication channel between the headset and the device or game, can also be disrupted. A simple denial of service attack can be hosted by multiple secondary sources substituting packets into the communication stream. These packets could be both meaningless and meaningful in the command and control structures, but either disposition would bring disruption to channel access. The implications are for disruption of human intentions and unintended actions.

The consequences of security failure in brainware devices are yet to be documented in sufficient numbers and scope, that regulatory requirements are implemented for device performance specifications. We also observed that with different brainware headsets that there were no standardised ways of doing a smile for example. This is something that the industry might look at in the future so that when a user is training a headset then a human characteristic is consistent between the different brands and different algorithms. The headsets are also sensitive to underlying emotions and can be used for feedback to the user and not just to an external control situation. For example the five electrode headset also reported to the user other parameters that included the user attention level, the user focus level, the user engagement level, the user interest level, the user relaxation level, and the user stress level. These emotional contexts are part of the feature extraction the brainware computes and provides as output. Our concern here is that not only is there information with an external control capability, but these headsets are also linked into an information feedback loop to the user. If either of these two information streams is compromised, then there are unplanned for consequences arising from the use of the technology.

REFERENCES

Allison, B., Graimann, B., & Gräser, A. (2007). Why use a BCI if you are healthy. Paper presented at the ACE Workshop-Brain-Computer Interfaces and Games.

Bonaci, T., Calo, R., & Chizeck, H. J. (2014). App stores for the brain: Privacy & security in Brain-Computer Interfaces. Paper presented at the International Symposium on Ethics in Science, Technology and Engineering, 2014 IEEE.

Da Silva, F. L. (1996). The generation of electric and magnetic signals of the brain by local networks. Comprehensive human physiology (pp. 509-531): Springer.

Darvas, F., Pantazis, D., Kucukaltun-Yildirim, E., & Leahy, R. (2004). Mapping human brain function with MEG and EEG: methods and validation. NeuroImage, 23, S289-S299.

Denning, T., Matsuoka, Y., & Kohno, T. (2009). Neurosecurity: security and privacy for neural devices. Neurosurgical Focus, 27(1), E7.

Donoghue, J. P. (2002). Connecting cortex to machines: recent advances in brain interfaces. Nature neuroscience, 5, 1085-1088.

Golub, M. D., Chase, S. M., Batista, A. P., & Byron, M. Y. (2016). Brain–computer interfaces for dissecting cognitive processes underlying sensorimotor control. Current opinion in neurobiology, 37, 53-58.

Jeunet, C., Jahanpour, E., & Lotte, F. (2016). Why standard brain-computer interface (BCI) training protocols should be changed: an experimental study. Journal of neural engineering, 13(3), 036024.

Kroeker, K. L. (2011). Improving Brain-computer interfaces. Communications of the ACM, 54(10), 11-14.

Lauer, R. T., Peckham, P. H., Kilgore, K. L., & Heetderks, W. J. (2000). Applications of cortical signals to neuroprosthetic control: a critical review. IEEE Transactions on Rehabilitation Engineering, 8(2), 205-208.

Li, Q., Ding, D., & Conti, M. (2015). Brain-computer interface applications: Security and privacy challenges. Paper presented at the IEEE Conference on Communications and Network Security (CNS).

Martinovic, I., Davies, D., Frank, M., Perito, D., Ros, T., & Song, D. (2012). On the Feasibility of Side-Channel Attacks with Brain-Computer Interfaces. Paper presented at the USENIX security symposium.

Millan, J. R., Renkens, F., Mourino, J., & Gerstner, W. (2004). Noninvasive brain-actuated control of a mobile robot by human EEG. IEEE Transactions on biomedical engineering, 51(6), 1026-1033.

Nijholt, A., Bos, D. P.-O., & Reuderink, B. (2009). Turning shortcomings into challenges: Brain–computer interfaces for games. Entertainment computing, 1(2), 85-94.

Ramadan, R. A., & Vasilakos, A. V. (2017). Brain computer interface: control signals review. Neurocomputing, 223, 26-44.

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain–computer interfaces for communication and control. Clinical neurophysiology, 113(6), 767-791.

Wyckoff, S. N., Sherlin, L. H., Ford, N. L., & Dalke, D. (2015). Validation of a wireless dry electrode system for electroencephalography. Journal of neuroengineering and rehabilitation, 12(1), 95.

INTELLIGENT FEATURE SELECTION FOR DETECTING HTTP/2 DENIAL OF SERVICE ATTACKS

Erwin Adi[1], Zubair Baig[2]
[1]Australian Centre for Cyber Security, School of Engineering and Information Technology, University of New South Wales, Australia
[2]Security Research Institute, School of Science, Edith Cowan University, Perth, Western Australia
e.adi@adfa.edu.au, z.baig@ecu.edu.au

Abstract

Intrusion-detection systems employ machine learning techniques to classify traffic into attack and legitimate. Network flooding attacks can leverage the new web communications protocol (HTTP/2) to bypass intrusion-detection systems. This creates an urgent demand to understand HTTP/2 characteristics and to devise customised cyber-attack detection schemes. This paper proposes Step Sister; a technique to generate an optimum network traffic feature set for network intrusion detection. The proposed technique demonstrates that a consistent set of features are selected for a given HTTP/2 dataset. This allows intrusion-detection systems to classify previously unseen network traffic samples with fewer false alarm than when techniques used in literature were employed. The results show that the proposed technique yields a set of features that, when used for network traffic classification, yields low numbers of false alarms.

Keywords: HTTP/2, feature selection Denial of Service, machine learning

INTRODUCTION

Hypertext Transfer Protocol (HTTP) has been the standard for web browser communication since the end of the 20th century. Until recently, most web communications were reliant on HTTP version 1.1, which was designed to transfer texts. As current web sites render large sized content such as audio and video, users routinely experience slow web browsing experience through HTTP/1.1. Hence, the new version, HTTP/2 (Belshe, Peon, & Thomson, May 2015), was designed to deliver web services at higher transfer rates, enhancing the end-user experience.

HTTP/2 has been accepted as the next generation standard for web communications. The protocol had its preliminary version deployed in 2010, was formally published in 2015, and is currently deployed by approximately 10 million websites globally, or 19% of all the websites (Usage of HTTP/2 for websites, November 2017). Major web browsers such as Mozilla, Chrome and Microsoft Edge support the protocol; and popular web sites such as Google services, Facebook and Twitter operate their web services employing the protocol.

As with all technological advances, the threat posed by the adversary class is ever existing. Adversaries can send a large volume of HTTP traffic towards a target HTTP/2 web service, causing resource exhaustion and eventual prevention of access for legitimate users. Such exploit is commonly known as a flood-based attack, or a Denial of Service attack. To detect flood-based attacks, previous studies have applied machine learning techniques. These techniques can learn from data samples, adapt to new environments, produce rule sets, and predict the class of unseen data. In detecting flood-based attacks, these known techniques construct a model of legitimate (normal) traffic, and classify attack traffic as instances where feature values extracted from the network traffic deviate from the baseline legitimate model.

A selection of studies that have previously classified traffic into flood-based and normal (Moore & Zuev, 2005; Mukherjee & Sharma, 2012; Baig, Sait, & Shaheen, 2013; Katkar & Kulkarni, 2013; Al-Jarrah et al., 2014) employed feature ranking and selection techniques to reduce the complexity associated with large scale data classification. Techniques such as Information Gain (Kullback & Leibler, 1951) and Gain Ratio are commonly used for feature *ranking*. The higher the rank of a feature, the more relevant the features are for the traffic classification process. Subsequently, the order of the rank number can be applied to *select* a set of most relevant features.

The main issue introduced with feature ranking and selection based on the above techniques is that they do not always yield the same subset of relevant features for varying datasets (Witten & Frank, 2005, p. 154). Specifically, different members of *{1 . . . n}* features are yielded, given different datasets. Over time, additional dataset samples are required to extend previous knowledge, to construct models that represent the current situation, and to classify with an acceptable degree of accuracy. Hence, there is a need to select a chosen set *{1 . . . n}* of features that will

yield a consistent list of members in the set, regardless of the data set choice. This study proposes a feature selection technique to address the aforementioned need. Specifically, this study proposes selection of a chosen set of features, given two different datasets that describe the characteristics of HTTP/2 traffic, to perform better than when HTTP/1.1 features are processed and ranked. The significance of this study is twofold: first, it proposes a novel feature selection technique for HTTP/2 traffic; and second, it examines how the technique can be applied to analyse HTTP/2 traffic based on variable datasets.

BACKGROUND

In detecting HTTP/1.1 flood-based attacks, reported studies have found that attack traffic showed higher number of packets when compared to the number of TCP connections that were established for a given target web service (Jung, Krishnamurthy, & Rabinovich, 2002; Ni, Gu, Wang, & Li, 2013). The introduction of HTTP/2 changed this concept significantly. The novel HTTP/2 mechanisms for communicating parties can be leveraged by adversaries to create previously unseen attack vectors to disrupt these services.

A recent study (Adi, Baig, & Hingston, 2017) showed that HTTP/2 flood-based attacks can be modelled to mimic the number of TCP connections observed on legitimate network traffic. The flood traffic, when examined through methods known in the literature to detect such attacks (Kumar, Joshi, & Singh, 2007; Lakhina, Crovella, & Diot, 2005; Rahmani, Sahli, & Kamoun, 2012), yielded a high number of False Alarms, indicating that the attack traffic mimicked normal traffic, thereby bypassing intrusion-detection systems. This showed that an urgent challenge exists to understand the characteristics of HTTP/2 flood-based attack traffic.

The above study (Adi et al., 2017), henceforth named Stealthy Attack Model, also proposed a set of 42 features as input to machine learning techniques to differentiate attack from normal network traffic. The study showed that the proposed set of features yielded fewer False Alarms compared to when HTTP/1.1 features were ranked and used for classification. These features were obtained from the statistical properties of captured HTTP/2 packets, which were grouped based on 3 features, namely, count, size, and lapse. The count features were obtained from the number of packets captured in an instance of time. The size feature was the total number of bytes of packets captured in an instance. The lapse feature was the time difference between the connection initiation of a packet (indicated by a SYN packet) and the time when the packet was captured. The study ranked the features with both Information Gain and Gain Ratio techniques, and analysed machine learning classification performance when applied on varying sets of ranked features. It showed that machine learning techniques such as Naïve Bayes, Decision Trees, JRip, and Support Vector Machines can be employed to classify HTTP/2 traffic into attack and normal class with good accuracies.

False Alarms is a notable performance measure employed by studies in the field of intrusion detection, to classify traffic into flood-based and normal (Wang, Zhang, Hei, Ji, & Ma, 2016; Manzoor, Kumar, et al., 2017; Suhasaria, Garg, Agarwal, & Selvakumar, 2017; Suganya, 2016; Osanaiye, Choo, & Dlodlo, 2016; Latif, Abbas, Latif, & Masood, 2015). False Alarms, described by the equation shown below, is defined as the percentage of instances incorrectly classified out of the total number of the whole instances *S* in a dataset. The False Positive *FP* is the ratio of normal traffic that a machine learning technique incorrectly identifies as attack traffic. The False Negative *FN* is the ratio of attacks that the technique incorrectly identifies as legitimate traffic.

$$False\ Alarm = \frac{\mathrm{FP} + \mathrm{FN}}{\mathrm{S}} \times 100\%$$

The percentages of False Alarms are illustrated on the Y-axis of a graph, as a function of 1 to *{1 . . . n}* features, presented on the X-axis of the graph. A lower percentage of False Alarms implies a better performance of the classifier. Observations based on human intervention are required to analyse which set of features are best employed for classification tasks to yield the lowest percentage of False Alarms. This was demonstrated in the Stealthy Attack Model analysis: different features sets were chosen to obtain the desired performance when different machine learning techniques were employed (Adi et al., 2017, Sec. 5.3).

The aforementioned study demonstrated that the results were observed and analysed to compare the performance when different sets of features were selected. On the other hand, the study reported in this paper places the observation before the classification results are obtained. The study also highlights that relationships exist between various features of a data set, and leverages this relationship to select relevant features.

DATA PROCESSING AND FEATURE SELECTION

Machine learning performance can be optimised through analysing relationships that exist in the data (Witten & Frank, 2005, p. 78). Learning involves examining relationships between objects rather than between instances. Of particular interest in this study is the concept of a sister in any family. Traditionally, when translated to a dataset representation, a sister is technically described as when two people can be labelled as *sister = yes* or *sister = no*. However, this representation can be prohibitive in terms of storage cost: describing the sister relationship of 100 people would require 100^2 = 10,000 instances. Describing such a relationship in real datasets poses a further problem: the number of objects that form relationships in datasets are unknown (Witten & Frank, 2005, p. 48). The solution to this problem is to inspect the objects and their relationships in the dataset before selecting features, to identify whether the two objects have the same parents (Witten & Frank, 2005, p. 48). This is illustrated in Table 1.

Table 1. A representation of a sister relationship

Name	Parent 1	Parent 2	Name	Parent 1	Parent 2	Sister
P	A	B	Q	A	B	Yes

Table 1 also lists P and Q as two objects, to examine if they are sisters. For the purpose of simplicity, the illustration ignores the gender of the objects. The table describes that P and Q are sisters because they have the same parents. This observation serves as the framework of the proposed feature selection technique, which is presented in the next section.

PROPOSED FEATURE SELECTION TECHNIQUE

This study observes that a sister relationship exists between count and size feature groups of the two datasets that were acquired from the Stealthy Attack Model (namely Stealthy Attack-1 and Stealthy Attack-2 that are discussed in the next subsection).

Referring to Table 1 as the framework, this study proposes that *P = count* and *Q = size*. These two parameters are sisters since they have the same parents. Parents were defined based on features of network traffic. One parent (A) was defined from packets generated by bots, that are uniform in their number of packets and packet sizes, compared to those generated by heterogeneous devices. In the Stealthy Attack Model, the bots were implemented by at most two C-libraries, i.e. curl (Stenberg, 1996–2016) and nghttp2 (Tsujikawa, 2015). This contrasts with real network traffic which is generated by different devices from different manufacturers, having different hardware/software implementations. Real network traffic generates heterogeneous packet sizes. Instead, the bots were generated through a model that dictates how client-to-server packets are formed. Particularly, Stealthy Attack Model launched only one HTTP/2 packet type (namely window update packet), where the packet size was defined by the C-libraries that implemented the bots. Furthermore, there were a limited number of bots that generated such network packets, causing the model to generate a less dispersed number of packets compared to normal traffic. This is illustrated in Figure 1, which is adapted from the Stealthy Attack Model study (Adi et al., 2017, Sec. 5.2).

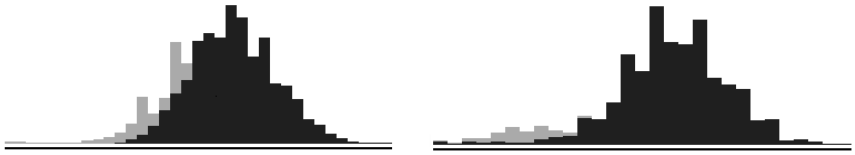


Figure 1. Distribution of attack (grey) and normal (black) feature values

Figure 1 illustrates the distribution of two different feature values, in this example, the count_syn (the number of SYN packets in an instance) and the size_rstAck (the total size of RST ACK packets in an instance), for the stealthy attack dataset. The X-axis represents various ranges of feature values, while the Y-axis describes the number of instances in the dataset that encompass such range of values for a given feature. The grey bars represent attack traffic generated by bots, while the black bars represent normal traffic. Figure 1 shows that the normal traffic is more distributed than the attack traffic: the black bars span wider than the grey bars; while the black bars forms a distributed bell-curve, the grey bars are lower relative to the black bars, and plane relative to the X-axis. This illustrates that the attack traffic is less heterogeneous in terms of packet count and sizes, when compared to normal.

The other parent (B) was defined as attacking bots attempting to mimic normal flows of normal data traffic. In the Stealthy Attack Model, two parameters were defined to mimic normal traffic: a stealthy factor and a delay between connections. The stealthy factor *sf* was defined as a flood of HTTP/2 packets launched with a probability of *1/sf*.

A delay between connections of *d* was prescribed as *d* ms added between successive TCP connections of a given session. These parameters defined the bots capable of controlling the number of packets launched from a malicious client towards a target server. As the bot traffic was controlled (i.e., operated below the radar), the number of packets and their sizes are less heterogeneous than those of normal traffic (as Figure 1 shows).

Henceforth, *P* and *Q* are sisters, since they both were derived from non-heterogeneous bots that attempted to mimic normal traffic.

Step Sister Algorithm

This study proposes the Step Sister technique to select a set of features based on inter-sister relationships, given two different datasets. The precondition of the algorithm is to have two sets of ranked features. In this study, the datasets Stealthy Attack-1 (S1) and Stealthy Attack-2 (S2) were ranked through application of the Information Gain technique (Kullback & Leibler, 1951).

The output of the algorithm is to have a *ChosenSet* of features; this set is initially assigned to an empty set. While there is no definition of how many features is sufficient, a finite number of 5 features was considered in this study for demonstration.

Two sets of ranked features were employed from the Stealthy Attack datasets (Adi et al., 2017). These are the Stealthy Attack-1 dataset and the Stealthy Attack-2 dataset. These features were ranked through application of the Information Gain technique, resulting in two ranked lists, S1 and S2, respectively. For the purpose of this study, the two features chosen to analyse HTTP/1.1 traffic, i.e. the count_app and the count_syn features (explained in the next subsection), were not selected as part of the *ChosenSet*. Furthermore, because this study observed that the count group has a sister association with the size group, this study disregarded the size_app and the size_syn features. These are shown on lines 5 - 7 of the algorithm.

To define a sister relationship, the algorithm examines the feature groups and feature types. Feature groups are features that share the same characteristics such as “count” and “size”. Feature types signify the packet types where the features were extracted from. For example, rstAck is a feature type, which is extracted from a network packet carrying RST ACK flag. Hence, a feature named count_rstAck is obtained from the number of packets (i.e. count) carrying RST ACK flags, observed in an instance of time. The sister association is found to be true when a feature *f* from one of the ranked list *S* being examined belongs to the “count” group, and the same type of feature where its group is “size” is found in the *ChosenSet*. The converse situation where *f* belongs to the “size” group, also demonstrate a sister association when the same type of feature of group “count” is found in the *ChosenSet*.

Henceforth, the Step Sister algorithm can be simplified as follows. The first step is to clean the two ranked list *S1* and *S2*, by eliminating a feature is its sister is already in the *ChosenSet*, starting from the highly ranked feature from each list. It selects a feature from the two lists *S1* and *S2* that is ranked higher. These steps iterate until there is a handful number of features in the *ChosenSet* with size *n*.

Stealthy Attack Datasets

The two attack models proposed in the Stealthy Attack Model were named Stealthy Attack-1 and Stealthy Attack-2. The Stealthy Attack datasets described attack and normal HTTP/2 traffic classes: the attack data was generated out of two attack models; and the normal data was obtained from simulating 5,200 bots that mimicked human behaviour when online.

The Stealthy Attack-2 dataset extended the Stealthy Attack-1 dataset through employing attacking bots that mimicked the distribution value of a highly relevant feature observed from the Stealthy Attack-1 data analysis. In both datasets, the features were ranked through employing both Information Gain and Gain Ratio algorithm. The result showed that 42 features could detect HTTP/2 flood-based attacks better, i.e. fewer False Alarms than when two HTTP/1.1 features were employed. The two HTTP/1.1 features used as a comparison were count_app, i.e. the number of Application Data packet observed in an instance, and count_syn, the number of TCP connection initiation observed in an instance.

RESULTS AND ANALYSIS

The *ChosenSet* from the Step Sister algorithm yielded the following set of features:

- size_rstAck, is the total size (in KB) of packets carrying RST-and-ACK TCP flags observed in one instance.

- size_tlsKey, is the total size (in KB) of TLS packets carrying key exchange observed in one instance
- count_encryptedAlert, is the total number of TLS packets carrying Encrypted Alert flags observed in one instance.
- lapse_rstAck max, is the maximum duration of time between packets carrying RST-and-ACK TCP flags within an observed instance, and a packet carrying SYN flag signifying its first connection initiation.
- size_tlsHello, is the total size (in KB) of TLS packets carrying Hello packet type observed in one instance.

The above set of features was employed to classify traffic described for both datasets, i.e. the Stealthy Attack-1 and Stealthy Attack-2 datasets. The classification analysis employed Weka (University of Waikato, 1993–2016), software that provides a collection of machine learning techniques, to analyse the classification performance in terms of False Alarm. Four machine learning techniques were employed: Naïve Bayes (NB), Decision Tree J48 (DT), JRip, and Support Vector Machines (SVM). Two other feature ranking algorithms that were employed in the Stealthy Attack Model, Information Gain (IG) and Gain Ratio (GR), serve as a comparison to analyse the classifier performance.

The Step Sister algorithm did not aim to yield better performance than what the Stealthy Attack Model study yielded. This is shown in Table 2. The False Alarm yielded by the machine learning techniques was compared when the features were selected and ranked by the Step Sister (SS) algorithm, Information Gain (IG) and Gain Ratio (GR). The values obtained for the SS columns were the outcomes of this study, while the values shown for the IG and GR columns were the outcomes of the Stealthy Attack Model study. To compare the correct values, only the most relevant 5 features were selected from the list of ranked features by IG and GR. The table shows that the numbers in column SS are not consistently lower than the values on the other columns. Hence, the SS algorithm did not seek to optimise the performance obtained in the Stealthy Attack Model.

Table 2. False Alarms produced (%) when different algorithms were employed

	Stealthy Attack-1			Stealthy Attack-2		
	IG	GR	SS	IG	GR	SS
NB	0.2892	0.2410	0.2410	0.0519	0.0519	0.0778
DT	0.0482	0.0723	0.0482	0.0519	0.0519	0.0519
JRip	0.0482	0	0.0723	0.0259	0.0259	0.0259
SVM	0	0	0	0	0.0778	0

Table 3. False Alarms produced (%) when different feature sets were employed

	Stealthy Attack-1		Stealthy Attack-2	
	HTTP/1.1	SS	HTTP/1.1	SS
NB	0.2651	0.2410	0.5189	0.0778
DT	0.1687	0.0482	0.2335	0.0519
JRip	0.1205	0.0723	0.2335	0.0259
SVM	0	0	0.3373	0

Consistent results can be seen when the machine learning performance was compared to those when the HTTP/1.1 features were employed. This is illustrated in Table 3. All of the False Alarm yielded by the machine learning techniques were lower when the features were selected by the SS algorithm, compared to those when HTTP/1.1 features were employed. Hence, the *ChosenSet* of five features consistently yielded better performance.

The *ChosenSet* features are more consistent than the set of features employed in the Stealthy Attack Model study: first, the Step Sister algorithm selected the same *ChosenSet* of features to be employed by machine learning techniques. This is different to the methods employed by the Stealthy Attack Model study, where different sets of features *{1 . . . n}* must be selected. Second, in the Stealthy Attack Model study, observations were required to choose the size n of the set *{1 . . . n}*. In contrast, the same five features selected by the Step Sister algorithm in this study were employed by the machine learning techniques. This demonstrated the aim of the study: to yield the same set of features to analyse different datasets.

CONCLUSION

This study proposed a technique, namely, *Step Sister*, which yielded a set of features to aid in machine learning-based classification of HTTP/2 flooding traffic. The analysis employed two datasets that described the characteristics of HTTP/2 flooding and legitimate traffic, respectively. Since the Step Sister technique yielded a consistent set of features for machine-learning based classification, intrusion-detection systems that employ this technique can operate without requiring human intervention to choose the size and members of the feature set. Furthermore, the proposed technique was tested on two varying datasets. The study demonstrated that the Step

Sister technique analysed both datasets to yield a chosen set of features. Machine learning techniques that employed these set of features yield lower False Alarms than when techniques known in literature were employed to analyse HTTP/2 traffic.

REFERENCES

Adi, E., Baig, Z., & Hingston, P. (2017). Stealthy denial of service (DoS) attack modelling and detection for HTTP/2 services. *Journal of Network and Computer Applications*, *91*, 1–13.

Al-Jarrah, O., Siddiqui, A., Elsalamouny, M., Yoo, P., Muhaidat, S., & Kim, K. (2014, June 30-July 3). Machine-learning-based feature selection techniques for large-scale network intrusion detection. *Distributed Computing Systems Workshops, Madrid, Spain* (p. 177-181). IEEE Xplore.

Baig, Z. A., Sait, S. M., & Shaheen, A. (2013). GMDH-based networks for intelligent intrusion detection [Journal Article]. *Engineering Applications of Artificial Intelligence*, *26* (7), 1731-1740.

Belshe, M., Peon, R., & Thomson, M. (May 2015). *Hypertext Transfer Protocol version 2 (HTTP/2)* (Report No. RFC 7540). Internet Engineering Task Force (IETF).

Jung, J., Krishnamurthy, B., & Rabinovich, M. (2002, May 7-11). Flash crowds and Denial of Service attacks: Characterization and implications for CDNs and web sites. In *Proceedings of the 11th International Conference on World Wide Web, Honolulu, Hawaii* (p. 293-304). New York: ACM.

Katkar, V. D., & Kulkarni, S. V. (2013, December 12-14). Experiments on detection of denial of service attacks using naive bayesian classifier. In *International Conference on Green Computing, Communication and Conservation of Energy, India* (p. 725-730). IEEE Xplore.

Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. *The Annals of Mathematical Statistics*, *22* (1), 79–86.

Kumar, K., Joshi, R., & Singh, K. (2007). A distributed approach using entropy to detect DDoS attacks in ISP domain. In *International Conference on Signal Processing, Communications and Networking, Honolulu, Hawaii* (pp. 331–337). IEEE Xplore.

Lakhina, A., Crovella, M., & Diot, C. (2005). Mining anomalies using traffic feature distributions. In *ACM SIGCOMM Computer Communication Review* (Vol. 35, pp. 217–228).

Latif, R., Abbas, H., Latif, S., & Masood, A. (2015). EVFDT: an enhanced very fast decision tree algorithm for detecting distributed denial of service attack in cloud-assisted wireless body area network. *Mobile Information Systems*, *2015*.

Manzoor, I., & Kumar, N. (2017). A feature reduced intrusion detection system using ANN classifier. *Expert Systems with Applications, 88,* 249-257.

Moore, A. W., & Zuev, D. (2005). Internet traffic classification using bayesian analysis techniques [Conference Proceedings]. In *ACM SIGMETRICS Performance Evaluation Review* (Vol. 33, p. 50-60). ACM.

Mukherjee, S., & Sharma, N. (2012). Intrusion detection using naive bayes classifier with feature reduction [Journal Article]. *Procedia Technology*, *4*, *119*-128.

Ni, T., Gu, X., Wang, H., & Li, Y. (2013). Real-time detection of application-layer DDoS attack using time series analysis [Journal Article]. *Journal of Control Science and Engineering, 2013,* 4.

Oikonomou, G., & Mirkovic, J. (2009, June 14-18). Modeling human behavior for defense against flash-crowd attacks. In *IEEE International Conference on Communications, Dresden, Germany* (pp. 1–6). IEEE Xplore.

Osanaiye, O., Choo, K. K. R., & Dlodlo, M. (2016). Distributed denial of service (DDoS) resilience in cloud: review and conceptual cloud DDoS mitigation framework. *Journal of Network and Computer Applications, 67,* 147-165.

Rahmani, H., Sahli, N., & Kamoun, F. (2012). Distributed Denial-of-Service attack detection scheme-based joint-entropy [Journal Article]. *Security and Communication Networks, 5* (9), 1049-1061.

Saleh, M. A., & Abdul Manaf, A. (2015). A novel protective framework for defeating HTTP-based denial of service and distributed denial of service attacks. *The Scientific World Journal.*

Stenberg, D. (1996–2016). *cURL* [Software]. Retrieved from https://curl.haxx.se/download.html

Suhasaria, P., Garg, A., Agarwal, A., & Selvakumar, K. (2017). Distributed Denial of Service Attacks: A Survey. *Imperial Journal of Interdisciplinary Research*, *3*(3).

Tsujikawa, T. (2015). *Nghttp2: HTTP/2 C library* [Computer Program]. Retrieved from https://nghttp2.org/

University of Waikato. (1993–2016). *Weka (version 3.8)* [Software]. Retrieved from http://www.cs.waikato.ac.nz/ml/weka/downloading.html

Usage of HTTP/2 for websites. (n.d.). Retrieved November 02, 2017, from https://w3techs.com/technologies/details/ce-http2/all/all

Wang, Y., Zhang, Y., Hei, X., Ji, W., & Ma, W. (2016). Game strategies for distributed denial of service defense in the cloud of things. Journal of Communications and Information Networks, 1 (4), 143{155. Witten, I. H., & Frank, E. (2005). *Data mining: Practical machine learning tools and techniques* [Book]. Morgan Kaufmann.

Zhou, W., Jia, W., Wen, S., Xiang, Y., & Zhou, W. (2014). Detection and defense of application-layer DDoS attacks in backbone web traffic [Journal Article]. *Future Generation Computer Systems*, *38*, 36-46.

A SRI LANKAN HACKING CASE STUDY

Ishan Senarathna, Matthew Warren
Deakin University Centre for Cyber Security Research, School of Information Technology, Faculty of Science, Engineering and Built Environment, Deakin University, Victoria, Australia
ishan.senarathna@deakin.edu.au, matthew.warren@deakin.edu.au

Abstract

The aim of the paper is to consider how hacking could impact a country that had historically experienced major cyber-attacks. The aim of the paper is to explore a cyber incident that occurred against the Sri Lankan president and how Sri Lankan authorities reacted to the incident. The paper will focus upon the motivations of the attack, the impact of the attack and how Sri Lankan authorities reacted to the situation.

Keywords: Hacking, Government and Sri Lanka.

INTRODUCTION

We have seen a rise in computer misuse at a global level; in many cases "Hackers" have been found responsible for these attacks. Hackers are often characterised as adolescent males in dark bedrooms that can cause damage to global IT systems through using their computers and computer skills. A more romantic perception portrays hackers as being determined cyber knights, who use personal codes of conduct to live by and are reminiscent of the great Arthurian knights (Warren and Hutchinson, 2003). Moreover, "hacker" is what computer-intruders choose to call themselves, not as a criminal pejorative, but as a noble title given to those "soaked through with heroic anti-bureaucratic sentiment" (Sterling, 1993). Hacking then, can describe the determination to make access to computers and information as free as possible. Hacking can involve the heartfelt conviction that beauty can be found in computers, that the fine aesthetic in a perfect program can liberate the mind and the spirit (Levy, 1984).

Contrasting this romantic perception is the way Bruce Sterling (1993), portrays "Hacking" in his book titled 'The Hacker Crackdown'. In Sterling's (1993) book, "Hacking" is described as the act of intruding into computer systems by stealth and without permission. However, Sterling's definition of "Hacking" is broader than the one used routinely by most enforcement officials with any professional interest in computer fraud and computer abuse. The enforcement officials' focus on "Hacking" relates to crimes committed with, by, through, or against a computer (Warren and Hutchinson, 2003).

But when happens when a country that has never experienced major cyber incidents becomes victim to a high profile cyber incident. In terms of the paper it reflects upon Sri Lanka. Sri Lanka is an island located of the coast of India and has a population of 22 million people, of the Sri Lankan population 7.1 million (32%) are Internet users (CIA, ND).

The paper intends to answer one key research question:

> What were the motives and impacts in relation to the Sri Lankan President's Cyber incident.

OVERVIEW OF SRI LANKA

The Sri Lankan authorities in anticipation of increased cyber security incidents that Sri Lanka could face and the growing ICT infrastructure across Sri Lanka, the Sri Lanka Computer Emergency Readiness Team | Coordination Centre (CERT|CC) was established as Sri Lanka's National CERT, by the ICT Agency of Sri Lanka to protect against Sri Lanka's future Cyber incidents, (Sri Lanka Cert, ND).

The role of the Sri Lanka CERT|CC has developed over time, it has now become the national centre for cyber security in Sri Lanka, mandated to protect the nation's information infrastructure and to coordinate protective measures against, and respond to cyber security threats and vulnerabilities (Sri Lanka Cert, ND).

Sri Lanka has had a history of cyber incidents, the following table depicts the distribution of various types of incidents reported to Sri Lanka CERT during 2016 (APCert, 2016). All the incidents reported to Sri Lanka CERT had been resolved satisfactorily.

Table 1: Sri Lankan Cyber Security Profile Incidents (2016) (AP, 2016)

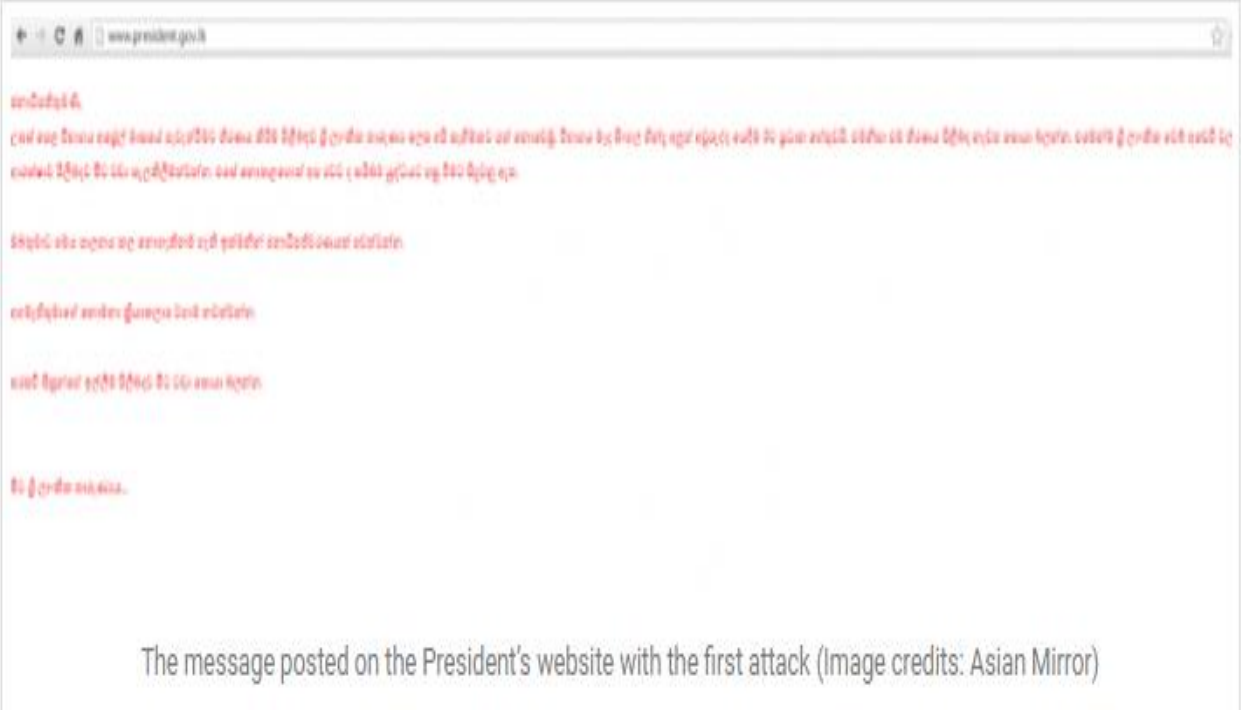
Type of Incident	Year 2016
Phishing	23
Abuse/Hate/Privacy violation (via mail)	32
Ransomware	10
Scams	12
Financial Frauds	16
Malicious Software issues	11
Web site Compromise	10
Compromised/hate/threat Email	16
Intellectual property violation	7
DoS/DDoS	4
Social Media related incidents	2200
Total	**2341**

The majority of issues that have been reported to the Sri Lankan CERT|CC related to social issues relating to social media and related issues such as cyber bullying, the "social media" type of cyber incidents represented 94% of the cases that Sri Lankan CERT|CC had to deal with and would relate to issues such as cyber bullying.

CASE STUDY

On the 25th August the official website of the Sri Lankan President Maithripala Sirisena, (www.president.gov.lk) suffered two cyberattacks on two consecutive days by a group who identified themselves as 'The Sri Lankan Youth' (BBC 2016, DNA 2016, Doole and Thomas 2016, Read Me News 2016, Yahoo, 2016).

The first attack took place on the 25th August 2016, the existing site was removed and replaced with a message. The message that was posted on the home page is shown in Figure1. The president site was hacked on Friday, August 26 with a message being posted in in Sinhala (Wollerton 2016). In the first message, the hacking group made a number of demands to the Sri Lankan government. One of these demands being a request to reconsider the decision to hold the GCE A/Level examination in April rather than in August. In addition, it also commented to the Government to be more conscientious regarding the security of Sri Lankan websites. If no action were to be taken with that regard, the country will have to be face a "cyber war" (Molloy 2016, Read Me News, 2016).



The message posted on the President's website with the first attack (Image credits: Asian Mirror)

Figure 1: The First Hacking Incident

The English translation of the first message from the hacking incident posted in Sinhalese was:

"Dear Mr. President,

We are extremely displeased about the decision to hold GCE A/L in April since the Sinhala/Hindu New Year falls in between the exam dates. Therefore, reconsider that decision. Furthermore, take care of the security of Sri Lankan websites. Or else, we will have to face a cyber war.

If you cannot control the situation hold a Presidential Election.

Stop the Prime Minister's irresponsible work.
Look more into the problems of the university students.

The Sri Lankan Youth"

The first message posted in Sinhalese was a political message which included the stopping of 'irresponsible conduct' by the Prime Minister and hold a presidential election if the president cannot control the situation. The message also directed the Government to pay more attention to the problems faced by university students.

After the first attack, the site was taken offline for a few hours and then the Presidents site returned to normal operations. But on the 26th August the Presidents site as hacked again, this time a message in English was posted (see Figure 2). The message made no demand, just stating that the site was down for maintenance and the site was quickly restored.

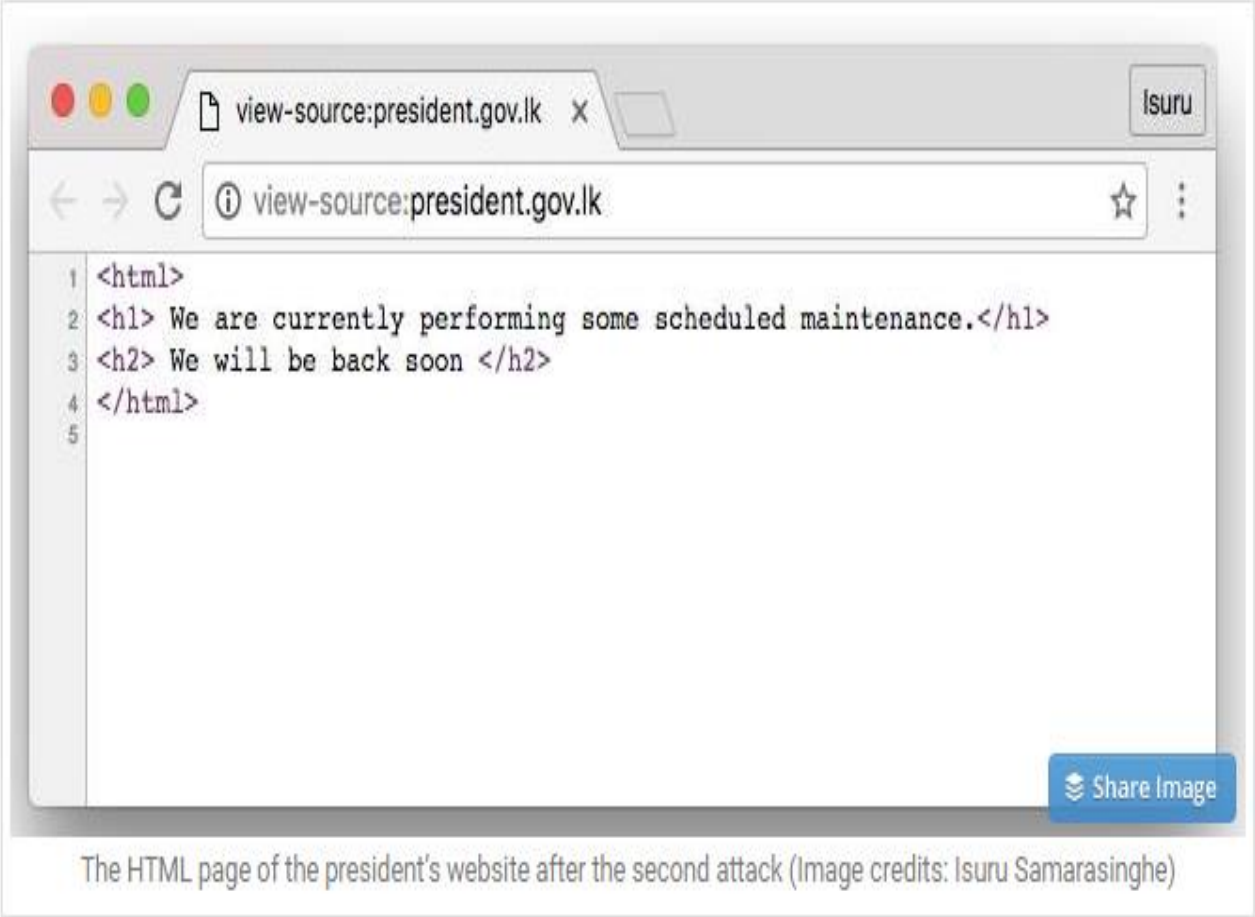


Figure 2: The Second Hacking Incident

WHO WAS RESPONISBLE

A group or individual known as the "The Sri Lankan Youth" was suspected to be responsible for the initial attack. Following the incident, a 17-year school boy from Kandy was arrested by the CID (Criminal Investigation Department) for allegedly hacking the website. Further, the CID arrested another 27 year old and charged him with hacking the President's website (Gossip News 2016). Under the Sri Lankan Computer Crimes Act, the unnamed boy who was arrested, and could face a possible fine of Rs. 300,000 (Sri Lankan rupee) and subsequent could face up to three years in jail (Doole and Thomas 2016, Read Me News, 2016).

The Sri Lanka Computer Crimes Act was enacted by the Sri Lankan Parliament and certified by the Speaker as the Computer Crimes Law No. 24 of 2007 with the aim of the protecting Sri Lanka against Cyber Crimes and these was the first arrests in connection to this law (Guardian, 2016).

According to CID, the arrested school boy had illegally accessed more than 37 websites and a Facebook account named 'arrow.lk' had been identified in connection with altering data and entering data (Doole and Thomas 2016). Further, the CID has informed that a group named 'Yakadaya Forums' on Facebook was involved in collecting and sharing information about security weaknesses links to certain hosting websites.(Doole and Thomas 2016).

DISCUSSION

The Sri Lankan President's web-site is a simple news website created for the purposes of disseminating information and news relating to President Sirisena and his activities and a way of connecting to the Sri Lankan public.

The President website was hosted on WordPress, a free and open-source content management system which is a popular choice for people to host their blogs and personal websites (Doole and Thomas 2016, Metzger 2016). The security issue here related to the fact that the WordPress site was not correctly configure and some blame poor security standards on the website as the cause of the problem (Metzger 2016).

The President's website hack was linked to a manipulation of the original program code (script) of the website. The hackers themselves had limited experience who use existing computer scripts or code to hack into computers. They lack the expertise to write their own program code to hack others websites and obtained the information they needed for the hack through the Facebook forums they were linked to (Doole and Thomas 2016). It was also determined that the second hacking incident was also linked to the first hacking incident (Daily News, 2016).

In terms of the paper's research question:

> What were the motives and impacts in relation to the Sri Lankan President's Cyber Incident.

In terms of the motives of the hacking incident the hacker's motivation was to highlight his displeasure that examinations had been scheduled for April, during the traditional Sinhala and Tamil New Year holidays. The motivations reflect that of a young person who was frustrated by the timing of a School exam and decided to hack the Presidents web-site to vent his frustration.

From a forensic perspective, not much information has been shared by the Sri Lankan authorities regarding the situation but following was determined:

1) After the first hacking incident, the Sri Lankan law enforcement agencies took over the operations of the Presidents web-site and very quickly reacted to the second hacking incident;
2) Very quickly the two individuals involved in the incident were identified and arrested by Sri Lankan authorities;
3) The security weakness on the President's WordPress site was quickly identified and corrected;
4) The Sri Lankan authorities used social media to collected information about the attackers and determine which Facebook groups where code exploits and other information had been exchanged.

The damage caused by the incident was limited and only impacted the credibility of the Sri Lankan authorities to protect the Presidents web-site.

The case just highlights a number of issues:

1) The problem of hosting government web-sites on third party web-sites where limited security systems may be in place;
2) The ability of unskilled attackers to use the Internet to download scripts to exploit security weaknesses, in this case being given the script via Facebook forums;
3) The role of media in portraying the defacement of the Sri Lanka Presidents web-site as a major "Cyber incident" and escalating the situation;
4) The capabilities of Sri Lanka authorities to quickly analyse the attacks and respond to the incident and arrest those connected.

CONCLUSION

The case is linked to a Sri Lankan teenager who hacked the president's website to try to reconsider the decision to hold the GCE A/Level examination in April rather than in August. In terms of this incident, it is a classic hacking incident based upon simple motivations of a single individual. The case does highlight the problem that governments have when using third party sites and services on behalf a national government official.

The outcome was that the case against the people arrested was dropped and the teenager behind the incident had to meet the President in person. During the meeting with the Sri Lankan President, he stated "It's our duty to encourage our youngsters to use their talents ethically" (Daily News, 2016).

REFERENCES

AP (Asian Pacific) Cert (2016) Annual Report, URL: https://www.apcert.org/documents/pdf/APCERT_Annual_Report_2016.pdfsite accessed: 7/10/17.
BBC (2016). "Sri Lankan teenager held over hacking of president's website." from http://www.bbc.com/news/world-asia-37214629. site accessed: 7/10/17.
CIA (nd) Sri Lanka, URL: https://www.cia.gov/library/publications/the-world-factbook/geos/ce.html, site accessed: 7/10/17.
Daily News (2016). URL: http://dailynews.lk/2016/11/18/local/99473, site accessed: 7/10/17.

DNA (2016). "Sri Lankan President Maithripala Sirisena's website hacked twice within two days." URL: http://www.dnaindia.com/world/report-sri-lankan-president-maithripala-sirisena-s-website-hacked-twice-within-two-days-2249709, site accessed: 7/10/17.
Doole, C. and Thomas K.C. (2016). "President's website Hack: The Full Story." URL: https://web.archive.org/web/20170118162646/http://www.ceylontoday.lk/print20160701CT20161030.php?id=5042, site accessed: 7/10/17.
Levy, S (1984). Hackers: Heroes of the Computer Revolution, Anchor Press, USA.
Guardian (2016) "Sri Lankan teenager hacks president's website to try to get exams delayed". URL: https://www.theguardian.com/world/2016/aug/30/sri-lankan-teenager-hacks-presidents-website-to-try-and-get-exams-delayed, site accessed: 7/10/17.
Gossip News (2016). "How Harshana and Janith who hacked President Website were cornered." URL: http://www.english.gossiplankanews.com/2016/08/how-harshana-and-janith-who-hacked.html, site accessed: 7/10/17.
Metzger, M. (2016). "Teenager hacks Sri Lankan president's website to protest exams." URL: http://www.scmagazineuk.com/teenager-hacks-sri-lankan-presidents-website-to-protest-exams/article/520647/, site accessed: 7/10/17.
Molloy, M. (2016). "Teenager accused of hacking president's website to try and get exams postponed ". URL: http://www.telegraph.co.uk/news/2016/08/30/teenager-accused-of-hacking-presidents-website-to-try-and-get-ex/, site accessed: 7/10/17.
Sterling B (1993). The Hacker Crackdown: Law and Disorder on the Electronic Frontier, Mass Market Paperback, USA.
Sri Lanka Cert (nd) http://www.slcert.gov.lk/aboutUs.php, URL: http://www.slcert.gov.lk/aboutUs.php, site accessed: 7/10/17.
Read Me News (2016). "The President's Website Was Hacked: Here's What We Know ". from http://www.readme.lk/presidents-website-hacked/, site accessed: 7/10/17.
Yahoo (2016). "Sri Lanka police arrest teen over hacking president's website." from https://www.yahoo.com/news/sri-lanka-police-arrest-teen-over-hacking-presidents-192221299.html, site accessed: 7/10/17.
Warren, M.J and Hutchinson W. (2003). Australian Hackers Ethics, Australian Journal of Information Systems, Vol 10, No 2, pp. 151 – 156.
Wollerton, M. (2016). "Sri Lankan teen hacks president's website to delay exams." from https://www.cnet.com/au/news/sri-lankan-teen-hacks-presidents-website-to-delay-exams/, site accessed: 7/10/17.

SECURITY VULNERABILITIES AND CYBER THREAT ANALYSIS OF THE AMQP PROTOCOL FOR THE INTERNET OF THINGS

Ian Noel McAteer[1], Muhammad Imran Malik[1], Zubair Baig[1,2], Peter Hannay[1,2]
[1]School of Science, [2]Security Research Institute, Edith Cowan University, Perth, Australia
imcateer@our.ecu.edu.au, mimalik@our.ecu.edu.au, z.baig@ecu.edu.au, p.hannay@ecu.edu.au

Abstract
The Internet of Things (IoT) expands the global Internet-connected network to encompass device-to-device, device-to-server, and server-to-server connectivity for an ever-increasing variety of end-user devices. IoT remains a somewhat amorphous entity, with little in the way of coordinated development, and is undermined largely by a manufacturer-driven scramble to be first-to-market with the latest innovation. Communication between IoT devices/servers relies on underlying protocols, which must be efficient and effective to establish and maintain reliability and integrity of data transfer. However, the lack of coordination during IoT's expansion has resulted in a variety of communications protocols being developed. AMQP (Advanced Message Queuing Protocol) originated from the financial sector's requirement for an improved messaging system that was fast, reliable and independent of end-user platform configurations. AMQP is an open-source server-to-server communications protocol which allows the addition of user-specific extensions. The software coding of such end-user-developed modules can be insufficient regarding threat-mitigation and can make the end product vulnerable to cyber-attack. Through this paper, we present vulnerability and threat analysis for AMQP-based IoT systems.

Keywords: AMQP (Advanced Message Queuing Protocol), AMQP Vulnerabilities, Application Layer Protocols, Internet of Things (IoT), IoT Architecture

INTRODUCTION

For more than a decade, the desire to have an increasing range of devices controllable via wired or wireless communications has seen an exponential rise in the number of Internet-of-Things (IoT) devices being available for the workplace and the home. No sector of our lives appears to be immune from the IoT invasion, whether it be military, medical, industrial, or domestic. The Internet-of-Everything (IoE) is rapidly becoming a reality.

The term IoT is believed to have been originally coined in 1999 (Ashton, 2009), though in hindsight devices such as the wireless telegraph can be considered to be IoT devices dating back to the early to mid-19th century (Foote, 2016). In the early 2000s, the first commercial domestic appliance IoT device, the Internet refrigerator, began being shipped by some manufacturers (Osisanwo, Kuyoro, & Awodele, 2015). During 2008, the number of devices connected to the Internet was considered to exceed the number of people on earth (Evans, 2011). In 2011, IPv6 protocol became available which allowed an address space of 2^{128} unique IP addresses. This vast number prompted Steven Leibson of the Computer History Museum, Mountain View, CA at that time, to say "we could assign an IPv6 address to EVERY ATOM ON THE SURFACE OF THE EARTH, and still have enough addresses left to do another 100+ earths!" (Bhalla, 2012). By 2025, it is estimated that 75 billion IoT devices will make up a market worth between US$3.9 trillion to US$11.1 trillion (Jacobs, 2017).

As with other fields of technological development, the security of such IoT devices has predominantly taken a backseat in a rush to be first to market with a new product (Arias, Wurm, Hoang, & Jin, 2015). For this reason, IoT has been considered as the future of the Internet (Gubbi, Buyya, Marusic, & Palaniswami, 2013) or as the death of the Internet (Vaughan-Nichols, 2016) depending on one's point of view. As part of this research, we present IoT as an imminent reality which is making a positive impact on modern day living and hence the paper will discuss the following:

1. Different protocols involved in enabling IoT concept.
2. The history of the AMQP protocol, along with various features.
3. The design architecture of AMQP.
4. The known vulnerabilities of the protocol.
5. Threats posed by these vulnerabilities.

BACKGROUND

IoT consists of much more than just connected end-user devices. Underlying an IoT operation is a framework to allow data extraction from each device for analysis, and data insertion back to each device for control purposes. Such frameworks apply to all IoT deployments in use today, and will equally apply to future IoT deployments that have not yet been developed, even for so-called independent or autonomous devices.

TechBeacon (2017) describes such an architecture as a four-stage process:

1. Stage 1 – IoT devices consisting of wireless sensors and actuators.
2. Stage 2 – Sensor data aggregation systems and analogue-to-digital conversion.
3. Stage 3 – Edge IT systems for data pre-processing and transmission.
4. Stage 4 – Back-end data systems for analysis, management, and storage.

Bradicich (2015) clarifies these four stages graphically, as shown in Figure 1.

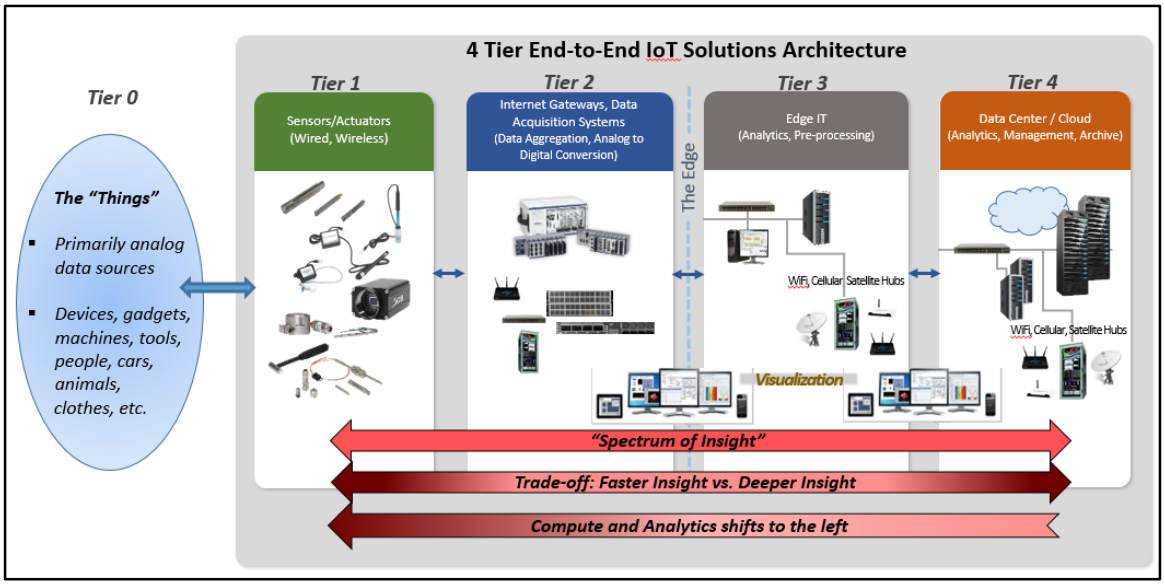


Figure 1. IoT Architecture Stages (Bradicich, 2015)

Overview of Application Layer Protocols in IoT

AMQP (AMQP, 2017a) is an open-standard application protocol middleware which enables server-to-server message communication. It is independent of both the platform and language that each of the end-servers may be using. Flow-controlled communication enables delivery options, such as at-most-once, at-least-once, or exactly-once. Delivery authentication is provided by SASL (Simple Authentication and Security Layer). Encryption is provided by TLS (Transport Layer Security), which is the successor to SSL (Secure Sockets Layer).

CoAP (Constrained Application Protocol) (CoAP, 2017) is an M2M (Machine to Machine) web transfer protocol for use with constrained nodes and constrained networks. It is based on the REST (Representational State Transfer) model. Since this is also true of HTTP, seamless connection to a web client interface is possible. CoAP uses minimal resources while providing encryption through DTLS (Datagram Transport Layer Security) equivalent to 3072-bit RSA.

DDS (Data Distribution Service) (DDS, 2017) is a protocol designed to interconnect devices and has its origins in high-performance environments, such as the military, industry and embedded systems. Its rapid delivery of messages simultaneously to an array of receivers make it suitable for real-time analytics and sensor monitoring.

IBM developed MQTT (Message Queuing Telemetry Transport) (MQTT, 2017) in 1999 and since 2013 has become an OASIS (Organization for the Advancement of Structured Information Standards) standard. It is an M2M IoT connection protocol, which uses publish/subscribe messaging transport methodology to connect a device

to a server. MQTT's lightweight architecture makes it suitable for situations where communications bandwidth is limited, such as remote locations or via satellite links.

RESTFUL (Service Architecture, 2017) provides a set of standards for communication between computer systems on the Internet. Its underlying HTTP-compliant architecture enables the communication and transfer of data between IoT devices using this protocol. It uses the same HTTP methods of GET, POST, PUT, and DELETE to conduct request/response interactions.

SMQTT (Secure Message Queuing Telemetry Transport) (SMQTT, 2017) is an extension to MQTT which encrypts a message before publication by the broker. The encrypted message is sent to multiple subscriber nodes where the message is decrypted using the same master key.

WebSocket protocol (WebSocket, 2017) provides full-duplex communication between clients and a remote server on a single TCP connection. WebSockets can be implemented in both web browsers, servers and indeed any application making use of the client/server paradigm. Apart from the initial handshake with HTTP servers, Web-Socket is an independent protocol that maintains a two-way exchange between client and server using TCP port 80 or TLS port 443 for encrypted traffic.

XMPP (Extensible Messaging and Presence Protocol) (XMPP, 2017) is one of four Instant Message (IM) protocols which has developed to satisfy the rapidly expanding information society's need for short message services. XMPP's use of Extensible Markup Language (XML) overcomes prior difficulties in connecting an IM system with a non-IM system. Several large public IM services, such as LJ Talk, Nimbuzz, and HipChat exclusively use XMPP. Other popular IM applications like WhatsApp, Gtalk and Facebook Chat use XMPP on their back-end servers.

AMQP Protocol History and General Features

AMQP was first proposed in 2003 by John O'Hara of JPMorgan Chase in London. In 2005 JPMorgan Chase commenced taking other firms on board to develop a working group for the project. These firms initially included the likes of Cisco Systems and Red Hat, but within a few years, the AMQP working group had expanded to include 23 partners.

In August 2011 the AMQP working group had attained OASIS membership, and two months later AMQP 1.0 was released at a conference in New York. Successful demonstrations of software running the protocol resulted in an OASIS Technical Committee being formed to develop the protocol into an international open standard.

In April 2014 OASIS AMQP was granted ISO and IEC International Standard approval, under the designation ISO/IEC 19464. Despite its development predominantly within the financial sector, AMQP has seen wide adoption as a server-to-server communications protocol within the IoT environment. AMQP is one of the few highly used application layer protocols with its prominent implementation India's Aadhaar project, considered as one of the world's largest biometric databases (Varma, 2010). AMQP is also being used in the Ocean Observatories Initiative infrastructure that uses sensor nets to bring readings ashore from ocean platforms and a global publisher-subscriber network to disseminate readings (Meisinger, 2010). Other notable implementations of AMQP include The Deutsche Börse, JPMorgan, NASA, AT&T (AMQP, 2017b).

AMQP offers the following protocol features (Ross, 2012):

Session Multiplexing

- Multiple Sessions can be carried over a single connection
- Sessions have independent message sequencing and flow control
- Interleaving of large messages

Full Duplex, Asynchronous Communication

- Within a session, messages can flow independently in both directions

Semantics of Message Hand-off

- AMQP formally defines the semantics of message transfer and settlement/acknowledgement
- Messages can be moved or copied
- Delivery guarantees:
  - Best Effort (Fire and Forget)
  - At Least Once

- Exactly Once
- Transactional Message Transfer
  - Local Transactions between Endpoints
  - Distributed Transactions

Data Security

- TLS/SSL
  - Encryption only, or
  - Encryption and authentication by x.509 certificates
  - Most APIs provide this option via the text of a URL: amqp://hostname vs amqps://hostname
- Extensible Authentication/Security Mechanisms
  - SASL – Simple Authentication and Security Layer
  - Supports numerous mechanisms:
    - ANONYMOUS
    - PLAIN
    - DIGEST-MD5
    - GSSAPI
    - NTLM

Flow Control

- Limits the number of messages that a producer can transfer at a time
- Greatly simplifies some difficult architectural problems
  - Many data sources sending to a data sink at the same time

Serialisation of Structured Data

- AMQP defines the wire-line format for data types
- The application designer does not need to be concerned with
  - The Endian order of a peer system
  - The Word Size of a peer system
- Rich API support is provided to access this feature
  - A Python program can send a message containing a Python Dictionary
  - A Java program can receive the message as a Hash Map (or a JMS Map Message)
  - A C++ program can receive the same message as an std::map

Message Metadata

- Like HTTP, AMQP allows messages to be annotated with headers
- Any number of application-specific headers may be placed in a message
- Headers are carried separately from the message body (which may be encrypted, encoded, and compressed)

Transport Independence

- Messages can be sent between services in many ways, all hidden from one's business logic

AMQP ARCHITECTURE

AMQP Design

The primary objective of AMQP was to replace existing proprietary and non-interoperable messaging systems that were considered to be hampering communications in the financial sector (Vermesan & Friess, 2014). It aimed to enable messaging at enterprise level between two platform systems of any configuration, provided that each of those platform systems, and the libraries that they are using, was AMQP-compliant (Indrasiri, 2016).

In AMQP, a shared queue space or broker is provided, which all participating applications can access. A message is sent from a client via a publisher to an exchange within the broker. Each message contains a routing key, which will be used by the broker to either assign the message to a particular queue, distribute the message to several queues, duplicate every message to multiple queues or distribute multiple messages to chosen queues as defined by the routing key.

A subscriber then receives its authorised messages from queues for delivery to its end client. Since each message is assigned to one particular subscriber, even though the same message may have been duplicated for another subscriber, only one subscriber can receive a particular message from any of the queues present.

Individual messages are tracked and accounted for to ensure that all messages are delivered as intended. For this to occur, all endpoints must acknowledge receipt of each message.

Richardson (2008) graphically represents the methodology of the AMQP protocol, as shown in Figure 2.

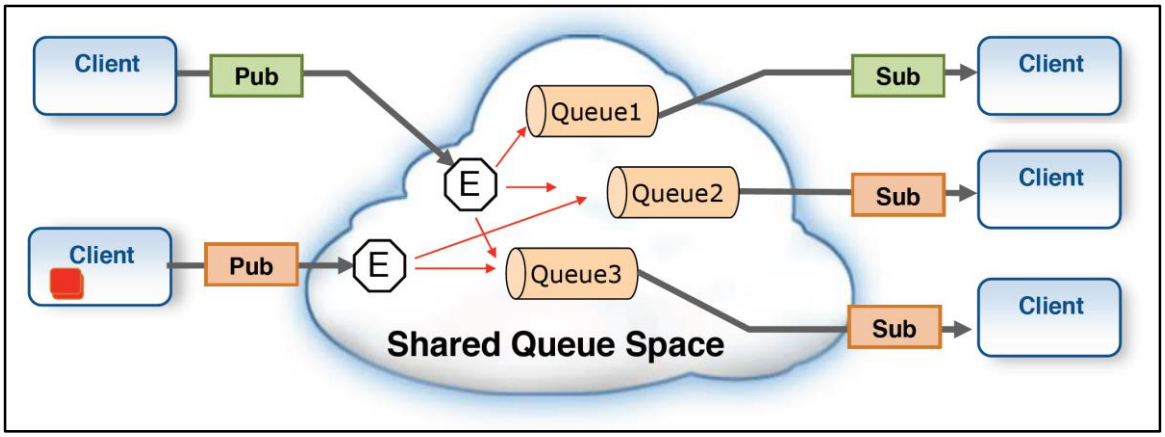


Figure 2. AMQP Methodology (Richardson, 2008)

Key: Pub = Publisher, E = Exchange, Sub = Subscriber

AMQP uses underlying TCP/IP-compatible protocols at the network and transport levels, which provides wire-level interoperability and reliability. It is a messaging solution offering flexibility, which Richardson (2008) summaries as follows:

1. any language (C, C++, C#, Python, Java, Javascript, Erlang, Lisp, Ruby, Tcl, PHP, …)
2. any model (native, .NET WCF, JMS, Mule, can do Caching)
3. any payload (binary, XML, SOAP, JSON, …)
4. any transport (TCP, SCTP, UDP, HTTP, …)
5. any scenario (desktop, router, wan, mobile, mesh, cloud)

Figure 3 shows a comparison between same-platform and cross-platform server integration, the latter being the issue AMQP was developed to resolve.

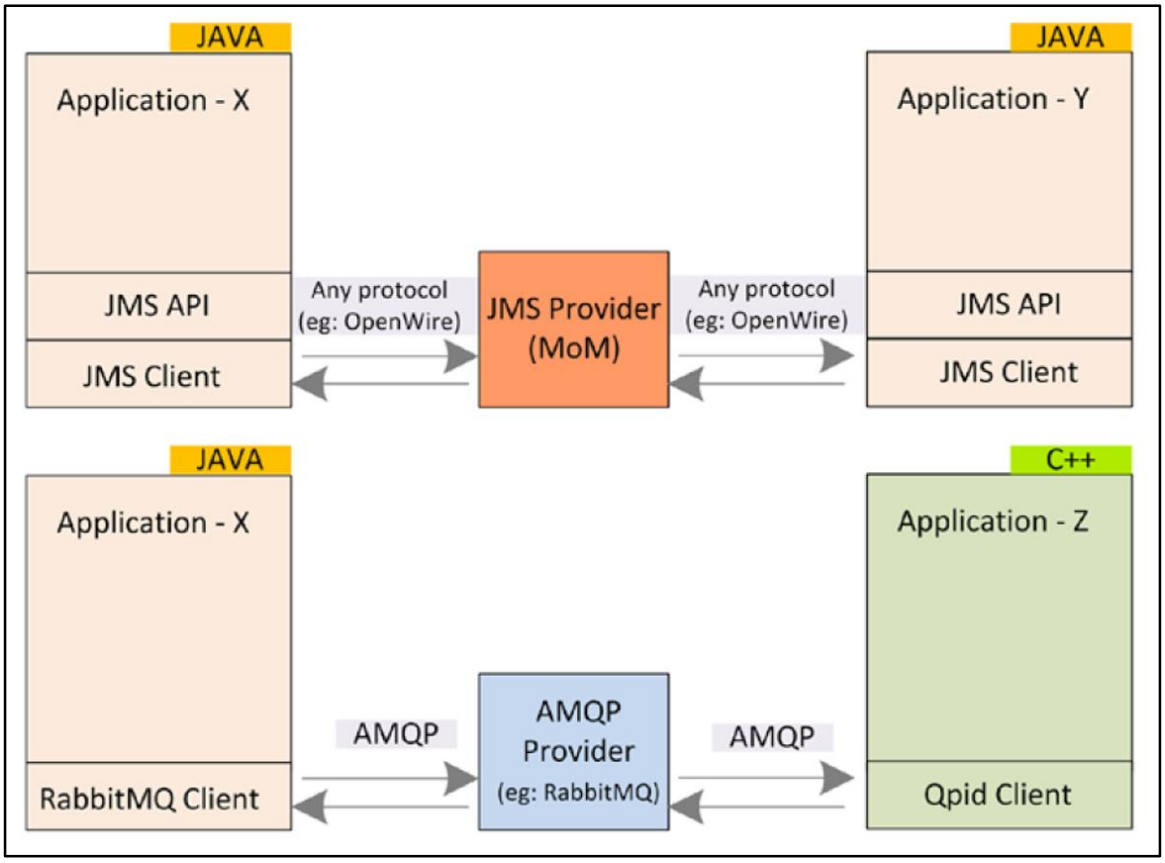


Figure 3. Same-Platform Integration via JMS vs Cross-Platform Integration via AMQP (Indrasiri, 2016)

Known Vulnerabilities

The research explored Common Vulnerabilities and Exposures (CVE) databases for the known vulnerabilities of AMQP. While searching and examining the CVE database for the terms *AMQP* and *IoT*, 17 vulnerabilities were identified with a disclosure date within the past decade:

Table 1. CVE Search Results for "AMQP IoT" (Common Vulnerabilities and Exposures, 2017)

Date Created	CVE ID	Details
18-Apr-2017	CVE-2017-7911	A Code Injection issue was discovered in CyberVision Kaa IoT Platform, Version 0.7.4. An insufficient-encapsulation vulnerability has been identified, which may allow remote code execution.
23-Mar-2017	CVE-2017-7243	Eclipse tinydtls 0.8.2 for Eclipse IoT allows remote attackers to cause a denial of service (DTLS peer crash) by sending a "Change cipher spec" packet without pre-handshake.

Date Created	CVE ID	Details
09-Mar-2017	CVE-2017-6780	A vulnerability in the TCP throttling process for Cisco IoT Field Network Director (IoT-FND) could allow an unauthenticated, remote attacker to cause the system to consume additional memory, eventually forcing the device to restart, aka Memory Exhaustion. The vulnerability is due to insufficient rate-limiting protection. An attacker could exploit this vulnerability by sending a high rate of TCP packets to a specific group of open listening ports on a targeted device. An exploit could allow the attacker to cause the system to consume additional memory. If enough available memory is consumed, the system will restart, creating a temporary denial of service (DoS) condition. The DoS condition will end after the device has finished the restart process. This vulnerability affects the following Cisco products: Connected Grid Network Management System, if running a software release prior to IoT-FND Release 4.0; IoT Field Network Director, if running a software release prior to IoT-FND Release 4.0. Cisco Bug IDs: CSCvc77164.
17-Jul-2017	CVE-2017-11408	In Wireshark 2.2.0 to 2.2.7 and 2.0.0 to 2.0.13, the AMQP dissector could crash. This was addressed in epan/dissectors/packet-amqp.c by checking for successful list dissection.
24-May-2016	CVE-2016-4974	Apache Qpid AMQP 0-x JMS client before 6.0.4 and JMS (AMQP 1.0) before 0.10.0 does not restrict the use of classes available on the classpath, which might allow remote authenticated users with permission to send messages to deserialize arbitrary objects and execute arbitrary code by leveraging a crafted serialized object in a JMS ObjectMessage that is handled by the getObject function.
02-May-2016	CVE-2016-4432	The AMQP 0-8, 0-9, 0-91, and 0-10 connection handling in Apache Qpid Java before 6.0.3 might allow remote attackers to bypass authentication and consequently perform actions via vectors related to connection state logging.
29-Jan-2016	CVE-2016-2173	org.springframework.core.serializer.DefaultDeserializer in Spring AMQP before 1.5.5 allows remote attackers to execute arbitrary code.
01-Jul-2015	CVE-2015-5240	Race condition in OpenStack Neutron before 2014.2.4 and 2015.1 before 2015.1.2, when using the ML2 plugin or the security groups AMQP API, allows remote authenticated users to bypass IP anti-spoofing controls by changing the device owner of a port to start with network: before the security group rules are applied.
25-May-2015	CVE-2015-4080	The Kankun Smart Socket device and mobile application uses a hardcoded AES 256 bit key, which makes it easier for remote attackers to (1) obtain sensitive information by sniffing the network and (2) obtain access to the device by encrypting messages.
09-May-2015	CVE-2015-2247	Unspecified vulnerability in Boosted Boards skateboards allows physically proximate attackers to modify skateboard movement, cause human injury, or cause physical damage via vectors related to an "injection attack" that blocks and hijacks a Bluetooth signal.
07-Jan-2015	CVE-2015-0862	Multiple cross-site scripting (XSS) vulnerabilities in the management web UI in the RabbitMQ management plugin before 3.4.3 allow remote authenticated users to inject arbitrary web script or HTML via (1) message details when a message is unqueued, such as headers or arguments; (2) policy names, which are not properly handled when viewing policies; (3) details for AMQP network clients, such as the version; allow remote authenticated administrators to inject arbitrary web script or HTML via (4) user names, (5) the cluster name; or allow RabbitMQ cluster administrators to (6) modify unspecified content.

Date Created	CVE ID	Details
09-Nov-2014	CVE-2014-8711	Multiple integer overflows in epan/dissectors/packet-amqp.c in the AMQP dissector in Wireshark 1.10.x before 1.10.11 and 1.12.x before 1.12.2 allow remote attackers to cause a denial of service (application crash) via a crafted amqp_0_10 PDU in a packet.
10-Apr-2014	CVE-2014-2814	Microsoft Service Bus 1.1 on Microsoft Windows Server 2008 R2 SP1 and Server 2012 Gold and R2 allows remote authenticated users to cause a denial of service (AMQP messaging outage) via crafted AMQP messages, aka "Service Bus Denial of Service Vulnerability."
12-Oct-2010	CVE-2009-5005	The Cluster::deliveredEvent function in cluster/Cluster.cpp in Apache Qpid, as used in Red Hat Enterprise MRG before 1.3 and other products, allows remote attackers to cause a denial of service (daemon crash and cluster outage) via invalid AMQP data.
21-Aug-2008	CVE-2012-4458	The AMQP type decoder in Apache Qpid 0.20 and earlier allows remote attackers to cause a denial of service (memory consumption and server crash) via a large number of zero width elements in the client-properties map in a connection.start-ok message.
21-Aug-2008	CVE-2012-4446	The default configuration for Apache Qpid 0.20 and earlier, when the federation_tag attribute is enabled, accepts AMQP connections without checking the source user ID, which allows remote attackers to bypass authentication and have other unspecified impact via an AMQP request.
14-Jun-2006	CVE-2012-3467	Apache QPID 0.14, 0.16, and earlier uses a NullAuthenticator mechanism to authenticate catch-up shadow connections to AMQP brokers, which allows remote attackers to bypass authentication.

The way the architecture of IoT has developed in recent years means that there are many protocol options available. The challenge, from the security aspect, is to ensure that the correct protocol is utilised in an appropriate environment. Often this will necessitate a solution which spans many protocols, and not just one (axway, 2015).

As with other IoT protocol software, the security of AMQP often suffers from the code being poorly written by many developers (Braue, 2015). Due to the open-source nature of AMQP, which allows for industry-specific extensions to be added, this exacerbates the variety of vulnerabilities that may be introduced to a particular messaging system based on AMQP.

Threats to AMQP Protocol

As the Internet of Things expands to encompass billions of devices around the world, the cybersecurity CIA triad of Confidentiality, Integrity, and Availability becomes as significant as ever. With an exponential growth in the number of IoT devices, so too is there a corresponding exponential growth in the number of lines of communication and data transfer, be they via wired or wireless connections. Indeed, in a situation where every device is capable of communicating with every other device, the number of communications channels equals $n(n-1)/2$, where n is the number of devices involved.

Every IoT communication channel is as vulnerable to potential man-in-the-middle cyber-attack as in a simple email communication between two end-users. The four types of such active attacks are:

1. Replay – An attack entity replays data between communication sessions to impersonate a user to obtain information.
2. Masquerade - An attack entity gains access to a system or performs a malicious act by posing as an authorised entity.
3. Modification - An attack entity performs additions or deletions to the network communication content.
4. Denial of Service – An attack that inhibits legitimate users from accessing computer services.

Despite AMQP using TLS/SSL-based encryption on an underlying TCP-based transmission protocol, resolute threat entities will still be able to intercept and decipher IoT communications, given sufficient time. Not only are we seeing IoT devices being introduced on the market with insufficient security measures (Arias et al., 2015), but we are also seeing IoT networks being compromised by the introduction of carefully-crafted botnets. An example of such an attack occurred at an unnamed University in the United States early this year (2017).

Cybercriminals were able to crack default or poorly-configured passwords in one IoT device via a brute force attack taking advantage of the device's inadequate security measures. Once this device was under their control, specially designed malware was able to be installed (Palmer, 2017).

The malware then spread from IoT device to IoT device by a botnet which again brute-forced weak or defaults passwords. As the botnet spread, it locked administrators out and repeatedly changed the password on infected devices with each malware update (Moss, 2017).

Within a short time frame, all 5,000+ devices were infected, and each device was making hundreds of DNS requests for seafood restaurants (Mezzofiore, 2017). The consequence of this DDoS attack was a severe slowing of the University's Internet access resulting in a loss of availability of resources required by students and staff (Palmer, 2017).

What makes this incident particularly interesting is that it is one of the few cases to date which has seen a botnet DDoS attack spread and then directed against the same network on which the infected devices are hosted. If such an attack is to involve the compromise of server-to-server communications hosting AMQP, then the potential would be for multiple IoT networks to be seeded with such internally-spreading infections; causing widespread compromise of AMQP-enabled IoT devices.

Figure 4 is a theoretical graphical interpretation of how the botnet attack on the University above may have been initiated and spread.

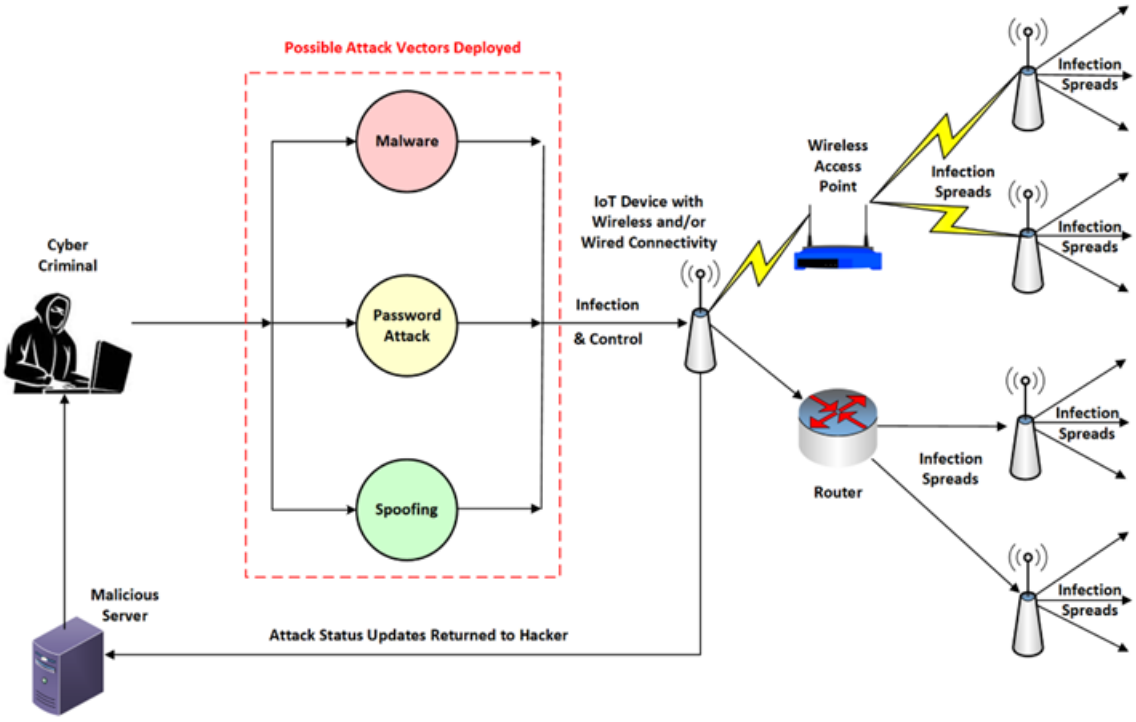


Figure 4. Hypothetical Interpretation of Attack Vector Methodology

Other software developers have created more direct AMQP-specific attack vectors. Enteletaor is a Message Queue & Broker Injection tool that can be used to implement attacks to AMQP-compatible messaging providers, such as RabbitMQ. It is also effective on non-AMQP-compatible message systems such as Redis, and ZeroMQ (GitHub, 2016). Features include:

1. Listing remote tasks.
2. Read remote task content.
3. Disconnect remote clients from Redis server (even the admin!)
4. Inject tasks into remote processes.
5. Make a scan to discover open brokers.
6. Try to discover user/passwords in authentication-protected brokers

CONCLUSIONS

From its origins in the finance sector, AMQP has developed far beyond its original scope to become a widely accepted messaging protocol throughout the field of IT, including the Internet of Things. Through the use of binding/routing keys, AMQP provides a powerful and flexible mechanism for sending messages from publishers to subscribers via exchanges and queues. AMQP's underlying TCP/IP-compatible framework adds to the diversity of its range of applications but also makes it vulnerable to an assortment of known weaknesses and methods of exploitation. Its usage as a server-to-server messaging protocol leads to the potential that its inherent TCP/IP-based vulnerabilities could be leveraged to propagate malware infections between networks. This research has reported various vulnerabilities that exist in AMQP protocol and how various threats can be used to exploit these weaknesses to make AMQP a susceptible protocol in IoTs list which otherwise has a very strong architecture. Therefore, the need for having an IoT device using AMQP protocol in a network must be carefully evaluated before AMQP becomes part of the system.

REFERENCES

AMQP. (2017a). AMQP. Retrieved from https://www.amqp.org/

AMQP. (2017b). Products and success stories. Retrieved from https://www.amqp.org/about/examples

Arias, O., Wurm, J., Hoang, K., & Jin, Y. (2015). Privacy and security in internet of things and wearable devices. *IEEE Transactions on Multi-Scale Computing Systems, 1*(2), 99-109. doi:https://10.1109/TMSCS.2015.2498605

Ashton, K. (2009). That 'internet of things' thing. Retrieved from http://www.rfidjournal.com/articles/view?4986

axway. (2015). In the Internet of Things, the thing talks back. Retrieved from https://www.axway.com/sites/default/files/resources/tools_and_tips/axway_checklist_internet_of_things_security_checklist_10_critical_issues_you_need_to_address.pdf

Bhalla, A. (2012). Making the transition to IPv6. Retrieved from http://www.wipro.com/blogs/making-the-transition-to-ipv6/

Bradicich, T. (2015). The 7 Principles of the Internet of Things (IoT). Retrieved from http://blog.iiconsortium.org/2015/07/the-7-principles-of-the-internet-of-things-iot.html

Braue, D. (2015). Small, unsophisticated developers perpetuating IoT security lapses: IBM. Retrieved from https://www.cso.com.au/article/560521/small-unsophisticated-developers-perpetuating-iot-security-lapses-ibm/

CoAP. (2017). CoAP. Retrieved from http://coap.technology/

Common Vulnerabilities and Exposures. (2017). Search results. Retrieved from https://cve.mitre.org/cgi-bin/cvekey.cgi?keyword=amqp+iot

DDS. (2017). DDS. Retrieved from http://portals.omg.org/dds/

Evans, D. (2011). The internet of things [Infographic]. Retrieved from http://blogs.cisco.com/diversity/the-internet-of-things-infographic

Foote, K. D. (2016). A brief history of the internet of things. Retrieved from http://www.dataversity.net/brief-history-internet-things/

GitHub. (2016). Enteletaor. Retrieved from https://github.com/cr0hn/enteletaor

Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. *Future Gener. Comput. Syst., 29*(7), 1645-1660. doi:http://10.1016/j.future.2013.01.010

Indrasiri, K. (2016). *Beginning WSO2 ESB*. Berkeley, CA: Apress.

Jacobs, J. (2017). What the future might hold for the internet of things. Retrieved from https://www.globalxfunds.com/what-the-future-might-hold-for-the-internet-of-things/

Meisinger, M. (2010). Ocean observatories initiative messaging service. Retrieved from https://confluence.oceanobservatories.org/display/CIDev/Messaging+Service

Mezzofiore, G. (2017). A university was attacked by its lightbulbs, vending machines and lamp posts. Retrieved from http://mashable.com/2017/02/13/internet-of-things-university-network-/#KXJv7vGDqOqj

Moss, S. (2017). University suffers DDoS attack from IoT vending machines. Retrieved from http://www.datacenterdynamics.com/content-tracks/security-risk/university-suffers-ddos-attack-from-iot-vending-machines/97808.fullarticle

MQTT. (2017). MQTT. Retrieved from http://mqtt.org/

Osisanwo, F., Kuyoro, S., & Awodele, O. (2015). *Internet refrigerator – A typical internet of things (IoT)*. Paper presented at the 3rd International Conference on Advances in Engineering Sciences & Applied Mathematics (ICAESAM'2015), London (UK). http://iieng.org/images/proceedings_pdf/2602E0315051.pdf

Palmer, D. (2017). How IoT hackers turned a university's network against itself. Retrieved from http://www.zdnet.com/article/how-iot-hackers-turned-a-universitys-network-against-itself/

Richardson, A. (2008). *AMQP business messaging for predictable, scalable, available SOA*. Paper presented at the Microsoft Architects Insight Conference 2008. http://download.microsoft.com/documents/uk/msdn/events/sol/sol03.pdf

Ross, T. (2012). Integrating the internet of things with AMQP. Retrieved from https://www.scribd.com/document/234021715/AMQP-IoT-pdf

Service Architecture. (2017). Representational State Transfer (REST). Retrieved from http://www.service-architecture.com/articles/web-services/representational_state_transfer_rest.html

SMQTT. (2017). SMQTT. Retrieved from http://smqtt.com/

TechBeacon. (2017). The 4 stages of an IoT architecture. Retrieved from https://techbeacon.com/4-stages-iot-architecture

Varma, P. K. (2010). Aadhaar: Scalability & data management challenges. Retrieved from https://www.cse.iitb.ac.in/~comad/2010/pdf/Industry%20Sessions/UID_Pramod_Varma.pdf

Vaughan-Nichols, S. J. (2016). Death of the internet: GIF at 11. Retrieved from http://www.zdnet.com/article/death-of-the-internet-gif-at-11/

WebSocket. (2017). About HTML5 WebSocket. Retrieved from https://www.websocket.org/aboutwebsocket.html

XMPP. (2017). XMPP. Retrieved from https://xmpp.org/

A CRITICAL ANALYSIS OF SECURITY VULNERABILITIES AND COUNTERMEASURES IN A SMART SHIP SYSTEM

Dennis Bothur, Guanglou Zheng, Craig Valli
Security Research Institute, School of Science, Edith Cowan University, Perth, Western Australia
d.bothur@ecu.edu.au, g.zheng@ecu.edu.au, c.valli@ecu.edu.au

Abstract

It is timely to raise cyber security awareness while attacks on maritime infrastructure have not yet gained critical momentum. This paper analyses vulnerabilities in existing shipborne systems and a range of measures to protect them. It discusses Information Technology network flaws, describes issues with Industrial Control Systems, and lays out major weaknesses in the Automated Identification System, Electronic Chart Display Information System and Very Small Aperture Terminals. The countermeasures relate to the concept of "Defence-in-depth", and describe procedural and technical solutions. The maritime sector is interconnected and exposed to cyber threats. Internet satellite connections are feasible and omnipresent on vessels, offshore platforms and even submarines. It enables services that are critical for safety and rescue operations, navigation and communication in a physically remote environment. Remote control of processes and machinery brings benefits for safety and efficiency and commercial pressure drives the development and adaptation of new technologies. These advancements include sensor fusion, augmented reality and artificial intelligence and will lead the way to the paradigm of "smart" shipping. Forecasts suggest unmanned, autonomous ships in international waters by 2035. This paper is the starting point for future research, to help mapping out the risks and protect the maritime community from cyber threats.

Keywords: maritime cyber security, smart shipping, autonomous shipping, vulnerabilities, and countermeasures

INTRODUCTION

Geographical isolation exposes mariners to a set of unique challenges such as navigating through rough weather and evading pirate attacks. Technology on ships plays a significant role to help manoeuvring through those conditions and it enables communication in situations of emergency and distress. Unfortunately, any type of technology has the potential to be used for malicious purposes. Cyber security awareness and culture is new on the agenda of the maritime community, but it must be taken seriously to avoid catastrophic consequences.

Universal satellite and data connectivity is one of the major advancements in seafaring, but this brings along a myriad of new risks. For instance, many critical systems on board rely on the Global Navigation Satellite System (GNSS) for safe navigation, communication, emergency response, and traffic control. However, disrupted or manipulated Global Positioning System (GPS) signals can send ships off their course and cause collisions, groundings, and environmental disasters. In 2016, multiple ships outbound from the United States (U.S.) reported GPS interferences which prompted the US Coast Guard to issue "Safety Alert 01-16 – GNSS – Trust, but Verify. Report Disruptions Immediately" (United States Coast Guard, 2016). In 2017, reports emerged of more than 20 vessels which noticed spoofed GPS signals that placed them about 25 nautical miles inland (Hambling, 2017). The source of the attack was attributed to tests performed by a nation-state. Adversaries are "testing the waters" but they already have the knowledge, tools, and motivation to launch attacks with potentially devastating outcomes. It is very alarming when we consider that this applies to naval vessels carrying advanced weaponry as well as the commercial shipping sector, which is part of the critical infrastructure and accounts for more than 90% of cargo transported globally (National Institute of Standards and Technology, 2017).

A host of weak spots in ship- and shore-based cyber systems has already been exposed by research conducted in the field. Unawareness or ignorance of these flaws leads many organisations to taking shortcuts in regard to applying and policing appropriate security measures. Additionally, rapid cycles of product development, implementation, maintenance, and decommissioning are overwhelming for the majority of maritime stakeholders.

The following section outlines critical vulnerabilities in common IT systems and Industrial Control Systems (ICS) on board. It explains the risks related to the heavy reliance on navigation and communication systems such as the Electronic Chart Display Information System (ECDIS), the Automated Identification System (AIS), and Very Small Aperture Terminals (VSATs). The section *Countermeasures* lays out current procedural and technological strategies to protect maritime infrastructure from malicious attacks and it describes the concept of *Defence in Depth*.

VULNERABILITIES

Vulnerabilities are flaws in a system that have the potential to be exploited by malicious parties. This section outlines vulnerabilities of several technologies used in IT networks, industrial control systems, navigation, and communication systems.

The purpose of critical services in the maritime community is to ensure the safety of people, equipment, and the environment. For example, the Global Maritime Distress and Safety System (GMDSS) is a set of standards and components to aid search and rescue operations for vessels in emergency situations. Each component of the GMDSS (e.g. VSAT terminals, AIS transponders) comes with a set of potential vulnerabilities. Other services include voice communications, crew welfare and entertainment systems, guest Wi-Fi, and video monitoring. These systems are perceived to be less critical to safety and operations and thus routinely left unpatched and exposed to attacks. Figure 1 summarises critical on-board systems which could be vulnerable to cyber-attacks. The weaknesses of each of these systems are explained in the sections below.

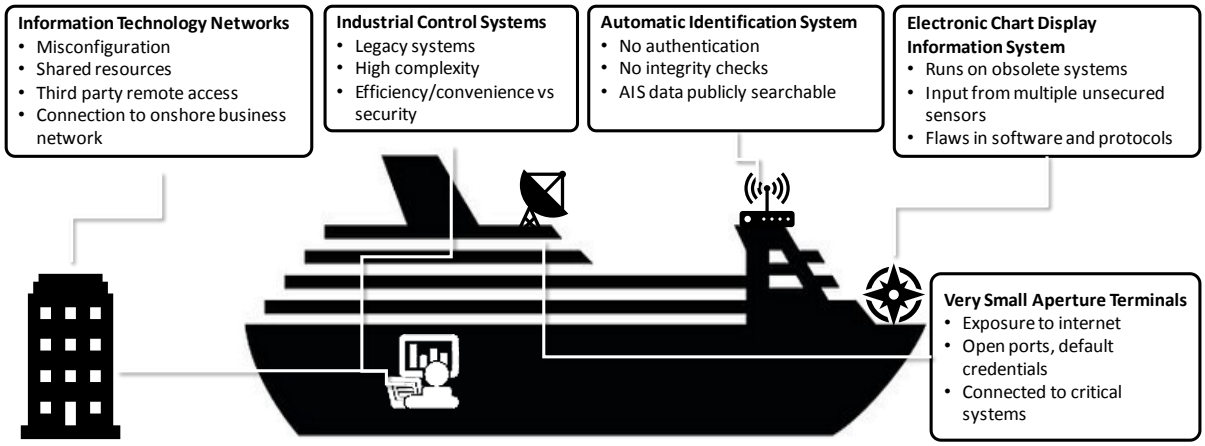


Figure 1 On-board systems where vulnerabilities could exist.

IT Networks

An Information Technology (IT) network is the fabric that integrates core business and operation systems on board and leverages shared databases and other systems. These systems can be used for accounting, cargo management, customs and shipping, human resource planning, and administration (Hudson Analytix Inc, 2017).

A malware outbreak in June 2017 paralysed IT networks across the world and caused significant business disruptions and loss of revenue. The worm "NotPetya" initially infected computers through a malicious update in an accounting software product. It then spread to attached systems, wiping or encrypting files and demanding ransom payments (Symantec Security Response, 2017). The Danish shipping company A.P. Moller-Maersk was one of the organisations that were hit the hardest and it reported a loss of $300M due to the significant system shutdowns and restoration efforts across its critical systems (Mimoso, 2017).

The event underlines that many organisations still lack a coherent approach to managing the cyber security of their systems. Business-critical software may be not updated or replaced because it is only compatible with other legacy systems or protocols. Some organisations do not have a regular patching or update regime and thus their antivirus software is outdated, or important application security updates are not patched. Networks are often inadequately segmented to manage access control, especially for third parties. IT systems should be scrutinised carefully as they provide a wide attack surface and many entry points for adversaries. All systems and endpoints must be secured, which in most cases does not happen. Third party access (e.g. for implementation, support and maintenance of equipment) further exposes vulnerable systems to the open world.

Critical control networks should be in a secured zone, isolated from the corporate IT network and the internet. However, economic pressure, regulations, and requirements for remote monitoring and control increase the need for a connection into the IT network. The design and configuration of the links between IT networks rarely consider authentication and encryption methods, thus exposing potential vulnerable and legacy system to the internet. IT systems on vessels are often connected with onshore facilities and this further increases the exposure to systemic and persistent threats (Baltic and International Maritime Council, 2017).

ICS – Industrial Control Systems

Industrial Control Systems (ICS) on ships assist to reduce human errors, increase resource efficiency, prolong equipment life, and ensure economic advantages (Wang & Zhang, 2000). ICS control and monitor parameters on board, including temperature, pressure, level, viscosity, flow control, speed, torque, voltage, current,

machinery and equipment status (Zaghloul, 2014). An array of devices and protocols from different vendors and technological eras and are often "bolted together" to provide interoperability. Most of these components were designed and programmed without any security in mind and data is transferred in plaintext. The onus of *securing* the components should be shared between the vendor, who follows a secure development framework, and the operator, who configures the components in line with industry standards and recommendations. The reality is often that either party assumes that the other party is responsible and no-one does it at all, leaving many critical weaknesses for attackers to exploit (Shoultz, 2017). It is crucial for integrators, implementers, and operators of ICS to understand the system's limitations and vulnerabilities of its components and protocols.

Primary control systems (hydraulic, electrical, automatic control) are vital to the ship's safe voyage. They are exposed to difficult environmental challenges, such as pressure, vibration, and humidity. These control systems are integrated via the ship's distributed IT network. A continuous link between IT networks and on-shore facilities enables remote access for monitoring, fault-finding, and troubleshooting, reduces site travel costs, and streamlines the collection and analysis of field data (Moxa Inc., 2017; Orbcomm, 2017). A major concern is that operators and engineers routinely bypass security for convenience and efficiency, which could have a cascading effect on the entire organisation (Zurich, 2014). This behaviour is attributed to the lack of awareness and skills, the commercial pressure to save time, and the plain non-adherence to security policies.

AIS – Automatic Identification System

The Automatic Identification System (AIS) is a ship- and shore-based Very High Frequency (VHF) radio broadcasting system. It is used for Vessel Traffic Services (VTS), search and rescue operations, accident investigation, and weather forecast (Australian Maritime Safety Authority). Reliance on the transmitted information is critical to situational awareness and collision avoidance at sea. AIS transponders communicate over the air without any authentication or integrity checks. Attackers can inject signals via a Software Defined Radio (SDR) and place fake "man-in-water" beacons, render the ship invisible and inject false weather reports (Balduzzi, Wihoit, & Pasta, 2013). Relying on the potentially incorrect information can lead to wrong decisions and catastrophic outcomes.

AIS data is publicly available via websites such as "Vesselfinder" (VesselFinder Ltd, 2017) and "Marinetraffic" (MarineTraffic, 2017). The International Maritime Organisation (IMO) has "condemned the regrettable publication on the world-wide web, or elsewhere" as it reveals a wealth of information on the vessel and its route which can be invaluable for a targeted attack (International Maritime Organisation, 2004).

ECDIS – Electronic Chart Display Information System

The Electronic Chart Display Information System (ECDIS) is mandated by the IMO for all commercial vessels and usually is installed on the bridge. ECDIS software implementations have an extensive list of weaknesses. Often the system runs on legacy computers (e.g. Windows XP desktops) for which no security updates are available. The maps are loaded onto the system either via the internet, or manually via USB or DVD. Sensor feeds comes from a multitude of other onboard systems such as Radar, Navigation Telex (Navtex), ICS, and satellite terminals. This provides a wide surface for a compromise. Dyryavyy (2014) audited commercial ECDIS software and highlighted some significant security risks that would allow an attacker to replace or delete files on the system or inject malicious content. Thus, tampered sensor data could be sent to ECDIS, which would influence decisions for navigation, and may cause collision or grounding.

VSAT – Very Small Aperture Terminal

A Very Small Aperture Terminal (VSAT) is a communications station used to send and receive data via a satellite network. The transceiver is installed above deck in line of sight of the satellite and a control unit below deck provides the interface to a PC. VSATs enable a range of communication and safety services including GMDSS, ECDIS, AIS, phone, internet, cargo management, vessel routing, radio integration, telemedicine, crew welfare, tele-training, and weather forecast. Santamarta (2014) tested a range of VSATs from multiple manufacturers and concluded that *all* the audited devices are vulnerable at the protocol and implementation level. They transmit in plain text without authentication, encryption, or integrity checks. This can allow an attacker to inject fake signals or malicious code to cause device to shut down or corrupt the system, disabling the ship from navigating safely.

A ship's geolocation is publicly available via AIS aggregators, but the real risk is that VSAT network interfaces can be identified on the internet, e.g. via the "Shodan Ship Tracker" (Matherly, 2017)). This can reveal manufacturer names, product codes, and other data which is useful for a potential attack. Vendors generally publish default credentials on their websites and many terminals run with unchanged default factory settings, including administrator usernames and passwords. Once an attacker found an open VSAT interface, they can change GPS coordinates, settings, and upload malicious software. This allows for further compromise of the network and may give up an entry point to critical control systems (Morse, 2017).

It is concerning that these systems are widely used by NATO and across critical infrastructure, as their exploitation could lead to catastrophic consequences.

COUNTERMEASURES

The concept of cyber security is novel to many maritime stakeholders and it is timely to raise awareness about the existing countermeasures. It is essential to create a common understanding that includes the shared responsibility amongst maritime stakeholders. The following section outlines a strategic direction to securing cyber technology on ships.

Defence-in-depth

Security is neither a product that can be bought off the shelf, nor a procedural blueprint that every organisation can apply in the same way. Securing maritime IT environments "in depth" creates an all-encompassing protection mantle and builds resilience to external and internal threats. This layered approach is depicted in Figure 2 and includes procedural and technical countermeasures on each layer (outlined thereafter).

Policy: Defence begins with the organisation's leadership, where *strategies* are formed, and *policies* are made.

Physical security: Physical measures to prevent intruders from entering the vessel by using guards, locks, alarms, and technical access control.

Perimeter security refers to measures which block attacks from entering the network through external communication connections.

Network security is concerned with the design, configuration and implementation of security zones, network segments, and other network based defences.

The layer of *Host security* protects computers and other endpoints with measures such as antivirus software and host based firewalls.

Application security prevents attackers from exploiting software flaws and entails secure software development, authentication, access control, and application vulnerability management.

Finally, the *Data security* layer addresses how to protect the information itself, whether it is in use, in transit, or at rest. Each layer plays a significant role in the overall security of maritime critical infrastructure.

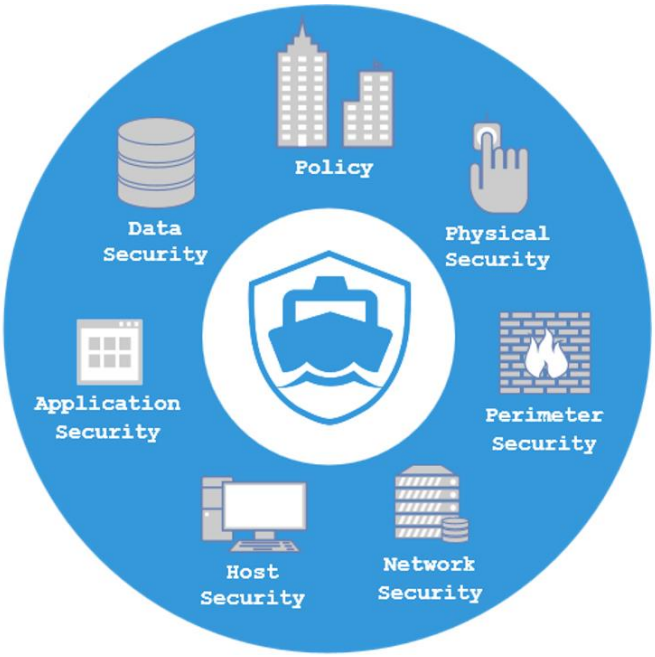


Figure 2 A systematic approach to defend the smart ship in depth

Security policies and procedures

The most important security asset on board is a vigilant employee. Policies and procedures give staff the necessary tools and guidance for their critical role within the organisation. Training and awareness programs enable staff to understand why and how each individual can help. Policy and procedural documents should be clearly communicated, published, and acknowledged. Policies must address how the content will be enforced and what the consequences are if ignored. The documents must be reviewed regularly to ensure they appropriately cover the organisation in a world of continuous technological advancements.

The policies should address and explain at least the following:

- Data recovery capability, backups, redundancy, business continuity and disaster recovery planning
- Administrator privileges, concepts of least privilege and the separation of duties
- Remote access control, use of encryption and Virtual Private Networks (VPN)
- Physical access, removable media controls, "Bring your own device" (BYOD)

- Acceptable personal use of IT systems
- Email, phishing, passwords rules
- Software upgrade, patch, and maintenance schedules
- Anti-virus/-malware software and signature updates
- White- or blacklisting and the use of third party software
- Onshore support and contingency planning
- Equipment disposal, and data destruction

(Baltic and International Maritime Council, 2017; International Maritime Organisation, 2016)

Technical security solutions

Policies require a physical or technical implementation to be monitored and enforced. Guards, locks, and security cameras protect equipment like the ECDIS and VSATs from unauthorised use. Secure network design and configuration should segment the network to prevent the direct exposure of devices and ICS to the internet. For example, a VSAT hub can act as an intermediate hop for remote connections to the terminal. Firewalls and intrusion prevention systems monitor and block the data traffic as it leaves and enters the ship's IT network. The dataflow between all nodes on the network, including ICS traffic and satellite and radio communications, should be mapped out and encrypted, e.g. by using a VPN. This way, even if signals were intercepted, the adversary could not easily read the message.

Network hardening refers to the secure configuration of hardware and software and the deactivation of unused features and accounts. It applies to firewalls, routers, switches, servers, voice communication equipment, and any other device on the network. Hardening includes disabling unused ports and services but also managing and installing updates, patches and bugfixes. Default usernames and passwords must be changed where possible. The use of complex passwords protects against automated port-scan and dictionary based login attempts, for example on the VSAT terminal. Access control systems should be configured and audited regularly to only allow users the access rights they need to perform their job.

X.509 certificate based authentication can secure the access to the ship's wireless network for authorised crew members and guests. It is recommended to create a separate wireless network (Virtual Local Area Network – VLAN) for guests to allow only minimal access to resources on the network. The usage of secure communications protocols like ssh, https, and sftp should be implemented and enforced where it is possible. Multi-Factor Authentication (MFA) can provide an additional layer of access security to sensitive systems and applications. Application whitelisting prevents staff from installing unapproved and potentially malicious programs. The threat of intentional or accidental data leakage can be mitigated with data-loss-prevention software (Mertens, 2014; Shoultz, 2017; Soullie, 2014).

DISCUSSION AND CONCLUSION

This paper analysed the current cyber security vulnerabilities and countermeasures in smart ship systems. It demonstrated that malicious attacks and incidents with devastating consequences are not only possible, but imminent. The maritime domain leverages cyber technologies to assure the safety and efficiency of operations at sea but vessels, platforms, satellites, and onshore facilities are increasingly interconnected, exposing them to an abundance of systemic and technology based threats.

The paper outlined the flaws in existing information technology networks on board and concluded that the high level of complexity and shared resources create a wide attack surface which should be segmented and secured systematically. Industrial Control Systems are often built on legacy infrastructure and implemented without security in consideration. Networking capabilities supplement these systems to allow remote control and troubleshooting but at the same time expose them to the IT network and its inherent vulnerabilities.

Vessel operators depend on the Automated Identification System for traffic and emergency services. Flaws discovered in the underlying hardware, software, and protocols would allow an adversary to manipulate the transmitted data and this could lead to navigational decisions with devastating outcomes. The Electronic Chart Display Information System equipment is equally unsecure and susceptible to malicious tampering. It often runs on legacy systems with well published vulnerabilities and the software itself can be misused to manipulate maps and sensor data which can compromise the safe navigation at sea.

Very Small Aperture Terminals provide geospatial capabilities for many critical services on board. The data flow between satellites and terminals is unencrypted and provides no integrity checks. Terminals can be tracked over the public internet and many interfaces are not locked down. An attacker could remotely log in to the exposed device and change critical information as well as using it as an entry point to pivot through the connected ship network.

The presented countermeasures were explained in the context of the multi-layered "Defence-in-depth" approach. Further recommendations were based on the implementation of security related policies and procedures. It was

suggested that technical security solutions are needed to implement and enforce said policies, and examples were presented relating to each layer of the "Defence-in-depth" approach.

The maritime domain is moving fast towards the paradigm of "smart" transportation with more innovative technologies and AI driven decision-making. While this implies a wealth of benefits for economy, critical infrastructure, and military, it also increases the complexity of the threats to the maritime community before we have discovered and mitigated the flaws in existing systems and legacy technology. The current coverage of research in the field of maritime cyber security is sporadic and fragmented, leaving many stakeholders unaware of the risks and unprepared to treat them.

The cyber security community has the urgent obligation to support the shipping industry with research, tools, security assessments, and education programs. In doing so, we will enable organisations to make informed strategic decisions and effectively allocate resources to protect each member and the maritime community as a whole.

Future work based on the present paper will aim to map all identifiable risks and to establish a status quo of the cyber security posture of the maritime sector.

REFERENCES

Australian Maritime Safety Authority. Automatic Identification System (AIS). Retrieved 01/09/2017 from https://www.amsa.gov.au/navigation/services/ais/.

Balduzzi, M., Wihoit, K., & Pasta, A. (2013). *Hey captain, where's your ship? attacking vessel tracking systems for fun and profit*. Paper presented at the The Eleventh Annual Hack in the Box (HITB) Security Conference in Asia. http://conference.hitb.org/hitbsecconf2013kul/materials/D1T1%20-%20Marco%20Balduzzi,%20Kyle%20Wilhoit%20Alessandro%20Pasta%20-%20Attacking%20Vessel%20Tracking%20Systems%20for%20Fun%20and%20Profit.pdf

Baltic and International Maritime Council. (2017). *The Guidelines on Cyber Security Onboard Ships*. Retrieved from https://www.bimco.org/-/media/bimco/news-and-trends/news/security/cyber-security/2017/industry-guidelines-cyber-security---june-2017.ashx

Dyryavyy, Y. (2014). *Preparing for Cyber Battleships – Electronic Chart Display and Information System Security*. Retrieved from https://www.nccgroup.trust/au/our-research/preparing-for-cyber-battleships-electronic-chart-display-and-information-system-security/

Hambling, D. (2017). Ships fooled in GPS spoofing attack suggest Russian cyberweapon. Retrieved 28/08/2017 from https://www.newscientist.com/article/2143499-ships-fooled-in-gps-spoofing-attack-suggest-russian-cyberweapon/#.WZy1mN2_kyQ.linkedin.

Hudson Analytix Inc. (2017). *Global Threats: Cybersecurity in Ports (Donald Duck, Daughters & Dollars)*. Paper presented at the Hemispheric Conference on Port Competitiveness & Security: Finding the Right Balance, University of Miami, Center for International; Business Education & Research. http://portalcip.org/wp-content/uploads/2017/03/Max-Bobys.pdf

International Maritime Organisation. (2004). *MSC 79/23 - Report of the Maritime Safety Committee on its Seventy-Ninth Session*. Retrieved from http://www.crs.hr/Portals/0/docs/eng/imo_iacs_eu/imo/msc_reports/MSC79-23.pdf?ver=2010-11-03-143734-000

International Maritime Organisation. (2016). *Measures to enhance Maritime Security: Report of the Working Group* (MSC 96/WP.9). Retrieved from http://12zc4845uhr73vbfjp3ubgkz.wpengine.netdna-cdn.com/wp-content/uploads/2016/05/Cyber-guidelines.pdf

MarineTraffic. (2017). marinetraffic.com. Retrieved 15/10/2017 from https://www.marinetraffic.com/en/p/contact-us.

Matherly, J. (2017). Shodan Ship Tracker. Retrieved 15/10/2017 from https://shiptracker.shodan.io/.

Mertens, M. (2014). Securing VSAT Terminals. *newtec.eu*. Retrieved 13/09/2017 from http://www.newtec.eu/article/article/securing-vsat-terminals.

Mimoso, M. (2017). Maersk Shipping Reports $300M Loss Stemming from NotPetya Attack. Retrieved 16/10/2017 from https://threatpost.com/maersk-shipping-reports-300m-loss-stemming-from-notpetya-attack/127477/.

Morse, J. (2017). Remotely hacking ships shouldn't be this easy, and yet ... *mashable.com*. Retrieved 12/09/2017 from http://mashable.com/2017/07/18/hacking-boats-is-fun-and-easy/#mjW1KLCj6aqb.

Moxa Inc. (2017). Industrial Ethernet for In-ship Communication. Retrieved from https://www.moxa.com/event/Net/2010/Maritime_microsite/In-ship_solution.htm.

National Institute of Standards and Technology. (2017). *Framework for Improving Critical Infrastructure Cybersecurity Draft Version 1.1*. Retrieved from https://www.nist.gov/sites/default/files/documents/cyberframework/cybersecurity-framework-021214.pdf

Orbcomm. (2017). SCADA System Monitoring. Retrieved 05/09/2017 from https://www.orbcomm.com/en/industries/natural-resources/scada-system-monitoring.

Santamarta, R. (2014). SATCOM terminals: Hacking by air, sea, and land. Retrieved 08/09/2017 from https://www.blackhat.com/docs/us-14/materials/us-14-Santamarta-SATCOM-Terminals-Hacking-By-Air-Sea-And-Land-WP.pdf.

Shoultz, D. (2017). Securely Connected Vessels: Vessel Communications and Maritime Cybersecurity. *maritimeprofessional.com.* Retrieved 07/09/2017 from https://www.maritimeprofessional.com/blogs/post/securely-connected-vessels-vessel-communications-and-maritime-15176.

Soullie, A. (2014). *Pentesting PLCs 101*. Paper presented at the Blackhat Europe 2014. https://www.blackhat.com/docs/eu-14/materials/eu-14-Soullie-Industrial-Control-Systems-Pentesting-PLCs-101.pdf

Symantec Security Response. (2017). Petya ransomware outbreak: Here's what you need to know. Retrieved from https://www.symantec.com/connect/blogs/petya-ransomware-outbreak-here-s-what-you-need-know.

United States Coast Guard. (2016). Safety Alert 01-16 - Global Navigation Satellite Systems - Trust, but Verify. Retrieved 06/10/2017 from http://www.dco.uscg.mil/Portals/9/DCO%20Documents/5p/CG-5PC/INV/Alerts/0116.pdf.

VesselFinder Ltd. (2017). Vessel Finder. Retrieved 15/10/2017 from https://www.vesselfinder.com/contact.

Wang, J., & Zhang, S. M. (2000). Management of human error in shipping operations. *Professional Safety*, *45*(10), 23-28.

Zaghloul, M. S. (2014). Online Ship Control System Using Supervisory Control and Data Acquisition (SCADA). *International Journal of Computer Science and Application*.

Zurich. (2014). *Beyond data breaches: global interconnections of cyber risk*. Retrieved from https://www.jasadvisors.com/custom/uploads/2014/04/Risk-After-Next-Whitepaper.pdf

THE 2017 HOMOGRAPH BROWSER ATTACK MITIGATION SURVEY

Tyson McElroy[1,2], Peter Hannay[1,2], Greg Baatard[2]
[1]Security Research Institute, [2] School of Science, Edith Cowan University, Perth, Western Australia
tjmcelro@our.ecu.edu.au, p.hannay@ecu.edu.au, g.baatard@ecu.edu.au

Abstract
Since their inception, International Domain Names (IDN) have allowed for non-Latin characters to be entered into domain names. This feature has led to attackers forging malicious domains which appear identical to the Latin counterpart. This is achieved through using non-Latin characters which appear identical to their Latin counterpart. This attack is referred to as a Homograph attack. This research continues the work of Hannay and Bolan (2009), and Hannay and Baatard (2012), which assessed the mitigation methods incorporated by web browsers in mitigating IDN homograph attacks. Since these works, time IDN mitigation algorithms have been altered, such as the one used in Mozilla Firefox (Gerv, 2017). This study evaluates browser homograph attack mitigation strategies in browsers released post-2011. In this study, we find a high level of effective multi-script mitigation across the browser families surveyed. Notable exceptions to this include a single version of Firefox in which the mitigation features were not present and ongoing omission of mitigation against single script attacks.

Keywords: IDN, homograph, homoglyph, internationalised domain names, browser security, phishing

BACKGROUND

Domain name resolution is a technology which allows for IP addresses to be encoded as a string of characters. In the early days of Domain Names, the technology only accepted a string of alphanumeric ASCII characters as input for domain names (Mockapetris, 1987). This limitation prevented international users from accessing certain domains in their respective languages. The introduction of IDNs allowed for specific domain names to be accessible to multiple languages through encoding domain names in the Unicode format ("Introduction to IDNs," 2016).

The proposed IDN solution made use of UTF-8 character encoding to allow for non-latin characters to be displayed. In order to enable existing DNS infrastructure to handle UTF-8 domains a system known as Punycode was developed. Punycode provides facility to represent IDNs as regular ASCII domain names, as such no changes are required for the majority of infrastructure (Costello, 2003). An example of an IDN would be the domain name 🐻.com, which would be represented as xn--n3h.com when converted to punycode.

IDN's have allowed for domains to be accessible from many global users. However, the use of Unicode characters has also allowed for phishing attacks to be possible against domain names (Spaulding, Upadhyaya, & Mohaisen, 2016). The use of non-Latin characters in the domain names allowed users to enter Unicode characters which appeared identical to the Latin counterpart, an example of this concept is shown in Figure 1 (Krammer, 2006). As a method to counteract these threats, Domains adopted a system known as Punycode to translate any Unicode characters into their ASCII representation ("Introduction to IDNs," 2016). Many browsers have adopted this technology, allowing users to see when a non-ASCII character is entered into a domain name.

Figure 1 - Example of Homoglyph for "g"

IDN's have continued to see considerable growth and development in recent years ("Key Numbers," 2016), reported that from 2010 to 2015, the total number of registered IDN's had doubled to 6.8 million. Despite browsers incorporating mitigation actions against IDN phishing attacks, they still pose a significant threat. In 2017 browsers such as Firefox, Google, Chrome and Opera were found to be vulnerable to a Punycode exploit (Kumar, 2017). This attack was possible due to a domain registration exploit, which allows the user to register a domain name in

Punycode format with foreign ASCII characters which appear identical to legitimate domains (Tseng, Ku, Lu, Wang, & Geng, 2013).

In the Wild Phishing Attacks using IDN Homographs

There have been numerous websites aiming to educate the public about the dangers of IDN homograph attacks, including epic.com, apple.com, and google.com (Hannay, 2012; Maunder, 2017; Zheng, 2017). However, the occurrence of homograph usage in confirmed phishing campaigns has been relatively minor with only a single confirmed major incident. In August the domain adobe.com was registered, subsequently the domain was used to distribute the Beta Bot malware, disguised as an update to the Adobe Flash Player software (Mimoso, 2017). Links to the website were distributed via email and Skype messages, requesting that the user install the fabricated update. Post infection the malware disables security software, then steals financial data and user credentials (Kaspersky Lab, 2017).

PREVIOUS WORK

Since the introduction of IDN's, many attacks have been possible due to the use of non-ASCII characters. The following literature review gives a brief overview of the types of attacks due to the introduction of IDNs. Table 1 provides an overview of of the types of attacks, covering single script, mixed-script, and whole script spoofing attacks which can take place through character substitution.

Table 1 – Examples of Single Script, Mixed Script, and Whole Script Spoofing

Script Type	String	Punycode	Comments
Single Script	EPlC.com	EPlC.com	Lowercase L used as replacement for uppercase I
Mixed Script	epic.com	xn--pic-qdd.com	Cryllic replacement used for 'e'
Whole Script	epic.com	xn--e1awd7f.com	Cryllic replacements used for 'epic'

Mixed-Script Spoofing

One of the most prevalent attacks which have been proven possible by IDN domains is mixed-script spoofing. Mixed-script spoofing generates domain names using visually indistinguishable characters form different script groups (Krammer, 2006). These characters appear almost the same to the end user but contain different Unicode values. These characters exist within Unicode due to the writing system being used, letter and number encodings, and legacy encoding values (Davis & Suignard, 2006). These visually indistinguishable characters are known as Homoglyphs. Due to the indistinguishable nature of certain Unicode characters, attackers can easily forge domain names which appear to be visually indistinguishable from other legitimate domain names. These domain names are referred to as holographs, as they are comprised of various characters from separate scripts. Homographs can be used to trick users into going to a malicious domain and as such have been used in various phishing schemes.

Whole-Script Spoofing & Single-Script Spoofing

Whole script spoofing and single script spoofing differ from the mixed-script spoofing approach. Due to the introduction of IDNs, entire domain names can be spoofed through substituting each character in a domain with one of a different script (Krammer, 2006). This attack relies on each character in a domain name having an indistinguishable counterpart in another script. Attackers can utilise this coincidence to generate a fully indistinguishable domain name. Another attack possible is single script spoofing. This attack uses characters from the same script to visually trick users into going to the domain (Gelernter & Herzberg, 2016). These attacks are more recognisable to end users, as attackers are not substituting Unicode character from different scripts (Krammer, 2006). Instead, domain names are constructed using characters which appear somewhat identical to their counterparts. An example of this is Latin 'o' and '0', which can be used to forge the domain 'www.g00gle.com'.

DEFENCE TECHNIQUES

Various defence mechanisms were adopted to address the security issues arising from the introduction of IDNs. The most prevalent defence against homograph attacks is displaying Unicode characters in a Punycode format. Punycode is used to encode a Unicode string into its appropriate ASCII string representation (Costello, 2003). Punycode values are prefixed with xn-- to represent the Unicode string. Values which contain ASCII characters

are interpreted as literal strings; however Unicode characters are transformed into their ASCII interpretation (Costello, 2003).

Web Browser Defence

As a means of defence against homograph threats, web browsers have begun implementing defence measures to notify users of potential issues with the domain name. One mitigation technique employed by Internet Explorer 7 is to display Punycode when mixed-script characters are detected in the domain name (Al Helou & Tilley, 2010). Another defensive technique which browsers incorporate is colour coding particular scripts. This technique shows characters in different colours based on the script to which they belong.

Another technique incorporated by Mozilla and Safari is to use a whitelisting approach. This security measure displays all IDNs in Punycode unless the domain is registered under a Top Level Domain (TLD) which has policies in place to prevent spoofing of the domain. As a requirement for this registering the domain with Homoglyphs, the owner must already own the western equivalent of the domain name. Mozilla Firefox still retains this whitelisting approach for handling IDNs but has since updated how the Punycode display works in 2012 (Gerv, 2017). The new algorithm employed determines if the entered domain name contains characters belonging the same script or if the characters are being pulled from one of the allowed predefined script combinations (Gerv, 2017). If the entered domain name is not within the pre-established whitelist for TLDs or if the domain name is using characters from illegal script combinations, a Punycode sample is displayed to the user. The previous whitelisting approach only remains for compatibility purposes with the domains registered with it but is no longer the primary method used for Homograph mitigation (Gerv, 2017). Some browsers such as Opera still retain the previous whitelisting approach which Firefox has since abandoned, but remains active for compatibility purposes with the registered domains.

In the Hannay and Baatard (2012) survey, various browsers were assessed to determine their effectiveness in mitigating Homograph attacks. In this previous study, numerous versions of the web browsers Mozilla Firefox, Internet Explorer, Google Chrome, Opera, and Safari were assessed regarding the mitigation techniques used against homograph attacks. The results from the previous study demonstrated that later versions certain browsers such as Google Chrome and Mozilla Firefox were highly effective in mitigating homograph attacks, while the latest versions of Internet Explorer and Safari were still vulnerable to some attacks at the time of the study.

In recent versions of browsers such as Firefox, the algorithm used to display Punycode has been altered. This modification indicates that later versions may not have the same homograph mitigation methods applied. This research updates the results seen in the Hannay and Baatard (2012) study by analysing the mitigation techniques adopted by browser versions released between 2011 and 2017. Through performing this investigation, the author answers the question of if mitigation techniques have been applied to browser versions post-2011 and if changes to the mitigation functions resulted in further vulnerabilities with certain versions.

RESEARCH METHOD

The testbed used for this investigation consists of a virtual machine running Windows 7 and an Ubuntu 14.04 Desktop virtual machine hosting sites containing single and mixed-script domain name. Various versions of the web browsers Mozilla Firefox, Chromium, Internet Explorer, and Opera were installed on the Windows 7 virtual machine. The versions which are tested consist of various versions from 2011 to 2017, which were not covered in the 2012 study. Due to issues obtaining previous versions of Google Chrome, the Chromium browser is used for this research. Given the total quantity of browser versions released per year, the versions covered in this research consist of those releases mid-year and at the end of the year. As a means of managing the browser versions installed on the tested, a snapshot is taken before any browser is installed on the environment. After each browser version is tested, the virtual environment is rolled back to the base install to prepare for the next iteration of testing.

Table 2: Test Domain Names

Domain Name Character Set	Domain Name
Single-Script	n0tasecuresite.com
Mixed-Script	notasecuresite.com

To perform this test in a controlled environment, two distinct websites were created, these were hosted on an Apache web server. The domain names of these sites were configured to use single-script and mixed-script

characters. The domain names used in the test are shown in Table 2. The single-script domain uses the Latin character '0' in place of an 'o', while the mixed-script site uses the U1086 Cryllic 'o' character. These sites are also using self-signed Secure Socket Layer (SSL) certificates which also correspond to the single-script and mixed-script characters used in the domain names. Finally, both the sites are configured to use Geolocation services which prompt the user to share their current location with the given domain. Through hosting websites using single-script and mixed-script character sets, It is possible to assess the mitigation techniques applied to the two domains securely.

To test the effectiveness of the mitigation methods applied to each browser version, four common attack vectors were identified corresponding to browser locations where the output is either the standard Unicode format or Punycode. This mitigation tactic is used to convey information to the user, regarding if any non-standard ASCII characters are detected. Through assessing various browser features, the research demonstrates how effective each browser is in conveying this information to the end user. These attack vectors used in this investigation are:

- The text shown in the browser's address bar, after the "Go" (or equivalent) button has been pressed.
- The text shown in the browser's status bar while the mouse is hovering over a hyperlink.
- The text shown when viewing prominent information about the website's SSL certificate.
- The text shown when the user is prompted to share their location using geolocation services.

As a means to assess how effective each browser version is with mitigating homograph attacks, a revised mitigation rating table similar to those presented in the 2012 study is provided. A value of zero was assigned should the browser not support a particular attack vector, for example, geolocation services. A value of negative one is given if a browser has implemented an attack vector, but is still open to homograph attacks. A value of positive one is given if a browser has implemented an attack vector and mitigated homograph attacks. Should the browser support or not support mitigation methods when displaying text in the browsers address bar, a value of positive two or negative two is given for this value. This value is due to this vector being the most prominent location for displaying Punycode. The browsers will receive two distinct ratings for single script and mixed-script mitigation. An example of the table used can be viewed in Table 3.

Table 3: Mitigation Table Structure

Address Bar	Status Bar	SSL Certificate	Location Request
-2 (No mitigation)	-1 (No mitigation)	-1 (No mitigation)	-1 (No mitigation)
0 (No Support)	0 (No Support)	0 (No Support)	0 (No Support)
+2 (Mitigated)	+1 (Mitigated)	+1 (Mitigated)	+1 (Mitigated)

RESULTS

The post-2011 versions of Internet Explorer, as shown in Table 4, demonstrated improvements to mitigating Homograph attacks. Following the introduction of geolocation services in version 9, version 10 incorporated a Punycode mitigation method for mixed-script attacks. Internet Explorer has yet to implement a method of displaying Punycode for the SSL certificate, therefore not all mitigation methods have been implemented in the browser as of yet. Internet Explorer provides no method of mitigating against single script attacks, as no notification was given when supplementing a Latin 'o' and a '0'.

As with Internet Explorer, single-script mitigation techniques are not present in any version of Firefox shown in Table 5. Version 17 of Firefox was not capable of displaying any form of Punycode in the address bar, SSL certificate, or geolocation request. This finding was likely a bug with versions from that time span, as this issue was later fixed. The only other issue with later Firefox versions was a lack of geolocation support present in version 26.

These results of Table 6 show that the Opera browser is highly effective in detecting mixed-script IDNs and implementing appropriate mitigation methods to notify the users. As with the other browsers, no support for mitigating single script spoofing has been implemented yet.

Table 4: Internet Explorer Mitigation Table

Version & Release Date	Address Bar Mitigation	Status Bar Mitigation	SSL Certificate Mitigation	Location Request Mitigation	Mitigation Rating
10 (2012 - 09) **11 (2013 - 10)**	Single Script – No Mitigation	Single Script – No Mitigation	Single Script – No Mitigation	Single Script – No Mitigation	-5
	Mixed-Script – Punycode	Mixed-Script – Punycode	Mixed-Script – No Mitigation	Mixed-Script – Punycode	+3

Table 5: Firefox Mitigation Table

Version & Release Date	Address Bar Mitigation	Status Bar Mitigation	SSL Certificate Mitigation	Location Request Mitigation	Mitigation Rating
13 (2012 – 06)	Single Script – No Mitigation	Single Script – No Mitigation	Single Script – No Mitigation	Single Script – No Mitigation	-5
	Mixed-script – Punycode	Mixed-script – Punycode	Mixed-script – Punycode	Mixed-script – Punycode	+5
17 (2012 – 11)	Single Script – No Mitigation	Single Script – No Mitigation	Single Script – No Mitigation	Single Script – No Mitigation	-5
	Mixed-script – No Mitigation	Mixed-script – Punycode	Mixed-script – No Mitigation	Mixed-script – No Mitigation	-3
22 (2013 – 06)	Single Script – No Mitigation	Single Script – No Mitigation	Single Script – No Mitigation	Single Script – No Mitigation	-5
	Mixed-Script – Punycode	Mixed-Script – Punycode	Mixed-Script – Punycode	Mixed-Script – Punycode	+5
26 (2013 – 12)	Single Script – No Mitigation	Single Script – No Mitigation	Single Script – No Mitigation	**No support***	-4
	Mixed-script – Punycode	Mixed-script – Punycode	Mixed-script – Punycode	**No Support***	+4
30 (2014 – 06) **34 (2014 – 12)** **38 (2015 – 05)** **43 (2015 – 12)** **47 (2016 – 06)** **50 (2016 – 11)** **55 (2017 – 08)**	Single Script – No Mitigation	Single Script – No Mitigation	Single Script – No Mitigation	Single Script – No Mitigation	-5
	Mixed-Script – Punycode	Mixed-Script – Punycode	Mixed-Script – Punycode	Mixed-Script – Punycode	+5

Table 6: Opera Mitigation Table

Version & Release Date	Address Bar Mitigation	Status Bar Mitigation	SSL Certificate Mitigation	Location Request Mitigation	Mitigation Rating
12 (2012 – 06)	Single Script – No Mitigation	Single Script – No Mitigation	Single Script – No Mitigation	Single Script – No Mitigation	-5
	Mixed-script – Punycode	Mixed-script – Punycode	Mixed-script – Punycode	Mixed-script – Punycode	+5
15 (2013 – 07)	Single Script – No Mitigation Mixed-script – Punycode	Single Script – No Mitigation Mixed-script – Punycode	Single Script – No Mitigation Mixed-script – Punycode	No Support	-4
	Mixed-script – Punycode	Mixed-script – Punycode	Mixed-script – Punycode	No Support	+4
18 (2013 – 11) 22 (2014 – 06) 26 (2014 – 12) 30 (2015 – 06)	Single Script – No Mitigation	Single Script – No Mitigation	Single Script – No Mitigation	Single Script – No Mitigation	-5
34 (2015 – 12) 38 (2016 – 06) 42 (2016 – 12) 47 (2017 – 08)	Mixed-script – Punycode	Mixed-script – Punycode	Mixed-script – Punycode	Mixed-script – Punycode	+5

Table 7: Chromium Mitigation Table

Version & Release Date	Address Bar Mitigation	Status Bar Mitigation	SSL Certificate Mitigation	Location Request Mitigation	Mitigation Rating
20.0.1123 (2012 – 05) 25.0.1323.1 (2012 – 11) 29.0.1541.0 (2013 – 06) 32.0.1700.6 (2013 – 11) 37.0.2017.2 (2014 – 05) 41.0.2243.0 (2014 – 12) 45.0.2431.0 (2015 – 06) 49.0.2593.0 (2015 – 12) 53.0.2763.0 (2016 – 05)	Single Script – No Mitigation	Single Script – No Mitigation	Single Script – No Mitigation	Single Script – No Mitigation	-5
	Mixed-script – Punycode	Mixed-script – Punycode	Mixed-script – Punycode	Mixed-script – Punycode	+5
56.0.2902.0 (2016 – 10)	Single Script – No Mitigation	Single Script – No Mitigation	Single Script – No Support	Single Script – No Mitigation	-4
	Mixed-script – Punycode	Mixed-script – Punycode	Mixed-script – No Support	Mixed-script – Punycode	+4
61.0.3153.0 (2017 – 09)	Single Script – No Mitigation	Single Script – No Mitigation	Single Script – No Support	Single Script – No Mitigation	-4
	Mixed-script – Punycode	Mixed-script – Punycode	Mixed-script – No Support	Mixed-script – Punycode	+4

The results from Table 7 show that Chromium is highly effective in mitigating against IDN homograph attacks. Like the other browsers analysed, no support for single-script mitigation has been added to any version. Chromium was shown to be highly consistent with the mitigation functions applied to each browser version. However, the most recent version of the browser does not accurately convey SSL certificate information to the user. This limitation results in the browser being unable to display Punycode for the domain name used in the SSL certificate.

The results from single-script mitigation are demonstrated in Figure 1. Given the lack of mitigation functions, each browser received a negative mitigation rating. The most consistent rating across all browsers was a -5. The versions which achieved a -4 rating were as a result of a lack of support for a given for that version.

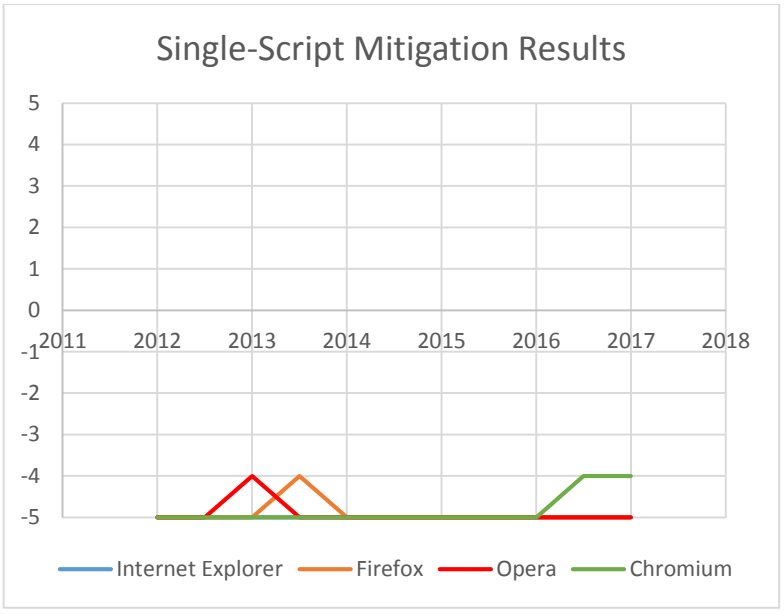


Figure 2 - Single-Script Mitigation Results

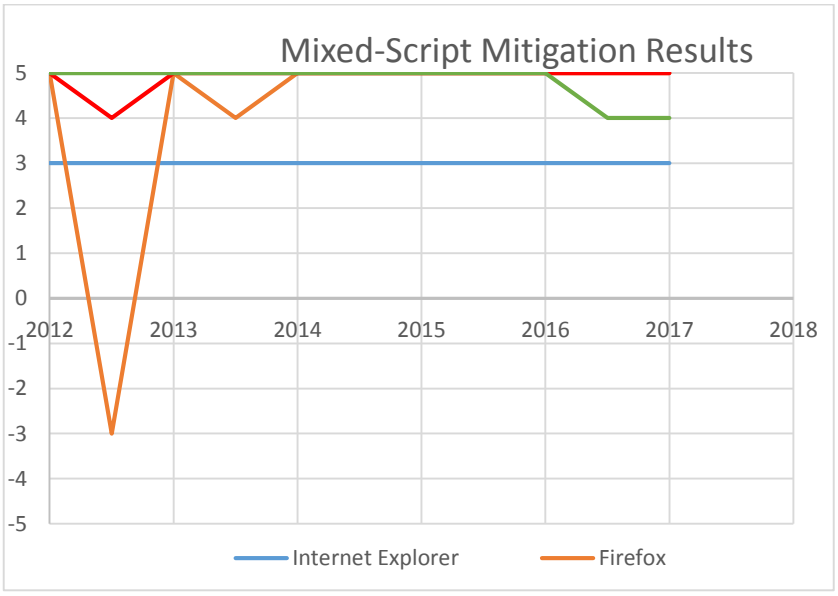


Figure 3- Mixed-Script Mitigation Results

The results from mixed-script mitigation are demonstrated in Figure 2. The majority of browser versions demonstrate consistent mitigation ratings across different releases. The drops in mitigation ratings appear to be from certain versions not supporting for a particular feature. However, the 2012 version of Firefox is shown to have the greatest drop in mitigation rating. This finding is a result of the version only displaying Punycode for the

hover feature of Hyperlinks. The issues discovered in this version were later amended in later versions, with the only other drop being the lack of geolocation support in the 2013 version. Internet Explorer has demonstrated a greater mitigation rating in later versions by implementing geolocation services. Opera was shown to be the most effective in implementing mitigation features for IDN homograph attacks, as it received the rating of +5 more consistently than the other browsers tested.

The uptake rate of new browser versions plays a significant role in determining the potential exposure of any particular browser version to hostile actors. A study conducted by Ion, Reeder, and Consolvo (2015) compared the cyber security practices of expert and non-expert users. One major finding of the study was that largest point of difference between the groups, was the importance placed on installing software updates in a timely fashion (Ion et al., 2015). Examining W3Counter data on browser usage shows us that as of August 2017, that 25.84% of users were running web browsers more than three months old (W3Counter, 2017). As such we can see that the potential exposure for vulnerabilities in specific browser versions, such as those seen in Firefox in 2012, may span many months from release of the software.

CONCLUSION

The results discovered in this research are representative of how IDN mitigation techniques have been implemented in various browsers in recent years. In extending the results of the previous study by Hannay and Baatard (2012), the results found in this study appear to indicate that mixed-script homograph attack mitigation has become a standard feature for most browsers. The majority of browsers analysed in this study demonstrated to implement appropriate mitigation techniques for IDN homograph attacks. The only degree of variance in the mixed-script results appears to be with different versions of particular browsers, which could be a result of bugs or changes in the algorithm. The results for single-script mitigation demonstrate that this is not a feature commonly implemented in browsers. Of the browsers analysed, none provided any support to notify the end user about a non-standard character being used in the domain name. This research suggests that mitigation against mixed-script homograph attacks has become a common feature for most browsers, while implementation of mitigation functionality for single-script spoofing attacks has not been undertaken.

REFERENCES

Al Helou, J., & Tilley, S. (2010). *Multilingual web sites: Internationalized Domain Name homograph attacks.* Paper presented at the Web Systems Evolution (WSE), 2010 12th IEEE International Symposium on.

Costello, A. M. (2003). Punycode: A bootstring encoding of unicode for internationalized domain names in applications (IDNA).

Davis, M., & Suignard, M. (2006). Unicode security considerations: Citeseer.

Gelernter, N., & Herzberg, A. (2016). Autocomplete Injection Attack. In I. Askoxylakis, S. Ioannidis, S. Katsikas, & C. Meadows (Eds.), *Computer Security – ESORICS 2016: 21st European Symposium on Research in Computer Security, Heraklion, Greece, September 26-30, 2016, Proceedings, Part II* (pp. 512-530). Cham: Springer International Publishing.

Gerv. (2017, April 17th). IDN Display Algorithm. Retrieved from https://wiki.mozilla.org/IDN_Display_Algorithm

Hannay, P. (2012). Google Awesome Edition. Retrieved from http://xn--goole-tmc.com/

Hannay, P., & Baatard, G. (2012). *The 2011 IDN homograph attack mitigation survey.* Paper presented at the Proceedings of the International Conference on Security and Management (SAM).

Hannay, P., & Bolan, C. (2009). *Assessment of Internationalised Domain Name Homograph Attack Mitigation.* Paper presented at the Australian Information Security Management Conference.

Introduction to IDNs. (2016, July 3rd). Retrieved from http://idnworldreport.eu/introduction-to-idns/

Ion, I., Reeder, R., & Consolvo, S. (2015). *"... No one Can Hack My Mind": Comparing Expert and Non-Expert Security Practices.* Paper presented at the SOUPS.

Kaspersky Lab. (2017). What is Beta Bot? Retrieved from https://usa.kaspersky.com/resource-center/definitions/beta-bot

Key Numbers. (2016, August 16th). Retrieved from http://idnworldreport.eu/facts-figures/number-of-idns-2/

Krammer, V. (2006). *Phishing defense against IDN address spoofing attacks.* Paper presented at the Proceedings of the 2006 International Conference on Privacy, Security and Trust: Bridge the Gap Between PST Technologies and Business Services.

Kumar, M. (2017, April 17th). This Phishing Attack is Almost Impossible to Detect On Chrome, Firefox and Opera. Retrieved. Retrieved from http://thehackernews.com/2017/04/unicode-Punycode-phishing-attack.html

Maunder, M. (2017, April 14th). Chrome and Firefox Phishing Attack Uses Domains Identical to Known Safe Sites. Retrieved from https://www.xn--e1awd7f.com/ & https://www.wordfence.com/blog/2017/04/chrome-firefox-unicode-phishing/

Mimoso, M. (2017, September 6th). IDN Homograph Attack Spreading Betabot Backdoor. Retrieved from https://threatpost.com/idn-homograph-attack-spreading-betabot-backdoor/127839/

Mockapetris, P. V. (1987). Domain Names-Concepts and Facilities *RFC1034*: IETF.

Spaulding, J., Upadhyaya, S., & Mohaisen, A. (2016). *The landscape of domain name typosquatting: Techniques and countermeasures.* Paper presented at the Availability, Reliability and Security (ARES), 2016 11th International Conference on.

Tseng, S.-S., Ku, C.-H., Lu, A.-C., Wang, Y.-J., & Geng, G.-G. (2013). *Building a Self-Organizing Phishing Model Based upon Dynamic EMCUD.* Paper presented at the Intelligent Information Hiding and Multimedia Signal Processing, 2013 Ninth International Conference on.

W3Counter. (2017). Browser & Platform Market Share - August 2017. Retrieved from https://www.w3counter.com/globalstats.php?year=2017&month=8

Zheng, X. (2017). IDN Homograph Example. Retrieved from http://xn--80ak6aa92e.com/

A REVIEW OF DATA BREACHES AND LOSSES THAT OCCURRED FROM LAPTOPS THAT WERE STOLEN OR OTHERWISE MISPLACED IN 2015 AND 2016

Samuel Griffith Wakeling[2], Peter Hannay[1,2], Zubair Baig[1,2]
[1]Security Research Institute, [2]School of Science. Edith Cowan University, Perth, Western Australia
sgwakeli@our.ecu.edu.au, p.hannay@ecu.edu.au, z.baig@ecu.edu.au

Abstract

This paper provides an analysis of what information can be found on laptops that may or may not have connections to an organisation of some form, the statistics of the number of laptops stolen or otherwise misplaced in 2015 and 2016, and the number of potentially affected people from each of the cases. As seen in many news articles, laptops are often stolen or otherwise misplaced by employees or contractors in an organisational environment. As discovered in this research, many laptops are stolen from vehicles or homes of employees rather than organisation's buildings, but not all. The majority of stolen or otherwise misplaced laptops have very little information security on them, and this increases the risk of a data breach once a third party has physical access to the device. The research finds that, with available information, only one laptop had used encryption for the personal, private and confidential information that was stored on the internal storage device. In total, this paper finds that 33 laptops were stolen or otherwise misplaced in 2015 and 2016. The healthcare industry had the largest number of potentially affected people, with 5,352,792 people, and an average of 334,350 across the 16 laptops. The government sector had the second highest impact, with a total of 1,000,865 potentially affected people. Out of the 33 laptops, the total number of potentially affected people was 6,598,995 affected people, with an average of 83,702 potentially affected people with each of the 33 laptops.

Keywords: Laptop; Stolen; Misplaced; Data Breach; 2015; 2016

BACKGROUND

What information can be stored on laptops?

A computer device can contain a large amount of information, which with the correct software and hardware tools can be forensically analysed to find information of interest. This information can include:

- Temporary files:
  - Application data.
  - Web browser data.
- Active directory (AD) user account on organisational devices with an active directory server.
- Offline files from network drives, including drives configured with active directory.
- Network configurations:
  - Wireless credentials.
  - IP address configurations.
- User credentials
  - Organisational systems
  - Banking system
  - Social media
  - Session data
- Intellectual property
- Other personal information
  - Black mail potential

Temporary files can include data that is stored on a local system by an application. For example, a database application may store information that was retrieved or worked on from a database in the temporary files on the system. By analysing this data, it is possible that personal and private information may be available on the system in the temporarily stored files from the database system.

Web browser data can include session data for emails, databases or other systems that require authentication. This information, which is stored in cookies, could allow an attacker to bypass the authentication on a system by restoring a previously open session before the server times out the session. Research finds that some versions of Firefox are known to store cookies and other session data in a SQLite database (Pereira, 2009), which can easily be recovered with the correct tools.

An active directory (AD) system in an organisation copies a user's account files to the system that is used to log onto the domain (Microsoft, 2016). This account data can include any private and confidential documents that a user may have stored in their personal account. By analysing this data, an attacker could potentially gain access to confidential information from an organisation and gain access to temporary internet files for internal web pages in the organisation. The attacker could then view these pages to find any vulnerabilities in the services, and view the data from the active directory user account to see if there is any confidential information stored by the user in their account or their temporary files folders.

A computer device may be configured to store files offline from network drives or other shared folders (Microsoft, 2017a) in an organisation, either with or without an active directory server. This data can be analysed for confidential information that is stored on the organisation's internal servers that has been copied to the laptop's internal storage device.

A computer device may also contain network configurations for corporate networks, which attackers could use to determine how to continue gaining further access to the information systems at their target organisation. This network information can include wireless credentials that can allow attackers to connect other devices to the organisation's network and gain further access to information, as Windows allows users to view the wireless pre-shared key in plain text (Microsoft, 2017c) when they have physical access to the device.

Issues involved in the theft of laptops

All the information described above shows how much information could potentially be stored on a single device. If these devices are stolen or otherwise misplaced, a third-party individual or group could have access to a large amount of information. This could potentially allow them to leak more information than that found on the stolen device, as they could use configuration data to gain further access to the information systems on personal or organizational networks and gain access to more data.

With access to this device, the individual or group in possession of the device could leak any information they find on the device, and with the amount of information that could be stored on a single device, this has the potential to destroy the security reputation of an organization.

Furthermore, an attacker could install malicious code to a device, then attempt to anonymously return the device to the organization to attempt to back-door their way into the organization. If a re-obtained device is not thoroughly examined for additional code or applications running on the device, it could cause potential leaks of more confidential information from the organization.
Following the loss or integrity breach of this information and data, there are a number of potential outcomes, including:

- Data loss
    - Data being deleted and no longer being accessible or useable to the user.
- Data misuse
    - Data being used for purposes other than the intended use.
        - Fraud
        - Unlawful or unethical releases of private information
        - Black mail
- Data damage
    - Data being either corrupted or encrypted making it inaccessible while it still exists on the storage device.
- Reputation loss
    - Loss of data integrity reputation to the organisation or individual that was the subject of the data breach.
- Commercial advantage to a third party engaging in information warfare
    - The potential for another organisation to possess the research and other intellectual property created and stored by the user of the laptop.

RESEARCH METHOD

The statistics for this paper were gathered by using an advanced search in Google News specifying news articles only, dated from 01/01/2015 to 31/12/2015 with the keywords "laptop data breach". Google then returns news articles from laptops that were reported stolen or misplaced in 2015 with the potential information that may have been stored on them and how much data could have been breached. The search can also be limited to "Australia Only" to get news results for data breaches in Australia instead of worldwide. However, this does not yield enough results for this paper, and so this search limitation can be removed for further results.

These same search queries were run, but the dates were changed to starting at 01/01/2016 and ending at 31/12/2016 to obtain the statistics for laptops that were stolen or otherwise misplaced in 2016.

Only news articles that specify that a data breach occurred due to a laptop being stolen or otherwise misplaced are in the scope of this paper, other forms of data breaches, including network attacks, are not in the scope of this paper and will be ignored in the search results of news reports or articles.

To get the information to be compared, each article was examined, and the industry section, information about how many people are affected by the data breach, and if the device was secured and how will be entered into a table that will be used to compare the statistics from all newspaper articles found, with references.
From this gathered information, comparisons and discussions can be made about:

- The total number of data breaches from stolen or otherwise misplaced laptops in 2015.
- The total number of data breaches from stolen or otherwise misplaced laptops in 2016.
- The total number of people affected by the breaches in both 2015 and 2016.
- The average number of people per breach.
- The security in place on each of the stolen devices, and the security of the data on each of the devices.
- The average security setup for the laptops.

If any of the above information is unavailable, average calculations can be made for the total number of available statistics for each data type, or it will be assumed that zero people were affected by the unspecified breaches.

For the general information in this paper, other papers and technology webpages will be reviewed and cited for their information regarding the potential data that can be stored on laptops, based on their possible configurations and uses within or external to an organization.

DISCUSSION

What information can be retrieved from a stolen laptop and how?

When an attacker or another third party has physical access to a storage device from a laptop computer, they can easily gain access to the data that is stored on the disk. Utilizing many freely available forensic acquisition tools and utilizing the correct hardware, attackers or other unauthorized individuals or groups can view all of the data that is stored on the laptop in their possession, including data that has been deleted but the sectors have not yet been over-written. One method often used to retrieve this data is called file carving (Gladyshev & James, 2017).

Active Directory and other locally stored files

For example, if a laptop is connected to an Active Directory server, it will contain all of the information that the user has on any other system connected to that network. Assuming the device does not have encryption setup for the entire storage device, all of the information that has been downloaded from the Active Directory server will be visible in plain text. This can allow analyzers to easily scan for particular file types of interest or temporary database files. If an analyzer finds that files could have been deleted from the internal storage device on the laptop, with the correct hardware and software, the analyzer would be "able to recover the deleted files" (Hanson, 2005).
A laptop may also contain offline files from a network drive at an organization. The type of data that would be stored here would depend on the information that is stored on the network drive in the organization. However, this data can range from personal information to freely available data such as application installers.

A laptop may also have credentials stored on it, such as wireless network credentials or usernames and passwords for internal or external websites and systems in the organization. Even if these are not stored in plain text, they can be dictionary or brute-force attacked by a powerful system or rainbow table to find the original passwords. This task can be easily performed on another system once the storage device from the stolen or misplaced laptop is obtained, and there are a number of open source, free to use tools that can achieve this (Eston, 2010).

Web Browsers

Modern web browsers store a large amount of temporary data on a system. If a laptop is acquired, there may be open session data stored in the browser's temporary files and cookies, even if the user has attempted to delete them through their web browser (Pereira, 2009). This issue can allow an attacker to bypass the authentication systems on a web-based service and use the currently authenticated session on the device. With this session, an attacker could access more information stored on the network than just the information that was stored on the obtained device, and the attacker could potentially use this network access to launch more internal network attacks on the storage systems at the organization.

Many users also choose to allow their web browsers to save their login information to many web pages and services. This information can be harvested in order to obtain and dictionary or brute force attack the credentials utilizing tools such as John The Ripper (Taiabul Haque, Wright, & Scielzo, 2014), to re-use them to obtain more information from the network.

Web history data from web browsers can be used to determine what research was conducted on the laptop, or, what sites may have stored temporary information on the laptop. As temporary files are not often deleted, even if the user requests it (Pereira, 2009), there may be information of interest that can be obtained from recently viewed web pages that are in the browsing history of the device. Also, forensic tools can be used to view any temporary information that may have been deleted on the device.

User Directory

As discussed above, Active Directory servers will send a connected laptop a large amount of data or information from the storage systems from the network (Microsoft, 2016). However, if the system is not part of an Active Directory server, there may still be a large amount of information that is stored on the system while the user was working either on or off-site with data and information.

While the device is either at an organization or connected to the network with a Virtual Private Network (VPN), any documents or other information/data that are accessed on the remote storage servers may be temporarily stored on the local storage device in the laptop. This data can be harvested for any temporary data that might contain confidential information.

If an organization has a web-based document management system, such as Microsoft's SharePoint, there may be temporary downloaded documents stored in the user directory folder on the laptop's storage device. These files may contain completed forms or other records potentially containing confidential data or information.

Temporary internet files that are stored on the system from accessing the websites on the internal networks could be analyzed to find and potentially exploit any vulnerabilities that are present on the web server(s).

System Configurations

System configurations could be analyzed to determine the network structure of the organization and how it is configured. This data could be helpful for further network attacks that could be used to obtain more information than the information that is present on the laptop's internal storage device.

If the system is configured to access the internet through a proxy server, these credentials and connection information can be used to determine where the proxy server is on the network, and how to connect to it, which could be used by an attacker to redirect traffic to a different location. If an automatic configuration script is available on the system, it can contain this information (Microsoft, 2017d).

Wireless network credentials can also be stored on a laptop, and these can be retrieved from the laptop's internal storage device (Microsoft, 2017c). These credentials can be used to easily gain access to any wireless networks that the laptop has been connected to, including home networks and networks at organizations.

Laptops that have been connected to an organization may have network drives configured in the operating system (Microsoft, 2017b). The configuration information for these network drives can be used to determine the network and storage structure of the organization, as the network paths show up in the file browser. Also, some of these network drives may have been configured to store some information and data offline, and this information and data will be stored on the laptop's internal storage device.

STATISTICS OF STOLEN OR OTHERWISE MISPLACED LAPTOPS IN 2015

Table 1: of all cases in 2015

Industry	Number of potentially affected people	Was the device password protected?	Was the device encrypted?	
Background Screening	100,000	Password protected	Unencrypted	(Greenberg, 2015c)
Healthcare	39,090			(Snell, 2015a)
Education	9,300	Password protected	Not encrypted	(Leventhal, 2015)
Healthcare	8,000			(Muckenfuss, 2015)
Healthcare	3,000		Not encrypted	(Lewis, 2015)
Healthcare	2,800		Encrypted	(WISN, 2015)
Healthcare	1,359	Password protected	Not encrypted	(Cantu, 2015)
Education	941	Password protected	Not encrypted	(Gallagher, 2015)
Financial		Password protected	Not encrypted	(Ilascu, 2015)
Healthcare				(Snell, 2015b)
Healthcare			"All of the personal information contained on the laptop was deleted"	(Greenberg, 2015d)
Law				(Greenberg, 2015b)
Military			Encrypted	(Shropshire-Star, 2015)
Security			Encrypted	(Greenberg, 2015a)

Table 2: Data analysis of cases in 2015

Industry Sector	Total number of breaches	Average number of potentially affected people	Total number of potentially affected people
Background Screening	1	100,000	100,000
Healthcare	7	7,750	54,249
Education	2	5,121	10,241
Financial	1	0	0
Law	1	0	0
Military	1	0	0
Security	1	0	0
Total	14	16,124	164,490

As some of the articles or sources did not specify the number of potentially affected people, the calculations of total numbers in this table assumes that 0 people were affected on the cases with no data on the number of affected people for calculating the averages.

STATISTICS OF STOLEN OR OTHERWISE MISPLACED LAPTOPS IN 2016

Table 3: Total of all cases in 2016

Industry	Number of potentially affected people	Was the device password protected?	Was the device encrypted?	
Healthcare	5,000,000			(Barth, 2016)
Government	>1,000,000^			(Press Herald, 2016)

Military	130,000			(Bevan, 2016)
Healthcare	52,076	The device's networking and employee's credentials were disabled.		(Belliveau, 2016b)
Healthcare	<25,000#	Password protected		(Devlin, 2016)
Healthcare	12,000	Password protected		(Martin, 2016)
Education	4,022			(Monegain, 2016)
Healthcare	3,119			(Snell, 2016)
Sport	>1,000*			(Clarke, 2016)
Healthcare	600		"Patient information on the hard drive was encrypted"	(Belliveau, 2016a)
Aged care	75		The device was unencrypted	(Belfast Telegraph, 2016)
Education		Password protected.		(Whitbourn, 2016)
Healthcare		Password protected	Not encrypted	(Masters, 2016b)
Healthcare			Unencrypted	(Olenick, 2016b)
Prison		Password protected	Not encrypted	(Jayanthi, 2016)
Security		Password protected		(Olenick, 2016a)
Vacation Properties				(Masters, 2016a)

* Article only specified "thousands of players".
^ Article only specified "millions".
# Article only specified "up to 25,000".

Table 4: Data analysis of cases in 2016

Industry Sector	Total number of breaches	Average number of potentially affected people	Total number of potentially affected people
Healthcare	9	588,727	5,298,543
Government	2	500,432	1,000,865
Military	1	130,000	130,000
Education	2	2,011	4,022
Sport	1	1,000	1,000
Aged care	1	75	75
Prison	1		
Security	1		
Vacation Properties	1		
Total	19	135,805	6,434,505

As some of the articles or sources did not specify the number of potentially affected people, the calculations of total numbers in this table assumes that 0 people were affected on the cases with no data on the number of affected people for calculating the averages.

DATA ANALYSIS OF ALL STOLEN OR OTHERWISE MISPLACED LAPTOPS IN 2015 AND 2016

Table 5

Industry Sector	Total number of breaches	Average number of potentially affected people	Total number of potentially affected people
Healthcare	16	334,350	5,352,792
Government	2	500,432	1,000,865
Military	2	65,000	130,000
Background Screening	1	100,000	100,000

Education	4	3,566	14,263
Sport	1	1,000	1,000
Aged care	1	75	75
Security	2	0	0
Financial	1	0	0
Law	1	0	0
Prison	1	0	0
Vacation Properties	1	0	0
Total	33	83,702	6,598,995

As some of the articles or sources did not specify the number of potentially affected people, the calculations of total numbers in this table assumes that 0 people were affected on the cases with no data on the number of affected people for calculating the averages.

SUMMARY OF STATISTICS

An analysis of 33 different cases of laptops being stolen or otherwise misplaced found that a total of 16 of these were from the healthcare industry. The largest total number of potentially affected people with 5,000,000 potentially affected people in one data breach was from the healthcare industry. In these cases, the system was either protected with just a password or no security information was available. No encryption was used on the majority of the laptops in the cases.

Figure 1 shows the percentage of devices stolen from each sector. As shown in the diagram, 49% of cases occurred in the healthcare industry. The second largest industry was the education industry, with 12% of the cases.

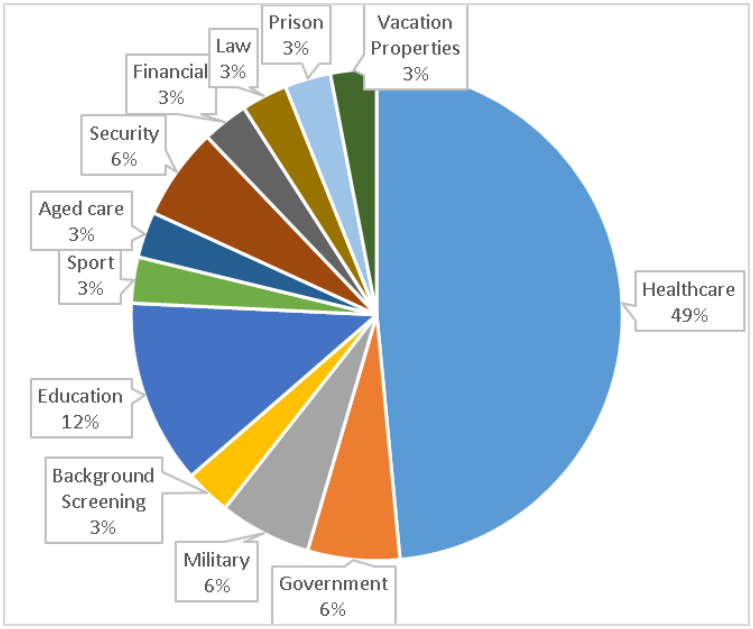


Figure 1: Number of cases

Figure 2 shows the percentage of people potentially affected in total by each industry sector. Healthcare had a total of 81% of the total number of potentially affected people, with the government sector following at 15% of the total number of potentially affected people.

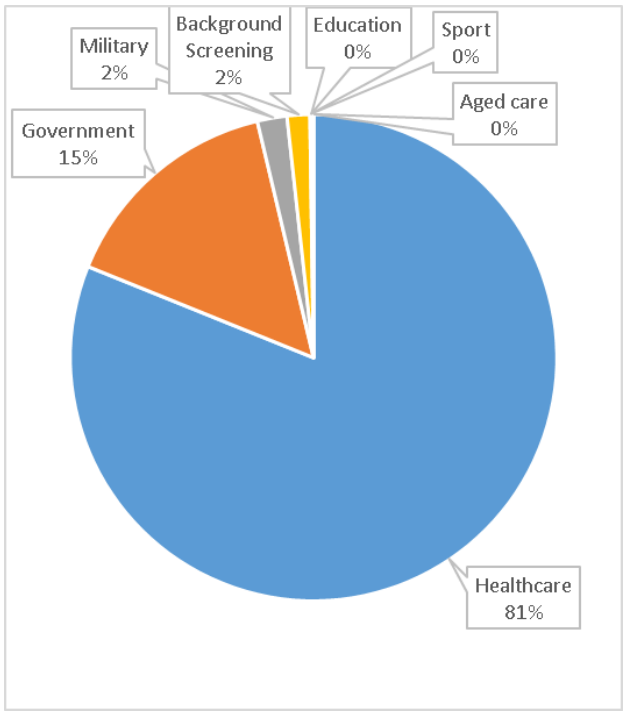


Figure 2: Number of potentially affected people

An analysis of available security information finds that there was no encryption used on the majority of laptops, and that they were simply protected with passwords in their operating systems. This means that all the information discussed in previous sections of this paper would be accessible in plain text, and an attacker can simply connect the storage device to another system or boot a live operating system on the stolen or misplaced laptop to gain access to the information on the system, allowing them to breach the integrity of the data of the potential number of affected people for each stolen or misplaced system.

CONCLUSION

In conclusion, this paper finds that very little security is used on laptops that contain personal information of employees or customers of businesses or other organisations. In some cases, the laptops belonged to the employees, rather than belonging to an organisation.

In total, the statistics of this paper finds that across the total of 33 analysed cases, an average of 83,702 people were potentially affected per stolen or otherwise misplaced laptop, with the total number of people potentially affected by stolen or otherwise misplaced laptops was 6,598,995 people.

The information in this paper can be used as a reference point for determining the security protocols that should be used on mobile devices, including but not limited to, laptops, phones, tablets and more. This paper's information is an example of the potential losses and breaches of integrity from a single stolen or otherwise misplaced unencrypted device from either an organisation or an employee's personal possessions.

Further research should be conducted yearly to monitor the statistics of stolen laptops, to check if data breaches from stolen or otherwise misplaced laptops continues to increase over time. If the statistics and information in this paper are used effectively in education and training, further research in this area should show a decline in the number of data breaches and potentially affected users from stolen or otherwise misplaced laptops in the future.

Further research over the past ten years should also be conducted to provide an accurate representation of the changes to the number of data breaches and people affected by stolen or otherwise misplaced laptops in the past, compared to the amount in 2015-2016 and the future. The data collected by all the research can be graphed and tabled for statistical and educational purposes with the intent to reduce the number of data breaches that occur from stolen or otherwise misplaced laptops in the future.

REFERENCES

Barth, Bradley. (2016). Personal laptop, possibly containing data on 5M patients, stolen from HHS facility. Retrieved 26-05, 2017, from https://www.scmagazine.com/personal-laptop-possibly-containing-data-on-5m-patients-stolen-from-hhs-facility/article/528761/

Belfast Telegraph. (2016). Nursing home fined for data breach after laptop with patients' details stolen. Retrieved 21-05, 2017, from http://www.belfasttelegraph.co.uk/news/northern-ireland/nursing-home-fined-for-data-breach-after-laptop-with-patients-details-stolen-34994692.html

Belliveau, Jacqueline. (2016a). Robbery at CA Practice Causes Possible Healthcare Data Breach. Retrieved 28-05, 2017, from http://healthitsecurity.com/news/robbery-at-ca-practice-causes-possible-healthcare-data-breach

Belliveau, Jacqueline. (2016b). Stolen Laptop Leads to Possible Healthcare Data Breach in KS. Retrieved 21-05, 2017, from http://healthitsecurity.com/news/stolen-laptop-leads-to-possible-healthcare-data-breach-in-ks

Bevan, Kate. (2016). 'Compromised' laptop implicated in US Navy breach of 130,000 records. Retrieved 21-05, 2017, from https://nakedsecurity.sophos.com/2016/11/24/compromised-laptop-implicated-in-us-navy-breach-of-130000-records/

Cantu, Tony. (2015). HealthSouth Rehab Hospital Warns of Potential Data Breach. Retrieved 19-09, 2017, from https://patch.com/texas/round-rock/healthsouth-rehab-hospital-warns-potential-data-breach-0

Clarke, Liz. (2016). Redskins employee's laptop stolen; NFL trying to determine extent of the breach. Retrieved 21-05, 2017, from https://www.washingtonpost.com/sports/redskins/redskins-employees-laptop-stolen-but-medical-records-not-feared-compromised/2016/06/01/3605b86e-2833-11e6-a3c4-0724e8e24f3f_story.html?utm_term=.25855e16c142

Devlin, Vince. (2016). Up to 25,000 could be affected by laptop stolen from New West employee. Retrieved 27-05, 2017, from http://missoulian.com/news/local/up-to-could-be-affected-by-laptop-stolen-from-new/article_88f2c565-a9dc-56f7-9620-25f4eb663d9b.html

Eston, Tom. (2010). Easy-To-Find Brute-Force Tools. InformationWeek(1284), 66.

Gallagher, Noel K. (2015). UMaine professor whose laptop was stolen violated university's data policy. Retrieved 25-08, 2017, from http://www.pressherald.com/2015/02/20/professor-whose-laptop-was-stolen-violated-university-systems-data-policy/

Gladyshev, Pavel, & James, Joshua I. (2017). Decision-theoretic file carving. Digital Investigation, 22(Supplement C), 46-61. doi: https://doi.org/10.1016/j.diin.2017.08.001

Greenberg, Adam. (2015a). Employee data on stolen Schlage Lock Company laptop. Retrieved 25-08, 2017, from https://www.scmagazine.com/employee-data-on-stolen-schlage-lock-company-laptop/article/532938/

Greenberg, Adam. (2015b). Personal data on laptop stolen from attorney with California law firm. Retrieved 25-08, 2015, from https://www.scmagazine.com/personal-data-on-laptop-stolen-from-attorney-with-california-law-firm/article/532899/

Greenberg, Adam. (2015c). SterlingBackcheck laptop stolen, contained data on about 100K individuals. Retrieved 25-08, 2017, from https://www.scmagazine.com/sterlingbackcheck-laptop-stolen-contained-data-on-about-100k-individuals/article/532911/

Greenberg, Adam. (2015d). Stolen DJO Global laptop contained patient data. Retrieved 25-08, 2017, from https://www.scmagazine.com/stolen-djo-global-laptop-contained-patient-data/article/536647/

Hanson, Doug. (2005). Computer forensic analysis. Law Enforcement Technology, 32(4), 8,10,12,14-16.

Ilascu, Ionut. (2015). Financial Institution Piedmont Advantage Loses Laptop with Customer Info. Retrieved 20-09, 2017, from http://news.softpedia.com/news/Financial-Institution-Piedmont-Advantage-Loses-Laptop-with-Customer-Info-474666.shtml

Jayanthi, Akanksha. (2016). Stolen laptop compromises PHI of California inmates. Retrieved 27-05, 2017, from http://www.beckershospitalreview.com/healthcare-information-technology/stolen-laptop-compromises-phi-of-california-inmates.html

Leventhal, Rajiv. (2015). University of Oklahoma Acknowledges Data Breach from Stolen Laptop. Retrieved 19-09, 2017, from https://www.healthcare-informatics.com/news-item/university-oklahoma-acknowledges-data-breach-stolen-laptop

Lewis, Dave. (2015). US Healthworks Suffers Data Breach Via Unencrypted Laptop. Retrieved 25-08, 2017, from https://www.forbes.com/sites/davelewis/2015/06/01/us-healthworks-suffers-data-breach-via-unencrypted-laptop/#2d9d57fb10c6

Martin, Kate. (2016). Stolen laptop tied to more than 12,000 accounts; CHI Franciscan says no evidence they were accessed. Retrieved 28-05, 2017, from http://www.thenewstribune.com/news/business/article122313219.html

Masters, Greg. (2016a). Laptop stolen from home of Welk Resorts employee, breach letters go out. Retrieved 27-05, 2017, from https://www.scmagazine.com/laptop-stolen-from-home-of-welk-resorts-employee-breach-letters-go-out/article/571096/

Masters, Greg. (2016b). Stolen laptop puts data of CVS customers in Alabama at risk. Retrieved 27-05, 2017, from https://www.scmagazine.com/stolen-laptop-puts-data-of-cvs-customers-in-alabama-at-risk/article/529115/

Microsoft. (2016). Folder Redirection, Offline Files, and Roaming User Profiles overview. Retrieved 20-09, 2017, from https://technet.microsoft.com/en-us/library/hh848267(v=ws.11).aspx

Microsoft. (2017a). Configure Offline Availability for a Shared Folder. Retrieved 10-10, 2017, from https://technet.microsoft.com/en-us/library/cc755136(v=ws.11).aspx

Microsoft. (2017b). Drive Map. Retrieved 11-10, 2017, from https://technet.microsoft.com/en-us/library/cc755136(v=ws.11).aspx

Microsoft. (2017c). How to find your wireless network password. Retrieved 29-09, 2017, from https://support.microsoft.com/en-us/help/4023501/surface-how-to-find-your-wireless-network-password

Microsoft. (2017d). Using Automatic Configuration, Automatic Proxy, and Automatic Detection. Retrieved 11-10, 2017, from https://technet.microsoft.com/en-au/library/cc985352.aspx

Monegain, Bernie. (2016). OHSU pays $2.7 million fine to HHS Office for Civil Rights for two HIPAA breaches. Retrieved 27-05, 2017, from http://www.healthcareitnews.com/news/ohsu-pays-27-million-fine-hhs-office-civil-rights-two-hipaa-breaches

Muckenfuss, Mark. (2015). UC RIVERSIDE: Computer stolen; data breach affects 8,000. Retrieved 25-08, 2015, from http://www.pe.com/2015/04/07/uc-riverside-computer-stolen-data-breach-affects-8000/

Olenick, Doug. (2016a). M. Holdings Security issues warning on possible data breach. Retrieved 27-05, 2017, from https://www.scmagazine.com/m-holdings-security-issues-warning-on-possible-data-breach/article/529460/

Olenick, Doug. (2016b). OptumRx customer records on stolen laptop compromised. Retrieved 28-05, 2017, from https://www.scmagazine.com/optumrx-customer-records-on-stolen-laptop-compromised/article/529045/

Pereira, Murilo Tito. (2009). Forensic analysis of the Firefox 3 Internet history and recovery of deleted SQLite records. Digital Investigation, 5(3), 93-103. doi: https://doi.org/10.1016/j.diin.2009.01.003

Press Herald. (2016). White House quiet about data breach. Retrieved 28-05, 2017, from http://www.pressherald.com/2016/04/04/white-house-quiet-about-data-breach/

Shropshire-Star. (2015). Laptop with MoD Donnington employee details stolen. Retrieved 20-09, 2017, from https://www.shropshirestar.com/news/2015/06/06/laptop-with-mod-donnington-employee-details-stolen/

Snell, Elizabeth. (2015a). ISMA Data Breach Reportedly from IT Head's Stolen Devices. Retrieved 20-09, 2017, from https://healthitsecurity.com/news/isma-data-breach-reportedly-from-it-heads-stolen-devices

Snell, Elizabeth. (2015b). PHI Safety Compromised After Texas Laptop Theft. Retrieved 19-09, 2017, from https://healthitsecurity.com/news/phi-safety-compromised-after-texas-laptop-theft

Snell, Elizabeth. (2016). Stolen Laptop Leads to Possible Health Data Breach in CO. Retrieved 26-05, 2017, from http://healthitsecurity.com/news/stolen-laptop-leads-to-possible-health-data-breach-in-co

Taiabul Haque, S. M., Wright, Matthew, & Scielzo, Shannon. (2014). Hierarchy of users′ web passwords: Perceptions, practices and susceptibilities. International Journal of Human-Computer Studies, 72(12), 860-874. doi: https://doi.org/10.1016/j.ijhcs.2014.07.007

Whitbourn, Michaela. (2016). Sydney University 'lost' computer containing sensitive student information. Retrieved 21-05, 2017, from http://www.smh.com.au/nsw/sydney-university-lost-computer-containing-sensitive-student-information-20160304-gnb4fa.html

WISN. (2015). Humana reports data breach that could affect up to 2,800. Retrieved 19-09, 2017, from http://www.wisn.com/article/humana-reports-data-breach-that-could-affect-up-to-2-800/6328907

A COMPARISON OF 2D AND 3D DELAUNAY TRIANGULATIONS FOR FINGERPRINT AUTHENTICATION

Marcelo Jose Macedo, Wencheng Yang, Guanglou Zheng, Michael N Johnstone
Security Research Institute, School of Science, Edith Cowan University, Perth, Western Australia
{m.macedo, w.yang, g.zheng, m.johnstone}@ecu.edu.au

Abstract

The two-dimensional (2D) Delaunay triangulation-based structure, i.e., Delaunay triangle, has been widely used in fingerprint authentication. However, we also notice the existence of three-dimensional (3D) Delaunay triangulation, which has not been extensively explored. Inspired by this, in this paper, the features of both 2D and 3D Delaunay triangulation-based structures are investigated and the findings show that a 3D Delaunay structure, e.g., Delaunay tetrahedron, can provide more feature types and a larger number of elements than a 2D Delaunay structure, which was expected to provide a higher discriminative capability. However, higher discrimination does not necessarily lead to better performance, especially in biometric applications, when biometric uncertainty is unavoidable. Experimental results show that the biometric uncertainty such as missing or spurious minutiae causes more negative influence on the 3D Delaunay triangulation than that on the 2D Delaunay triangulation in three out of four experimental data sets.

Keywords: Biometrics, Fingerprint, Delaunay triangulation, Delaunay tetrahedron

INTRODUCTION

In computer security, personal authentication properties are derived from knowledge, ownership of an object or some physical property of the self, commonly expressed as something you know (e.g., a password), something you have (e.g., a smart card) or something you are (e.g., a fingerprint). It is the last of these properties that is of interest here. Biometrics, derived from the Greek for life (βιο) and to measure (μετρικός) is especially attractive as the other methods of authentication tend to be less secure. People choose poor passwords or misplace access cards. Fortunately, people do not need to remember their fingerprint and cannot lose it (except under somewhat extreme circumstances).

The fingerprint is a complex set of segments formed by ridges and valleys that are unique to each person. Fingerprint matching is a process used to classify and match the fingerprint based on a group of features that are extracted from fingerprint images and used during the comparison process to minimise the effects of biometric uncertainty or error. Biometric uncertainty, which is generated during the fingerprint image acquisition process by image distortion, translation, and rotation, is undesirable in fingerprint matching due to the fact that it can generate unreliable results as noted by Yang et al. (2015) and Wang et al. (2017). In other words, two fingerprint images never look exactly the same even if they are scanned one after another which can affect the reliability of the fingerprint extraction and comparison process.

To mitigate the negative influence of biometric uncertainty, especially non-linear distortion, many local structures have been explored. Examples include tree-based structures (Moayer and Fu, 1986), n-nearest minutiae based structures (Liu et al., 2011) and minutia-pair based structures (Ahmad et al., 2011). From these examples, Delaunay triangulation has shown to be an effective structure which has been subsequently widely applied in fingerprint-based authentication systems, as it is able to provide some favorable local and global features (Yang et al., 2014). For example, in a Delaunay triangulation, each minutia will form a stable structure with its neighbors. Even if there is a certain degree of non-linear distortion, the local structure is still invariant. Moreover, missing or spurious minutiae only influence the local triangles that contain those missing or spurious minutiae, thus the error is localised.

There is considerable literature using Delaunay triangulation to achieve promising matching performance and security. For instance, in Parziale and Niel (2004), a Delaunay triangulation net is formed by a set of minutia points. Several invariant features such as rotation and translation are extracted from each Delaunay triangle and utilised to conduct matching. In the method of Deng and Huo (2005), some edge pairs are considered as best-matching pairs in a first round of matching. Then a second round is conducted, with matching based on global features and a matching score is computed under the guidance of the best-matching edge pairs obtained in the first round. In Yin et al. (2005), some similar minutiae pairs are chosen from both template and query fingerprint

images and considered as references for alignment. Then the aligned feature sets based on the references are used for fingerprint matching. There are other Delaunay triangulation-based fingerprint matching methods which can be found in Liang et al., (2006), Wang and Gavrilova (2006), Liang et al. (2007), Zhang and Yan (2007), Yang et al. (2012), Yang et al. (2014), and Yang et al. (2013).

All of the abovementioned Delaunay triangulation-based methods use only the 2D pattern of a Delaunay triangulation (as shown in Figure 1.a). However, the 3D Delaunay Triangulation (as shown in Figure 1.b) also exists as shown by Maur (2002), but it has not been extensively explored, suggesting that it might be a fruitful avenue to traverse as a research challenge and solution for application to real-world problems, e.g., fingerprint matching. Each unit, e.g., tetrahedron (a tetrahedron is called a 3D triangle in this paper), of 3D Delaunay triangulation, would be expected to have better discriminative ability than a unit, e.g., Delaunay triangle, of a 2D Delaunay triangulation, as the former have more properties. A tetrahedron can provide more feature elements than a triangle. For example, a 3D triangle has six edges, whilst a Delaunay triangle has only three edges (Frey et al., 1998) as shown in Figure 2 and Figure 3, respectively.

In this paper, the feature differences between 2D and 3D Delaunay triangulation-based structures are investigated. Also, through experiments, the effect of spurious and missing minutiae on both 2D and 3D Delaunay triangulation is studied, which lays a basis for further application of 3D Delaunay triangulation in fingerprint matching.

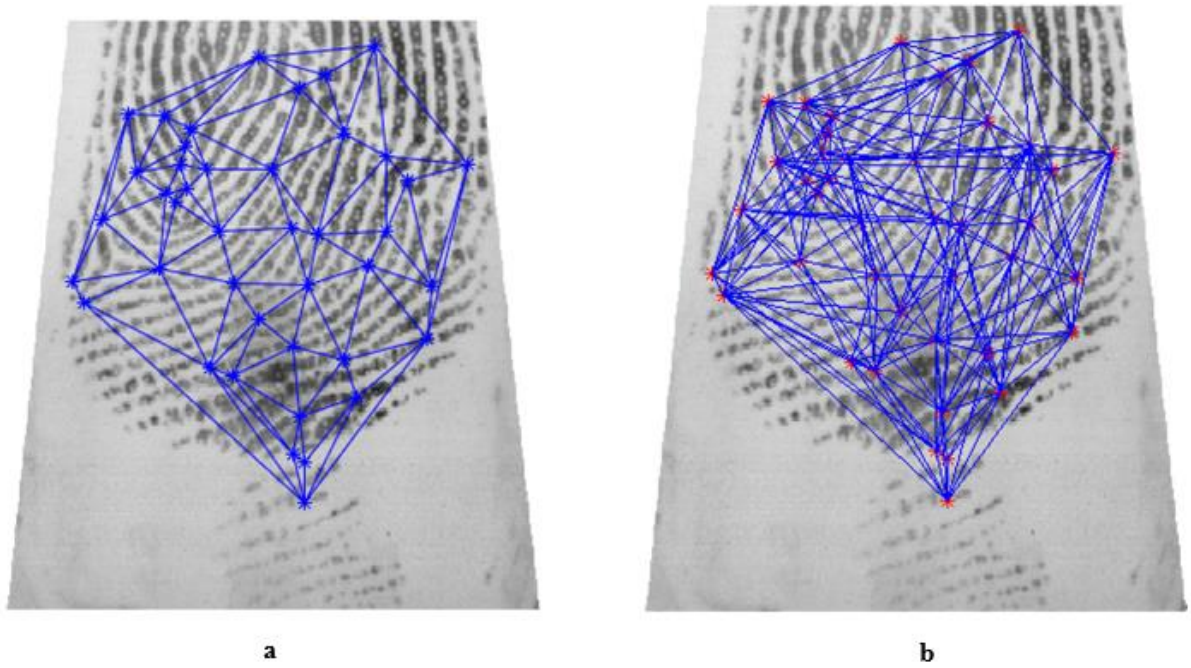


Figure 1. An example of (a) a 2D Delaunay triangulation, and (b) a 3D Delaunay triangulation

FEATURE DIFFERENCES BETWEEN A 2D TRIANGLE AND A 3D TRIANGLE

Definition of a 2D triangle and its features

In mathematics, a triangle is defined as a three-sided polygon which has a sum of 180 degrees for all its internal angles, as shown in Figure 2. A triangle is the simplest form of a polygon and has been studied and applied in areas such as construction due to its strength and precision. Any triangle belongs to one of three types, Equilateral, Isosceles or Scalene. Specifically, the Equilateral triangle has three sides of equal length and three equal angles, the Isosceles triangle has two sides of equal length and two equal angles, and the Scalene has no equal sides of equal length and no equal angles. Trigonometry studies the relationship between angles and sides of these shapes and there are many features that can be extracted from them as listed in Table 1.

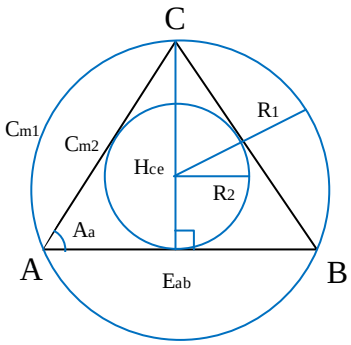


Figure 2. An example of a 2D triangle

Table 1. Features that can be extracted from a 2D triangle

Feature type	Definition	# Elements
Edge	A line connects two points, e.g., E_{ab}	3
Angle	The curvature of a combination of two lines with a common endpoint, e.g., A_a	3
Radius of the inscribed circle	A straight line from the centre of the inscribed circle to the circumference, e.g., R_2	1
Radius of the circumscribed circle	A straight line from the centre of the circumscribed circle to the circumference, e.g., R_1	1
Altitude (Height)	The length of a line segment through a vertex and perpendicular to the opposite edge, e.g., H_{ce}	3

Definition of a 3D triangle and its features

A 3D triangle defined in this paper is a tetrahedron as shown in Figure 3. It is formed by the four triangular faces which will form a group of four vertices and six edges. Table 2 shows features that can be extracted from a 3D triangle.

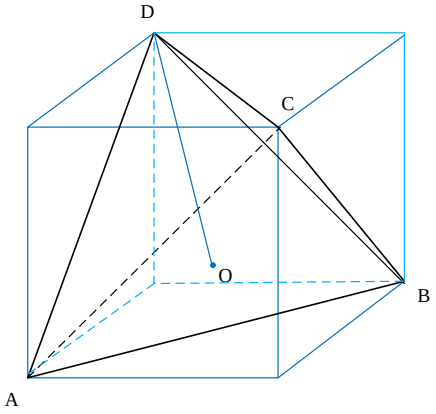


Figure 3. An example of a 3D triangle

Table 2. Features that can be extracted from a 3D triangle

Feature type	Definition	# Elements
Edge	As per a 2D triangle in Table 1	6
Angle	As per a 2D triangle in Table 1	12
Radius of the inscribed circle	As per a 2D triangle in Table 1	1
Radius of the circumscribed circle	As per a 2D triangle in Table 1	1
Altitude (Height) of a face	As per a 2D triangle in Table 1	12
Faces	A triangular shape, e.g. Δ_{ABC}	4
Altitude (Height)	The length of a line segment through a vertex and the perpendicular to a line containing the base shape, e.g., H_{DO}	4

From a comparison of tables 1 and 2, it can be seen that a 3D triangle can provide more features than a 2D triangle, in terms of feature type and number of elements, which means the former may be able to provide higher discrimination and lead to potential better matching performance in a 3D triangle based fingerprint authentication system than a system that uses 2D triangles.

EFFECT OF BIOMETRIC UNCERTAINTY IN 2D AND 3D DELAUNAY TRIANGULATION

The aim of this section is to demonstrate how missing and spurious minutiae can affect the Delaunay triangulation, a well-known issue of fingerprint matching. The number of missing or spurious minutiae depends on the quality of the fingerprint image that is used during the minutiae extraction process and this factor will likely influence the matching performance.

In experiments, the fingerprint image "1_1.tif" from the public database FVC2002 DB2 was chosen as the experimental subject and a set of minutiae M_p were extracted from it using the software Verifinger 4.0 SDK from Neurotechnology (Verfinger, 2010). A specific percentage of minutiae based on M_p are randomly removed or added to simulate biometric uncertainty.

Effect of missing minutiae to 2D and 3D triangle Delaunay triangulations

In this test, a certain number of points (10% and 20% of M_p) are randomly chosen and removed from the minutiae set M_p, the new minutiae sets are used to generate 2D and 3D Delaunay triangulations. Thus, four cases are generated:

Case 1: 10% of points are removed from minutiae set M_p and the new minutia set is used to generate 2D Delaunay triangulation, $DT_{2D\text{-}M10}$.

Case 2: 20% of points are removed from minutiae set M_p and the new minutia set is used to generate 2D Delaunay triangulation, $DT_{2D\text{-}M20}$.

Case 3: 10% of points are removed from minutiae set M_p and the new minutia set is used to generate 3D Delaunay triangulation, $DT_{3D\text{-}M10}$.

Case 4: 20% of points are removed from minutiae set M_p and the new minutia set is used to generate 3D Delaunay triangulation, $DT_{3D\text{-}M20}$.

The generated 2D Delaunay triangulations, e.g., $DT_{2D\text{-}M10}$ and $DT_{2D\text{-}M20}$, and 3D Delaunay triangulations, e.g., $DT_{3D\text{-}M10}$ and $DT_{3D\text{-}M20}$, in Cases 1 to 4, are shown in Figure 4, where the small yellow circles represent the points (10% of M_p) that are removed from the original set M_p and the green ones represent the points (20% of M_p) that are removed from the original set M_p.

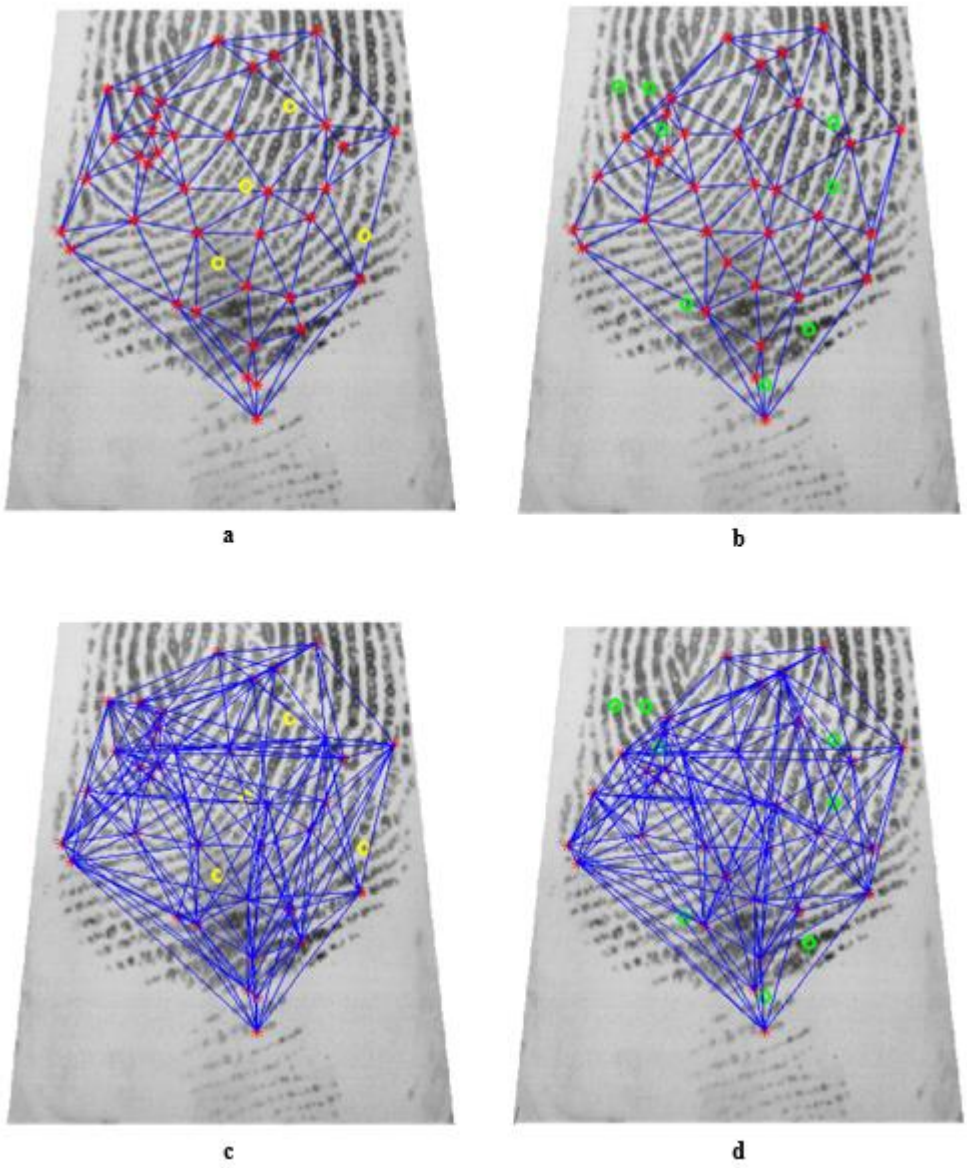


Figure 4. The generated 2D and 3D Delaunay triangulations in Cases 1 to 4

Effect of spurious minutiae in 2D and 3D triangle Delaunay triangulations

In this test, a certain number of points (10% and 20% of M_p) are randomly generated and added into the minutiae set M_p, the new minutia sets are used to generate 2D and 3D Delaunay triangulations. As before, there are four cases generated:

Case 5: 10% of points are added into minutiae set M_p and the new minutia set is used to generate 2D Delaunay triangulation, $DT_{2D\text{-}S10}$.

Case 6: 20% of points are added into minutiae set M_p and the new minutia set is used to generate 2D Delaunay triangulation, $DT_{2D\text{-}S20}$.

Case 7: 10% of points are added into minutiae set M_p and the new minutia set is used to generate 3D Delaunay triangulation, $DT_{3D\text{-}S10}$.

Case 8: 20% of points are added into minutiae set M_p and the new minutia set is used to generate 3D Delaunay triangulation, $DT_{3D\text{-}S20}$.

The generated 2D Delaunay triangulations, e.g., $DT_{2D\text{-}S10}$ and $DT_{2D\text{-}S20}$, and 3D Delaunay triangulations, e.g., $DT_{3D\text{-}S10}$ and $DT_{3D\text{-}S20}$, in Cases 5 to 8, are shown in Figure 5, where the small yellow circles represent the points (10% of M_p) that are added into the original set M_p and the green ones represent the points (20% of M_p) that are added into the original set M_p.

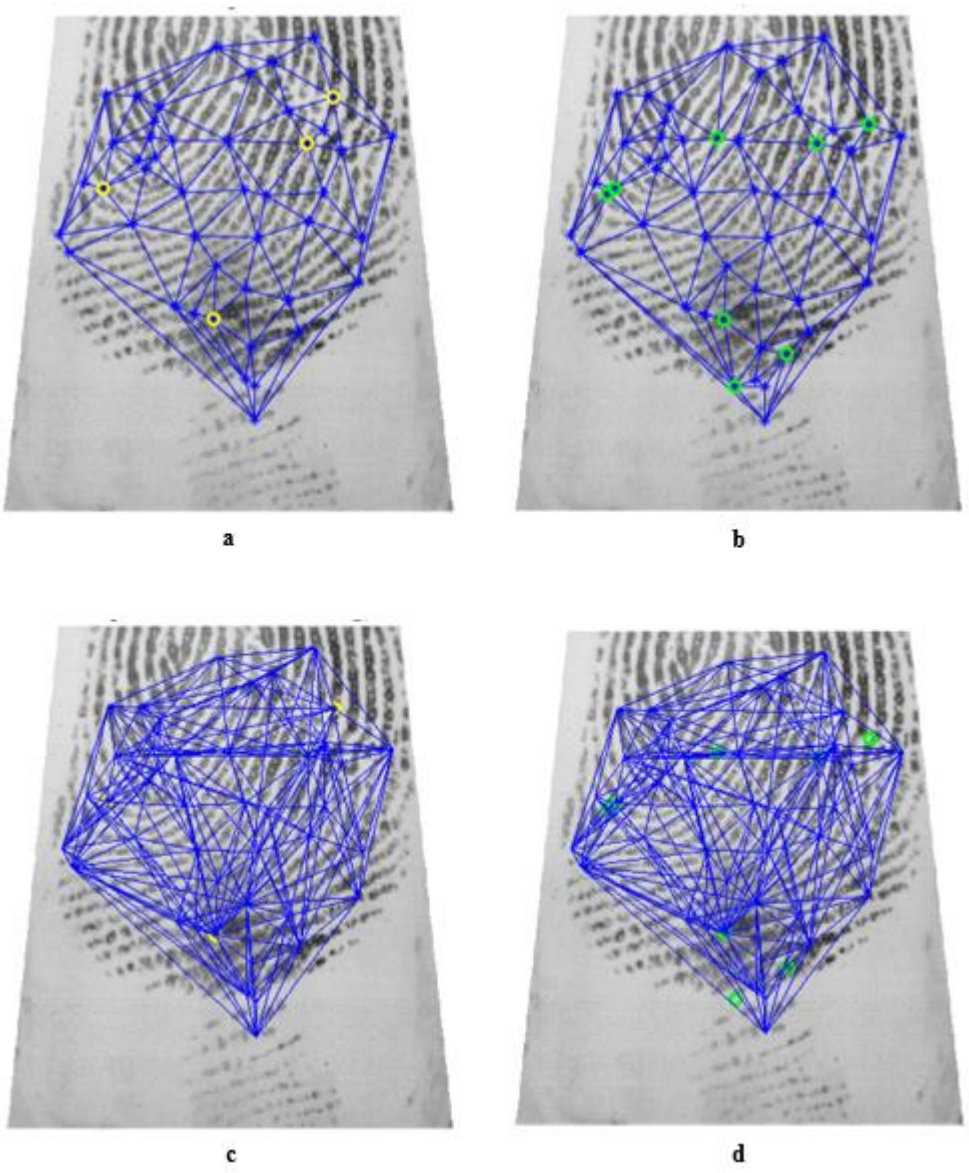


Figure 5. The generated 2D and 3D Delaunay triangulations in Cases 5 to 8

The effects of missing and spurious minutiae to 2D and 3D Delaunay triangulations are listed and compared in Table 3, from which it can be seen that when 10% of total minutiae are missing, compared with original number of units, 10.81% is reduced in 2D triangulation, while 13.37% is reduced in 3D triangulation, which means that missing minutiae have more negative influence on 3D triangulation than that on 2D triangulation. However, a contrary result is reported when 20% is missing. In the cases of spurious minutiae, it can be seen that spurious minutiae create more negative influence on 3D Delaunay triangulation with 11.05% and 24.42% of original number of units is increased (compared with only 10.81% and 21.62% in 2D Delaunay triangulation), when 10% and 20% spurious minutiae are added, respectively.

Table 3- Comparison between 2D and 3D Delaunay triangulation under biometric uncertainty

Structure type	Units created with minutiae set M_p	Units created with 10% points missing from M_p	Units created with 20% points missing from M_p	Units created with 10% points added into M_p	Units created with 20% points added into M_p
2D triangles	74 100%	66 89.19% (-10.81%)	57 72.03% (-27.97%)	82 110.81% (+10.81%)	90 121.62% (+21.62%)
3D triangles	172 100%	149 86.63% (-13.37%)	133 77.33% (-22.67%)	191 3D triangles 111.05% (+11.05%)	214 124.42% (+24.42%)

CONCLUSION

In this paper the feature differences between 2D and 3D Delaunay triangulation-based structures are investigated and it is demonstrated that a 3D triangle has more feature types and a larger number of elements, thus it may possess higher discriminative ability than a 2D triangle. However, the use of 3D Delaunay triangulation does not mean absolute better matching performance, especially under the presence of biometric uncertainty, e.g., missing or spurious minutiae. Biometric uncertainty is simulated by randomly adding or removing points from a generated minutiae set. The experimental results show that 3D Delaunay triangulation is more sensitive to the missing and spurious minutiae than a 2D Delaunay triangulation in three cases out of four. In future research, to explore and potentially reduce the negative influence of biometric uncertainty on the 3D Delaunay triangulation is needed.

ACKNOWLEDGEMENT

This paper is supported by Early Career Grant Scheme of ECU Australia through project G1003411.

REFERENCES

Ahmad, T, Hu, J. and S. Wang (2011). Pair-polar coordinate-based cancelable fingerprint templates, *Pattern Recognition*, 44 pp. 2555-2564.

Deng H. and Huo Q. (2005) Minutiae Matching Based Fingerprint Verification Using Delaunay Triangulation and Aligned-Edge-Guided Triangle Matching. In: Kanade T., Jain A., Ratha N.K. (eds) Audio- and Video-Based Biometric Person Authentication. AVBPA 2005. Lecture Notes in Computer Science, vol 3546. Springer, Berlin, Heidelberg.

Frey, P.J., Borouchaki, H., and George, P-L., (1998). 3D Delaunay mesh generation coupled with an advancing-front approach, *Computer Methods in Applied Mechanics and Engineering*, 157 pp. 115-131.

Liang, X, Asano, T and Bishnu, A. (2006). Distorted fingerprint indexing using minutia detail and Delaunay triangle, IEEE 3rd International Symposium on Voronoi Diagrams in Science and Engineering, pp. 217-223.

Liang, X, Bishnu, A and Asano, T. (2007). A robust fingerprint indexing scheme using minutia neighborhood structure and low-order Delaunay triangles, *IEEE Transactions on Information Forensics and Security*, 2 pp. 721-733.

Liu, E., Zhao, H., Liang, J., Pang, L., Xie, M., Chen, H., Li, Y., Li, P., and Tian, J. (2011). A key binding system based on n-nearest minutiae structure of fingerprint, *Pattern Recognition Lett*., 32 pp. 666-675.

Maur, P. (2002). Delaunay triangulation in 3D, Technical Report DCSE/TR-2002-02, Department of Computer Science and Engineering, University of West Bohemia.

Moayer, B. and Fu, K-S (1986). A Tree System Approach for Fingerprint Pattern Recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8 pp. 376-387.

Parziale G. and Niel A. (2004) A Fingerprint Matching Using Minutiae Triangulation. In: Zhang D., Jain A.K. (eds) *Biometric Authentication. Lecture Notes in Computer Science*, vol 3072. Springer, Berlin, Heidelberg.

VeriFinger. (2010). VeriFinger, S. D. K. Neuro Technology.

Wang, C. and Gavrilova, M.L. (2006). Delaunay Triangulation Algorithm for Fingerprint Matching, 3rd International Symposium on Voronoi Diagrams in Science and Engineering, pp. 208-216.

Wang, S., Yang, W., and Hu, J. (2017). Design of Alignment-Free Cancelable Fingerprint Templates with Zoned Minutia Pairs, *Pattern Recognition*, 66. pp. 295-301.

Yang, W., Hu, J., and Wang, S. (2012). A Delaunay Triangle-Based Fuzzy Extractor for Fingerprint Authentication, 11th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom), IEEE2012, pp. 66-70.

Yang, W., Hu, J., and Wang, S. (2013). A Delaunay triangle group based fuzzy vault with cancellability, 6th International Congress on Image and Signal Processing (CISP), 2013, pp. 1676-1681.

Yang, W., Hu, J., and Stojmenovic, M. (2014). NDTC: A novel topology-based fingerprint matching algorithm using N-layer Delaunay triangulation net check, 7th IEEE Conference on Industrial Electronics and Applications (ICIEA), IEEE2012, pp. 866-870.

Yang, W., Hu, J., and Wang, S. (2014). A Delaunay Quadrangle-Based Fingerprint Authentication System With Template Protection Using Topology Code for Local Registration and Security Enhancement, IEEE *Transactions on Information Forensics and Security*, 9 pp.1179-1192.

Yang, W., Hu, J, Wang, S., and Chen, C. (2015). Mutual dependency of features in multimodal biometric systems. *Electronics Letters*. 51. 234–235.

Yin, Y., and Zhang, H., and Yang, H.K. (2005). A method based on Delaunay triangulation for fingerprint matching, Proc. International Society for Optical Engineering, 5779 pp. 274-281.

Zhang, Q. and Yan, H. (2007). Fingerprint orientation field interpolation based on the constrained delaunay triangulation, *International Journal of Information and Systems Sciences*, 3 pp. 438-452.

LITERATURE-BASED ANALYSIS OF THE INFLUENCES OF THE NEW FORCES ON ISMS: A CONCEPTUAL FRAMEWORK

Zahir Al-Rashdi, Dr Martin Dick, Dr Ian Storey
School of Business Information Technology and Logistics, RMIT University, Melbourne, Australia
zahir.al-rashdi@rmit.edu.au, martin.dick@rmit.edu.au, ian.storey@rmit.edu.au

Abstract

This paper presents an analysis that arose from a comprehensive review of the academic and professional literature of two areas – information security management systems (ISMS) and information resources – and their relationship with information security. It analyzes the role of ISMS in protecting an organization's information environment and infrastructure. It has identified four key areas that strongly influence the safety of information resources: cloud computing; social media/networking; mobility; and information management/big data. Commonly referred to as 'new forces', these four aspects are all growing exponentially and are not easily controlled by IT. Another key finding of the paper is that organizations aiming to protect their information resources need to adapt their ISMS in an environment where these new forces have exposed them to a range of external entities and influences.

Keywords: accountability, cloud computing, service provision, information security, information resources

INTRODUCTION

Gartner identified a range of emerging technologies that strongly influence information resources, which have been referred to as 'new forces' (Gartner 2013). These new forces – cloud computing, social media/networking, mobility, and information management/big data – cannot be controlled by IT groups (internal or external) and have been predicted as a high priority for IT spending in the next decade (Smith 2013). They are all growing at a fast pace (Gartner 2013), and the traditional controls (firewalls, malware detection, etc.) that IT groups have used to protect an organization's information resources are generally unable to cope with their rapid evolution on a daily basis (Crompton 2015). The following are examples of how these new forces are impacting on the frameworks that IT groups have installed to protect organizations' information resources:

- **Cloud computing:** The new service delivery styles and options offered by the cloud computing trend shift responsibility for security to external providers, where IT groups retain only partial responsibility for security and service delivery (Scholtz 2013).
- **Social media/networking:** This is creating different, more extensive aspects of collaboration, and there has been a change in users' behavior and in the communities in which they work (Kim 2012). Again, this technology has opened up the organization's information resources to external influences.
- **Mobility:** A wide range of new access has been directed at different applications and data, and end users have been offered a broad variety of device options (Markelj & Bernik 2012) – the heterogeneity of location and device.
- **Information management/big data:** This has altered the relationship of technology to information consumption, as the data now flows from different federated sources in either structured or unstructured forms (Marchand 1985). This revealed data is analyzed using new methodologies foreign to various IT groups (Gartner 2013).

These examples show that these new forces have exposed organizations' information resources to a wide range of external entities and influences, which has significant implications for the design, management and use of the organization's information security management systems (ISMS). IT groups need to respond to this exposure by considering the information security aspects against access and exchange of information. This proves challenging as they also need to continue to meet the individual's expectations – those that are now more knowledgeable about the use of technology and are clearly requiring the capabilities these new forces are providing. The existence of the new forces encourages and demands IT groups to work toward reformation of their information security practices.

This paper presents a conceptual framework for understanding the impact of the emergence of these new forces on ISMS. According to Von Solms (2005)), ISMS are the processes and procedures used within an

organization to secure the information environment through information security, operational management and information security compliance management. A perfect ISMS is a complete and systematic management system that involves "management of humans, processes, and technologies" (Suhaimi, Goto & Cheng 2013, p. 31), in order to establish, implement, operate, monitor, review, maintain and optimize security to ensure confidentiality, integrity and availability of information.

In summary, the main motivation of this paper is to provide an in-depth understanding of the conceptual factors that comprise ISMS when used in the context of organizations' information security and information resources, to ensure their commitment to the security of business practices and compliance to address the rapid growth of the new forces (Siponen & Willison 2009).

BACKGROUND

There are two key concepts which need to be examined in the context of this paper: (1) ISMS; and (2) the concept of information resources.

Information security management systems (ISMS)

Hong et al. (2003) defined ISMS as technical methods, along with managerial processes, practically applied to information resources such as hardware, software and data, to ensure that organizational assets and personal privacy are protected. Research on ISMS has produced a considerable amount of definitions, embodying the different spheres of ISMS research. Both academics and practitioners have differing views and interpretations of the ISMS concept. For example, Eloff and Eloff (2003) argued that an ISMS can be defined as a management system employed to secure an information environment within an organization, with a good establishment and maintenance process and procedure to manage information technology security. The management and execution of ISMS requires some necessary actions: (1) identify the information security requirements; (2) ensure the right strategies are in place to meet these requirements; (3) verify the continuous evaluation and measurement of achieved objectives/results; and (4) ensure the compatibility and usability of both protection strategies and the ISMS by reviewing and improving them over time (Yeniman Yildirim et al. 2011).

Based on the above varying definitions of ISMS, it is clear that ISMS entails a number of components that form the information environment within an organization: (1) information system technology resources (hardware); (2) information system human resources (IT skilled people); (3) information system software resources (software); and (4) information system data resources (data). In this context, this study has used the ISMS definition offered by Hong et al. (2003)), because it covers the most important aspects of technology used in any organization, and is considered the most valuable asset to an organization in today's world.

Information resources

Information resources comprise hardware, software, data and IT human resources in an organization, and represent the main source for information (Bharadwaj 2000). The effective use of information resources is often a key indicator of an organization's ability to achieve a high level of information security and organizational privacy; although such a high investment in IT can also be a key indicator of failure, if not properly adjusted and controlled (Nolan 1994).

What is the relationship between information resources and ISMS?

Most organizations today are heavily dependent on the use of IT and information resources – the foundation of an organization, representing a key element of its growth and survival (Bharadwaj 2000).

Thus, there are escalating organizational concerns about information security including privacy and protection, risk management, and the management of information resources, which is ever-increasing. A proper solution therefore needs to be implemented to secure information resources. Some information resources are sensitive; meaning that a cost-optimal solution for secure access to information located across different servers or databases is needed, along with guidelines to ensure that the security and privacy of sensitive, unclassified information is not leaked. A combination of the latest technologies and strong human IT skills would strengthen organizational capability to safeguard information. Similarly, successful organizations should use technology and human IT skills to ensure the protection of organizational information resources and personal

privacy. ISMS is commonly considered a socio-technical system that encourages a combination of both technical and human elements (Siponen & Willison 2009).

RESEARCH METHODOLOGY

A systematic literature review of ISMS, information resources, the new forces and their relationship with an organization's information security was conducted using an adapted version of the methodology of Okoli and Schabram (2010). The key steps in the methodology are:

1. *Purpose of the Review* – As described in the introduction of this paper
2. *Protocol and Training* – As only one reviewer was used, training did not have to be done and the protocol focused on the intersection between the four new forces and ISMS.
3. *Searching for the Literature* - via two sources: (1) papers published in academic and professional literature; and (2) industrial reports published by well-known organizations such as Gartner, Microsoft, IBM, Cisco, and Business News Daily. When accessing the literature, the following keywords were used to search IEEE Xplore, ScienceDirect, the ACM Digital Library, ProQuest, and Google Scholar: cloud computing; cloud computing security; information security; ISMS; information systems; information resources; vulnerability; risk assessment and risk management; threats and information security breaches; auditability and trust; confidence; social media/networking; information security for cloud computing; information systems security; information technology security; information security management; cyber security; information resources; the new forces of information resources; information assurance; and information security practices and standards. The preliminary results were sourced from 500+ scholarly articles, industry standards and technical reports;
4. *Practical Screen and Quality Appraisal* – A review of the abstracts eliminated many of these papers and those remaining were then appraised as to their relevance to the purpose of the review. In total, over 350 papers were eliminated, leaving 150 papers were left remaining as suited to the literature review.
5. *Data Extraction* – The remaining papers were then examined thoroughly and relevant data was extracted.
6. *Synthesis of Studies* – A thematic analysis was performed based on the extracted data.
7. *Writing the review* – a summarized form of the review is presented in this paper.

CONCEPTUAL UNDERSTANDING OF THE NEW FORCES

The main outcomes of this review were that the scope of information resources has changed dramatically and is not really limited to one single source. That is extensive exchange of data across the Internet between different organizations has become common practice; it is not a choice anymore, but rather a mandatory task, an organization is forced to undertake. Such data exchange is represented by a variety of ways/media, made up of cloud computing, mobility, social media/networking, and information management/big data. These are a group of new technologies which are strongly interacting with information resources, which is why they are referred to as the 'new technology forces' or 'new forces'.

This review also revealed that traditional IT security methods are often unable to cope with information security issues that arise from the new forces of technology. ISMS can and needs to be frequently adjusted to deal with the evolving information security aspects revealed across the various information resources including cloud computing, mobility, social media/networking, and big data.

In addition, there are some internal characteristics that are likely to influence to what extent organizations apply ISMS, such as the business model, data holdings, technologies and applications, and the privacy risks they may raise for clients. For example, organizations managing highly sensitive data would need to implement the largest amount of ISMS components. Thus, each organization needs to customize its ISMS to accommodate the new forces.

Due to this vast and sophisticated technology (Prensky 2011), information resources are rapidly growing (Webster 2014). The scope of information resources (Kazan et al. 2012) has dramatically shifted from internal to external management (Lawrence et al. 2013), including the exchange of information Lacity and Hirschheim

(1993). Such changes are driven by different aspects of technology, based on the new forces – cloud computing (Xu 2012), social media/networking (Kim 2012), mobility (Tokuyoshi 2013), and information management/big data – which are all growing strongly and are not easily controlled by IT (Smith 2013). The organizational approach towards ISMS is therefore transforming, based on the new requirements of these new forces (Scholtz 2013). Figure 1 below presents a conceptual framework arising from the analysis that shows the overall interactions between the four new forces , information resources and ISMS.

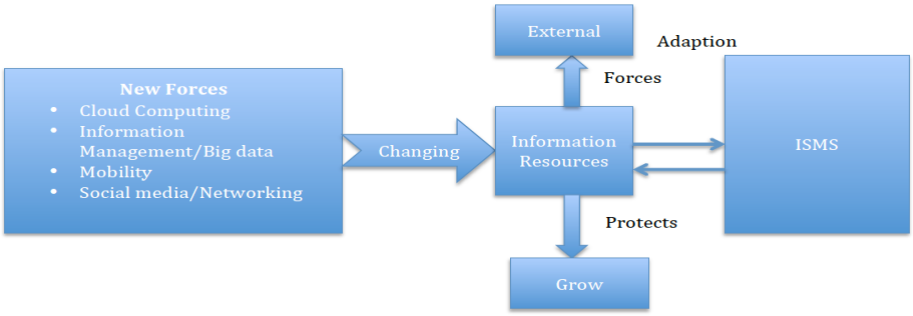


Figure 1: Conceptual Framework of the interaction of the New Forces with ISMS and Information Resources

Social media/networking

Social media/networking is the emergence of a new paradigm on the internet which allows communication and collaboration between online users and family, friends, social groups and other communities via social media channels and tools including Twitter, Facebook, MySpace and YouTube (Kim 2012).

Founded on the above definition, social media networks are considered one of the most influential factors for information resources (Xiang & Gretzel 2010), as a lot of information is gathered and exchanged (Kaplan & Haenlein 2010) among different internet users. Today, online communications via social networking applications is growing strongly and is used both personally and professionally (Jansen et al. 2009). Millions of users (Ellison 2007) are being attracted by social network sites (SNSs) such as MySpace, Facebook, Twitter and YouTube. These social media networks are receiving considerable attention from today's business executives, decision-makers and leaders, who are trying to introduce new ways of increasing their profits (Xiang & Gretzel 2010) by using social media applications such as Wikipedia, YouTube, Facebook, Second Life, and Twitter, which means lots of information (Kaplan & Haenlein 2010) is going to be exchanged and transferred.

The social networking sites (SNSs) are being integrated with user's daily practices (Hathi 2009) through mobile connectivity, blogging, photo sharing and video sharing as new communication tools (Ellison 2007). This indicates that the of sensitive information will be transferred and exchanged with different interests (Kim 2012). For example, Facebook (Edosomwan et al. 2011) is one of the most popular and strongest growing web applications in social networking services, where more than 500 million users all over the world use it either for work (Duggan & Brenner 2013) or pleasure, like playing games (Foster, Francescucci & West 2010). The emergence of social media and online social networking applications created real revolution to the working environment (Sturdevant 2011) in large enterprises. For example, IBM reported that (Hathi 2009) more than 40% of their employees preferred to work from customer location or home rather than attending IBM premises. Similarly, Cisco (Cisco 2010) reported that over 60% of their employees believe that productivity no longer means to work from the office, but rather productivity depends on knowledge and the ability to share that knowledge. Thus, lots of enterprises allow their employees to access online social networking sites; however this brought many security breaches to business and organisations (Kaplan & Haenlein 2010). For example, the usernames and passwords of Facebook and MySpace being sold to underground networks where sensitive information was stolen by cybercriminals (Shulman 2010), meanwhile their accounts were being hijacked (Kim 2012) and the stolen accounts used by hackers for phishing scams (Kaplan & Haenlein 2010).

Therefore, information security is a new and increasing challenge in the field of cyber security brought along by the use of social networks and the integration of ISMS within organisation to closely and properly monitor the use of different aspects of social media became a mandatory practice to eliminate the breach of information that would occur through this new rapidly-growing force.(Edosomwan et al. 2011).

Mobility

The use of mobile devices such as laptops, smartphones and PDAs to access data has become far more frequent in recent times (Markelj & Bernik, 2012). This increase in the use of mobility raises questions about corporate data security and privacy (Miller, Voas & Hurlburt 2012), which should not be exposed or compromised (Rose 2013). The future of cyber risk prevention is an area that needs to be urgently addressed by researchers, technology experts and policymakers (Kenny 2014). Most mobile device users are not proactive in dealing with information stored in their mobile by taking proper protection actions (Tokuyoshi 2013). Therefore, losing the mobile device could result in exposure of sensitive information, which is sometimes much more important and valuable than the mobile itself (Markelj & Bernik 2012). Thus, corporations' information security policies and risk management procedures and should be constantly reviewed and upgraded (Markelj & Bernik 2012).

There are many studies (Escherich 2014; Miller, Voas & Hurlburt 2012; Morrow 2012; Rose 2013; Tokuyoshi 2013) that have focused on mobility, especially the use of mobile devices in the workplace. According to a recent survey conducted by Gartner Inc. in 2013 which measured the use of personal devices for business use, or bring your own devices (BYODs) (Escherich 2014), almost one-quarter acknowledged they had experienced a security issue with their private device; yet only 27% of those users who contributed to the survey reported the issue to their employer. In summary, it would appear that mobility raises many problems for an organization's ISMS, and most will need to accelerate an introduction of "the right mix of mobile security defences to balance protection, governance and user flexibility" (Escherich 2014, p. 1).

Information management/big data

Information management in a business and organizational context means a collection of data from different departments, made ready for processing, with decisions made at the higher levels of the organization – mostly strategic rather than tactical – and external information becoming more relevant than information sourced internally (Marchand 1985). Information management is now considered a critical resource to ensure privacy (Bélanger & Crossler 2011) due to the advanced use of IT and information systems like intranets (Curry & Stancich 2000).

Dhillon (1999)) contended that one of the main issues in regards to current information management is information security – that is, the protection of information assets of the body corporate, due to the massive use of IT and growth dependent on electronic network resources.

In the past five years, the media/entertainment industry has shifted to digital recording, production and delivery, and as a result large amounts of rich content including user viewing behaviors are collected. Furthermore, significant amounts of data are now collected within the transportation, logistics, retail, utilities and telecommunications sectors via various media/technology such as GPS transceivers, RFID tag readers, smart meters and cell phones; call data records CDRs and all collected data are used to optimize operations and drive operational business intelligence (BI) to realize immediate business opportunities (Kaisler et al. 2013).

This literature review indicates that a huge amount of organizational information is being shared among users and organizations, within the provision of information management. Thus, the security of information management is deemed one of the key factors for measuring the quality of service by organizations or service providers. The implementation of ISMS would eliminate many of these information management security issues.

Cloud computing

Cloud computing relates to the use of online computing services, and is considered an on-demand IT service or product based on a business model. This is where users can access software and hardware via cloud services including SaaS, Platform as a Service (PaaS) and Infrastructure as a Service (IaaS), which are managed remotely by third parties (Sreenivas et al. 2013). It can therefore be perceived as significantly opening an organization's information resources to the external world.

Cloud computing has expanded into one of the fastest-growing portions of the IT industry (Chung & Chun 2014), and has become a promising business concept where a huge amount of information for both individuals and enterprises is now placed. However, transformation of data distribution and storage in the cloud has generated a real concern towards data privacy and data protection, and a rise in questioning about how safe the

cloud environment is. Such questioning should be considered by organizations before making the decision to deploy their business into the cloud (Subashini & Kavitha 2011).

Securing information that has been transferred to the cloud is a critical issue for the success of information systems, communications and information security (Zissis & Lekkas 2012). As can be seen, there is a clear need for organizations to adapt ISMS to this new, fast-growing force.

ADAPTING ISMS TO THE NEW FORCES

The new forces – mobility, social media/networking, cloud computing, and information management/big data – are rapidly changing the information resources of organizations. Consequently, the ISMS with its core mission of protecting the information resources of the organization must adapt. In particular, increased external openness of information resources must be considered when adapting an organization's ISMS to these new forces. The following section details the adaptions that were identified from the literature analysis to allow organizations to adapt their ISMS to the impacts of the four new forces. Of course, this analysis is limited to adaptions that have already been identified in the existing literature in this area.

Organizations need to move from applying traditional information security controls such as firewalls and malware detection, and pay more attention to users behavior or internal staff that are using their own devices. This is achievable by ensuring a solid and acceptable change in the security infrastructure, design and implementation of controls to minimize preventative measures and balance the use of policies, controls, rights and responsibilities. Such balance would maximize human potential by increasing trust and independent decision-making (Scholtz 2013).

One way that organizations need to adapt is by moving from static to dynamic defences. Gartner (2013) has predicted that one of the likely scenarios of the information security changes caused by these new forces is that by 2020 the allocated budget of enterprise information security will rise to 60%, up from less than 10% in 2013, for rapid detection and response approaches.

Another key adaption is the development of ISMS policy that allows for these new forces. For example, in terms of mobility, it has been recognized that "security policies are still incomplete in many organizations, and contain gaps and other inconsistencies that don't measure up to business obligations" (Escherich 2014, p. 1). Such policy gaps need to be addressed.

Another significant adaption aspect is the strategic consideration of the new forces. First, how they will impact on the potential for interruption/disruption to IT within the business, including evaluating the risk associated with enterprise information security. Second, the organization needs to strategically determine the required investment in ISMS when adopting the new forces, and how much needs to be allocated to adapt it to an acceptable level of risk.

In addition, it should also be highlighted that these new forces are not a comprehensive list of every technology that is ready for adoption or incorporation into the strategic planning process. Enterprises should use them as a starting point and customize them based on their industry, unique business needs, technology adoption model, and which category their business is classified, and then customize their ISMS accordingly. Table 1 summarises the impacts of the new forces on Information Resources and ISMS

New Force	Impact on Information Resources	Impact on ISMS
Social Media/Networking	• Significantly increased information is gathered and exchanged • Interaction between personal and professional data and integration with user's daily practices • Use of social media as a business tool	• Many security breaches to business and organisations due to the extensive use of information brought along by the use of social networks (Edosomwan et al. 2011).
Mobility	• Increased usage and generation of information due to the very rapid uptake of mobile technology • The rapid advancement in	• Increased vectors for cyber attack, including much easier physical loss and theft • The addition of personal

	functionality of mobiles is also causing increased information resources	devices to the business environment due to BYOD • Poor security practices in many mobile devices
Information Management/Big Data	• The control of information physically became more complex. • The uncoordinated and fragmented nature of much big data • The need for much better understanding of the information resources to allow knowledge management and machine learning	• Privacy becoming a much more important issue • Coping with information overload
Cloud Computing	• Made storage of information resources much more affordable and manageable	• Management of the 3rd party outsourcing relationship with regards to the ISMS • Ensuring compatibility with external ISMS

Table 1 Impact of New Forces on Information Resources and ISMS

CONCLUSIONS AND FUTURE WORK

This study has conducted an extensive analysis of literature relating to ISMS, including the various aspects of information resources for information security to develop a model of how the new forces influence an organization's ISMS. It has been recommended here that organizations adapt their ISMS to accommodate the changes these new forces have on information resources, with a number of suggestions made on how the ISMS should be adapted; although this is far from exhaustive. This research is part of an ongoing research program in this area.

This study is expected to have several important implications for practitioners and researchers. The findings will likely contribute to the growing awareness of the importance of the proper implementation of ISMS; resulting in a better understanding of the importance of information security factors that motivate policymakers to adopt ISMS projects prior to, during and after their implementation. It will also assist information security decision-makers to evaluate what has happened and why in terms of any security issues. This study also provides a holistic and heuristic ISMS framework that enables a theoretical-based description and analysis of the gap that exists between the ideal and the actuality of the institutional and technical environments of ISMS implementation in regards to the protection of information resources.

Finally, this study is expected to contribute to the body of knowledge of information systems and information security as part of the development of new theory. It provides a practical contribution, where new insights for ISMS implementers and investors will help improve their services. Developing this model will be achieved by conducting a series of case studies to examine the real-life experiences of organizations in ensuring information security when adopting ISMS, including analyzing the four conceptual factors that underlie the new forces.

REFERENCES

Bélanger, F & Crossler, RE 2011, 'Privacy in the digital age: a review of information privacy research in information systems', *MIS Quarterly*, vol. 35, no. 4, pp. 1017-42.

Bharadwaj, AS 2000, 'A resource-based perspective on information technology capability and firm performance: an empirical investigation', *MIS Quarterly*, pp. 169-96.

Chung, D & Chun, SG 2014, 'ADOPTION AND IMPLEMENTATION OF CLOUD COMPUTING SERVICES: A RAILROAD COMPANY CASE', *Issues in Information Systems*, vol. 15, no. 2.

Cisco 2010, *Cisco, Social Media: Cultivate Collaboration and Innovation*.

Crompton, JC 2015, *Gartner's Top 10 Strategic Technology Trends For 2015* January 26, <http://blogs.sap.com/innovation/innovation/gartners-top-10-strategic-technology-trends-2015-webinar-recap-02113450

Curry, A & Stancich, L 2000, 'The intranet — an intrinsic component of strategic information management?', *International Journal of Information Management*, vol. 20, no. 4, pp. 249-68.

Dhillon, G 1999, 'Managing and controlling computer misuse', *Information Management & Computer Security*, vol. 7, no. 4, pp. 171-5.

Duggan, M & Brenner, J 2013, *The demographics of social media users, 2012*, vol. 14.

Edosomwan, S, Prakasan, SK, Kouame, D, Watson, J & Seymour, T 2011, 'The history of social media and its impact on business', *Journal of Applied Management and Entrepreneurship*, vol. 16, no. 3, pp. 79-91.

Ellison, NB 2007, 'Social network sites: Definition, history, and scholarship', *Journal of Computer-Mediated Communication*, vol. 13, no. 1, pp. 210-30.

Eloff, JH & Eloff, M 2003, 'Information security management: a new paradigm', paper presented to Proceedings of the 2003 annual research conference of the South African institute of computer scientists and information technologists on Enablement through technology, pp. 130-136.

Escherich, M 2014, 'Gartner Survey Shows U.S. Consumers Have Little Security Concern With BYOD', viewed 19 May 2014.

Foster, MK, Francescucci, A & West, BC 2010, 'Why users participate in online social networks', *International Journal of e-Business Management*, vol. 4, no. 1, p. 3.

Gartner 2013, *Analysts to Explore Emerging Business Strategies at Gartner Symposium/ITxpo 2013*, Gartner, <http://www.gartner.com/newsroom/id/2613016

Hathi, S 2009, 'How social networking increases collaboration at IBM', *Strategic Communication Management*, vol. 14, no. 1, p. 32.

Hong, K-S, Chi, Y-P, Chao, LR & Tang, J-H 2003, 'An integrated system theory of information security management', *Information Management & Computer Security*, vol. 11, no. 5, pp. 243-8.

Jansen, BJ, Zhang, M, Sobel, K & Chowdury, A 2009, 'Twitter power: Tweets as electronic word of mouth', *Journal of the American society for information science and technology*, vol. 60, no. 11, pp. 2169-88.

Kaisler, S, Armour, F, Espinosa, JA & Money, W 2013, 'Big data: Issues and challenges moving forward', paper presented to System Sciences (HICSS), 2013 46th Hawaii International Conference on.

Kaplan, AM & Haenlein, M 2010, 'Users of the world, unite! The challenges and opportunities of Social Media', *Business horizons*, vol. 53, no. 1, pp. 59-68.

Kazan, W, Font, A, Akmal, M, Grossberg, SD, Penov, FP, Truelove, BN, Chandrasekaran, V & Bahrainwala, S 2012, *PRESENTING CONTENT ITEMS SHARED WITHIN SOCIAL NETWORKS*, US Patent 20,120,151,383.

Kenny, J 2014, *Privacy problems are here to stay*, Blouin News USA,Manhattan, <http://blogs.blouinnews.com/blouinbeattechnology/2014/05/29/privacy-problems-are-here-to-stay/

Kim, HJ 2012, 'Online social media networking and assessing its security risks', *International Journal of Security and Its Applications*, vol. 6, no. 3, pp. 11-8.

Lacity, MC & Hirschheim, R 1993, 'The information systems outsourcing bandwagon', *Sloan management review*, vol. 35, no. 1, p. 73.

Lawrence, S, Peterson, S, Gregory, M, Gourley, CR, Korth-McDonnell, P & Stewart, J 2013, *Linking a retail user profile to a social network user profile*, Google Patents.

Marchand 1985, 'Infonnation management: strategies and tools in transition', *Information Management Review*, vol. 1, pp. 27-34.

Markelj, B & Bernik, I 'Mobile devices and corporate data security'.

---- 2012, 'Mobile devices and corporate data security', *International Journal of Education and Information Technologies*, vol. 6, no. 1, pp. 97-104.

Miller, K, Voas, J & Hurlburt, G 2012, 'BYOD: Security and Privacy Considerations', *IT Professional*, vol. 14, no. 5, pp. 53-5.

Morrow, B 2012, 'BYOD security challenges: control and protect your most sensitive data', *Network Security*, vol. 2012, no. 12, pp. 5-8.

Nolan, R 1994, 'Note on estimating the value of the IT asset', *Harvard Business School*, vol. 9, pp. 195-7.

Okoli, C & Schabram, K 2010, 'A guide to conducting a systematic literature review of information systems research'.

Prensky, M 2011, 'Digital wisdom and homo sapiens digital', *Deconstructing digital natives. New York and London: Routledge*, pp. 15-29.

Rose, C 2013, 'BYOD: An Examination Of Bring Your Own Device In Business', *Review of Business Information Systems (RBIS)*, vol. 17, no. 2, pp. 65-70.

Scholtz, T 2013, *Gartner Says the Nexus of Forces is Transforming Information Security*, 1, Gartner, India, Goa, <http://www.gartner.com/newsroom/id/2613016

Shulman, A 2010, 'The underground credentials market', *Computer Fraud & Security*, vol. 2010, no. 3, pp. 5-8.

Siponen, M & Willison, R 2009, 'Information security management standards: Problems and solutions', *Information & Management*, vol. 46, no. 5, pp. 267-70.

Smith, DM 2013, 'The Top 10 Strategic Technology Trends for 2014', paper presented to Symposium ITXPO 2013.

Sreenivas, V, Narasimham, C, Subrahmanyam, K & Yellamma, P 2013, 'Performance evaluation of encryption techniques and uploading of encrypted data in cloud', paper presented to 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT).

Sturdevant, C 2011, 'Socializing the enterprise', *eWeek*, vol. 28, no. 1, pp. 34-.

Subashini, S & Kavitha, V 2011, 'A survey on security issues in service delivery models of cloud computing', *Journal of Network and Computer Applications*, vol. 34, no. 1, pp. 1-11.

Suhaimi, A, Goto, Y & Cheng, J 2013, 'An analysis of software supportable tasks in information security management system life cycle processes', paper presented to International Conference on Information and Social Science, Nagoya, Japan.

Tokuyoshi, B 2013, 'The security implications of BYOD', *Network Security*, vol. 2013, no. 4, pp. 12-3.

Von Solms, S 2005, 'Information security governance–compliance management vs operational management', *Computers & Security*, vol. 24, no. 6, pp. 443-7.

Webster, F 2014, *Theories of the information society*, Routledge.

Xiang, Z & Gretzel, U 2010, 'Role of social media in online travel information search', *Tourism management*, vol. 31, no. 2, pp. 179-88.

Xu, X 2012, 'From cloud computing to cloud manufacturing', *Robotics and computer-integrated manufacturing*, vol. 28, no. 1, pp. 75-86.

Yeniman Yildirim, E, Akalp, G, Aytac, S & Bayram, N 2011, 'Factors influencing information security management in small-and medium-sized enterprises: A case study from Turkey', *International Journal of Information Management*, vol. 31, no. 4, pp. 360-5.

Zissis, D & Lekkas, D 2012, 'Addressing cloud computing security issues', *Future Generation Computer Systems*, vol. 28, no. 3, pp. 583-92.

CORE ELEMENTS IN INFORMATION SECURITY ACCOUNTABILITY IN THE CLOUD

Zahir Al-Rashdi, Dr Martin Dick, Dr Ian Storey
School of Business Information Technology and Logistics, RMIT University, Melbourne, Australia
zahir.al-rashdi@rmit.edu.au, martin.dick@rmit.edu.au, ian.storey@rmit.edu.au

Abstract

This paper proposes 9 core elements of information security accountability in the area of cloud computing. The core elements were determined via a series of 18 case studies with Omani government organisations that were actively using and/or providing cloud computing. 36 interviews were conducted and then analysed using a grounded theory methodology As a result of the analysis, responsibility, transparency, assurance, remediation, accountability support environment, flexible change process, collaboration, mechanisms and commitment to external criteria. The research also found that the emphasis on specific core elements is context-dependent and that there was considerable variation in emphasis amongst the case study organisations.

Keywords: Accountability, Cloud Computing, Information Security, Case Study, Grounded Theory

INTRODUCTION

Cloud computing is growing at a dramatic rate (Weins 2017). Such rapid growth over the past decade, combined with the changes cloud computing can cause in the structure and operations of an organisation means information security needs to be more closely examined. Accountability is a core concern for information security in cloud computing, representing most importantly the trust in service relationships between clients and cloud service providers (CSPs). Without evidence of accountability, there will be a lack of trust and confidence in cloud computing by decision makers and it will be considered as an added level of risk, since a client's essential services will be controlled and managed by a third party. The combination of the two factors of significantly increased usage of cloud computing in the last decade and that this involves an outsourcing arrangement raises the need for improved understanding by organisations of many aspects of the cloud computing relationship. Accountability in information security is an important aspect that needs to be examined in a serious manner. Research in this area of information security for the cloud has also been largely technical in nature and management issues such as accountability have not been examined extensively.

When information security accountability is not present, a lack of trust and confidence in cloud computing often develops among business management (Ko et al. 2011; Muppala, Shukla & Patil 2012; Pearson 2013; Rashidi & Movahhedinia 2012), which is then considered an added level of risk (Cayirci 2013; Gellman 2012; Guitart et al. 2013; Morin, Aubert & Gateau 2012; Rajani, Nagasindhu & Saikrishna 2013). Cloud outsourcing renders the process of maintaining data security and privacy, supporting data and service availability, and demonstrating compliance far less transparent (Rajani, Nagasindhu & Saikrishna 2013).

This paper presents a model of the core conceptual elements that determine information security accountability in cloud computing – a primary concern that represents the trust in service relationships between clients and cloud service providers (CSPs) (Pearson & Wainwright 2013).

BACKGROUND

Past research on cloud computing accountability has produced various definitions, embodying different spheres of accountability research. There is a wide variety of views of accountability among academics and practitioners. Accountability in computer science has been referred to as a limited and imprecise requirement met by reporting and auditing mechanisms (Cederquist et al. 2005; Pearson 2011). Yao et al. (2010) considered accountability the way of making the system accountable and trustworthy via the combination of mechanisms. Muppala, Shukla and Patil (2012) defined accountability as accepting ownership and responsibility towards all actions in a standardized manner, regulated by an acknowledged organisation such as the Organisation for Economic Co-operation and Development (OECD) which published privacy guidelines in 1980. In addition, Rush (2010) defined accountability as the reporting and auditing mechanisms that obligate an organisation to be answerable for its actions.

Ko et al. (2011) considered accountability only one component of trust in cloud computing; the others are security mechanisms (e.g. encryption), privacy (protection of confidential data), and auditability. Similarly, the Galway project on privacy regulators and professionals defined accountability as the safeguarding of personal information, acting as a responsible steward and taking responsibility for protecting, managing and appropriately using that information beyond legal requirements, including being held accountable for any misuse (The Centre for Information Policy Leadership 2009).

The Centre for Information Policy Leadership (2011) also identified accountability in relation to privacy as "the acceptance of responsibility for personal information protection. An accountable organisation must have in place appropriate policies and procedures that promote good practices which, taken as a whole, constitute a privacy management program. The outcome is a demonstrable capacity to comply, at a minimum, with applicable privacy laws. Done properly, it should promote trust and confidence on the part of consumers, and thereby enhance competitive and reputational advantages for organisations" (The Centre for Information Policy Leadership 2011, p. 1).

RESEARCH METHODOLOGY

This research builds on earlier work (Al-Rashdi, Dick & Storey 2015), that detailed a comprehensive review of accountability in relation to cloud computing, via two primary literature sources: academic and professional literature; and industrial reports published by high-profile organisations (e.g. Gartner and Microsoft). More than 450 articles, standards and reports were examined, with the literature revealing four core, interrelated elements of accountability: transparency; responsibility; assurance; and remediation. Accountability is not a 'one-size-fits-all' approach, as each organisation has separate needs, such as those with highly sensitive data.

That research was extended using an interpretive qualitative case study approach. This approach was chosen based on the maturity of this research area, as it was felt that a qualitative approach could be used here to obtain an in-depth understanding or 'very rich' picture of information security accountability in cloud computing – a phenomenon common to the qualitative approach (Kvale 1989). Grounded Theory was the chosen research methodology with case studies used for data collection (Eisenhardt 1989). It should be noted that the extended research was carried out independently from the initial research and that only after both research projects had been completed were comparisons of the results made.

This study was conducted via 18 case studies within Omani government organisations that both use and provide cloud computing. 36 staff (senior and middle managers, information security specialists, regulators, and cloud service providers) with over two years of experience in cloud computing were interviewed via open-ended questions. The interviews were conducted during official working hours in the interviewee's offices, with each session lasting about an hour and half. Over 120,000 words were obtained, with all interview transcripts analysed via an inductive and interpretivist qualitative approach. The specific approach was a Grounded Theory (Strauss & Corbin 1998) methodology that was based on Eisenhardt's approach (Eisenhardt 1989) to using Grounded Theory in case studies. This is a highly accepted variant in the area of grounded theory with over 43,000 citations to the relevant paper on Google Scholar. It adapts grounded theory by focusing specifically on case studies and by also incorporating theory in a compatible way. Analysis was conducted using standard grounded theory methods such as open coding, axial coding and constant comparison. This analysis then led to the coding that has been formalised and presented in this paper. NVivo 11 software (QSR_International 2016) was used as the tool for coding the themes that arose.

CORE ELEMENTS OF INFORMATION SECURITY ACCOUNTABILITY

The goal of this research was to understand what an organisation needs to do to achieve information security accountability in a cloud computing context. It should be noted that this is different from achieving information security, as an organisation considered accountable for information security may still have corresponding breaches. Indeed some aspects of information security accountability such as remediation may never come into play if such a breach does not occur.

In order to determine if an organisation is accountable for its information security, the first step is to determine and define the core elements of information security accountability. Previous research (Al-Rashdi, Dick & Storey 2015) used literature analysis to determine the following four core elements of information security accountability – Assurance, Remediation, Responsibility, and Transparency. The case study research undertaken here, independently found that these four elements were considered by respondents to be core elements in achieving information security accountability. The four elements are defined as follows:

- **Assurance** - a well-founded belief that all relevant actors are complying with governance and ethical measurements, and promoting the implementation of practical mechanisms that support them (The Centre for Information Policy Leadership 2010).
- **Remediation** - "the method by which an organisation provides remedies for those whose privacy has been put at risk" (The Centre for Information Policy Leadership 2010, p. 1).
- **Responsibility** - the acknowledgment and assumption of responsibility by relevant actors that they have introduced or have in place appropriate policies and procedures (Ko, Lee & Pearson 2011).
- **Transparency** – the availability and provision of relevant information in a clear and timely manner to the relevant actors (Pearson & Charlesworth 2009).

Figure1 shows the nine core elements and any sub-elements that make up the element. This model arose out of the analysis of the data collected in the 18 case studies.

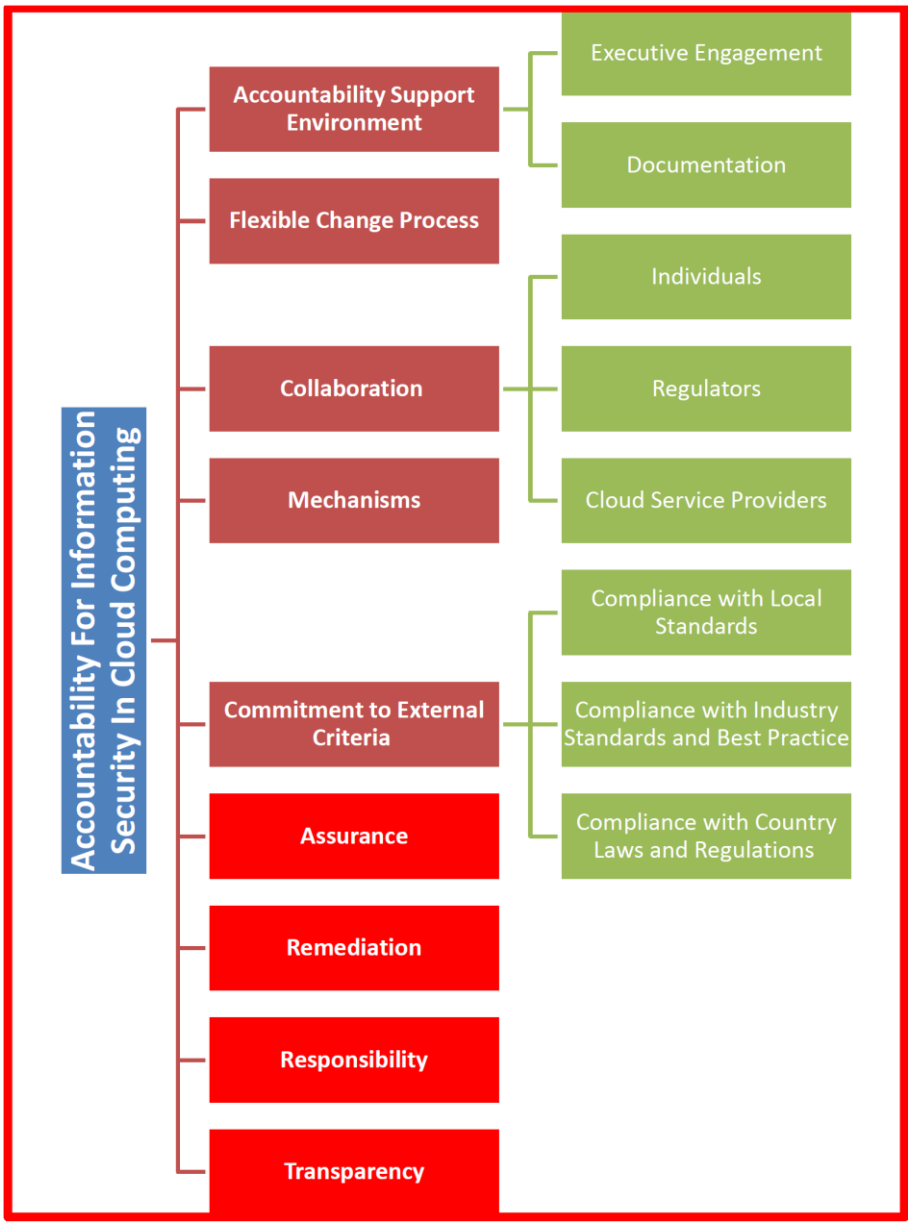


Figure 1: The core elements of information security accountability

As the four core elements of responsibility, transparency, assurance and remediation have been documented individually in the related literature and brought out and described in our earlier work (Al-Rashdi, Dick & Storey 2015) the following sections provide further detail on each of the five new core elements and their sub-elements.

Accountability Support Environment

An Accountability Support Environment in the context of cloud computing is the level of support provided by the organisation in regards to the implementation of accountability and that by providing an environment which supports accounrtability, that the overall level of accountability is increased. The study found that there were two main sub-elements that influenced the level of support of accountability.

- *Executive engagment* is how direct the engagement and follow-up by executives in the development of information security policies. "The executive oversight shows a high level of engagement towards reviewing and re-issuing the implemented policies to match the national or international accepted criteria and country law. This boosts our effort, of information security towards accountability …." Respondents emphasized that fulfilling all of these aspects of compliance, engagement in developing the processes and procedures, and regularly reviewing those policies enhances confidence between CSPs and clients. More confidence then leads to better accountability and minimizes risks.

- *Documentation* is all processes and procedures in information security as a reference for future work, which provides evidence of accountability. A failure to provide documentation leads to significantly reduced accountability according to respondents "It is vital to document all processes and procedures in information security as a reference for future work, which also provides evidence of accountability. Such documentation should keep an up-to-date record of all information security aspects implemented in organizations, with a proper means of implemented solution..".

Flexible Change Process

The respondents were clearly of the opinion that the change process in the context of accountability for cloud computing requires flexibility in relation to altering ICT policies, plans, processes and procedures. The case studies show that the Information Technology Authority supervises the adoption and use of cloud computing by Omani government organisations, but it has worked with them

> *"I can say this terms and conditions applied to change in process, plans and Information Technology Authority policies in the negotiation face within government clients. So overall, that is all based on the outcome revealed from the continued risk assessment used on the customer's site*."

Many respondents emphasized that although ITA policies are comprehensive, plans still need to be tailored to fit specific business needs. Similarly, they pointed out that process changes offered by ITA can allow clients to integrate new protocols that are not addressed by existing ITA policy, e.g. one ministry needing to include health-related security policies into their information security arrangements. An inflexible change policy was considered to not allow organisations to be properly accountable to all their requirements.

A flexible change policy is not without its associated challenges. A typical comment was "*From my experience, there would be some challenges, maybe because of the flexibility in the change process, and there is always conflict between flexibility and security efficiency*."

Collaboration

The third new core element uncovered by the analysis was *Collaboration*. In the context of accountability for information security in the cloud this is primarily about cooperation between the CSP and the organisation. Collaboration ensures that every one of the partners are responsible for their actions and held accountable breaches of regulations. Everything will be transparent and instantly hold a discussion about the non-compliance or breached policies and figure out how to overtake the issue immediately and in cooperation with the client. This ensures the client that the CSP is doing what has been agreed too in the contractual part or assigned SLA. A secondary level of collaboration to achieve accountability is between the organisation and two other entities: regulators and data subjects/and data subjects/individuals.

> *"Compliance and collaboration processes between individuals, business partners and regulators are touching the surface of accountability. In fact, this is part of the selection process that ITA, G-Cloud division is following to filter the government entity according to the list of criteria and standards (local and international) along with applicable law are part of this process."*

By enhancing collaboration, respondents felt that achieving accountability was noticeably easier than situations where collaboration was minimal or non-existent asit built trust between the business partners.

Mechanisms

Mechanisms in the context of information security accountability for cloud computing is the way that cloud service providers assured the clients on how SLA terms and conditions executed in the real world. The assurance on how do the clients trace any breach or non-compliance aspects by ITA as Government Service Provider in Oman. A list of mechanisms has to be established by the organisation and performed in a way that ensures the information security policies and privacy effectively practised and appropriately implemented. The mechanisms might include tools, training and education. The recommended tools might be used to facilitate decision making about appropriate data use and protection, whereas training along with education sessions can be used to identify how to use those tools, and processes. An example of the type of mechanism that needs to be implemented to achieve information security accountability was described by one of the respondents

> *"We do provide the customers with shared service portal. It is a dashboard of the statistics of basically the SLAs and the business process. So it gives the customer that kind of credibility, it tells you very transparently if your SLA and information security policies are right or not. All staff involved in collecting and managing the personal and migrated data is mandatory to commence in training and education session about the use of the portal."*

A failure to put these mechanisms in place undermines information security accountability in the cloud computing context.

Commitment to External Criteria

Commitment to External Criteria in the context of information security accountability for cloud computing is a willingness and capacity to adopt and use organisation-external policies and practices. Such external criteria include local policies (e.g. ITA policies in Oman), industry standards and best practices (e.g. ISO27001, ITIL, CCSK, HIPAA etc.) and government law and regulation (e.g. Oman Electronic Government Architecture Framework (OeGAF) Standard, E-transaction law, privacy law etc.). In the context of information security, failure to commit to these standards indicates a lack of knowledge of the complexity of information security and the inability of any one organisation to be able to deal with the wide and varied threats in the area. As one respondent indicated:

> *"organisations must implement policies connected to appropriate external criteria that can be found in the country law, or industry best practices to protect the information migrated to cloud and privacy of individuals and business partners."*

An organisation that does not adopt and use these external criteria as a reference point is not seen by the respondents as being accountable for information security. Compliance with these external criteria was seen as very important - *"We are not tolerant with vendors/suppliers and individual users for non-compliance."*

THE IMPORTANCE OF CONTEXT

Achieving information security accountability is a complex task for any organisation. The first step is to understand what are the elements that make up information security accountability. It is important to realize that though there are nine elements, the level at which any specific element needs to be driven to is highly context dependent. A level of transparency for one organisation that provides an acceptable level of accountability, may be totally inadequate for another organisation.

It should be noted that the types of organisation that were in the case studies have many similarities. They are all Omani government departments with a high need for data security to protect highly sensitive government data, large complex business processes, and significant infrastructure. This will to some extent hide the requirements for variation in the level of emphasis on specific elements of information security accountability that a more diverse set of case studies might have discovered. But even within this set of case studies, the elements have significant variation. As an example, some of the case studies use private cloud to ensure accountability due to the nature of their requirements, others are able to use private cloud providers while others are satisfied with the government cloud provider.

Overall, it needs to be understood, that the nine element model is not prescriptive and that it must be used in a context-sensitive way that is dependent on the needs of the specific organisation that is attempting to achieve information security accountability.

CONCLUSION

This paper has sought to understand how accountability in cloud computing can be conceptualised. The wide range of existing research into accountability in cloud computing has used a technical approach and has been quantitative, and has generally not addressed the conceptual issues. The enormous growth in moving businesses to cloud computing, mainly due to its flexibility, cost-effectiveness and scalability, and the corresponding absence of a specific cloud computing accountability framework, highlights the growing need for research in this area. Previous research used an extensive analysis of the literature relating to cloud computing and accountability for information security to develop a model of the key conceptual elements (transparency, responsibility, assurance and remediation) relating to this issue. A set of 18 interpretive qualitative case studies were then conducted to examine the situation independently of the literature. A series of interviews were conducted within 38 interviewees from different government entities. Five more key elements were identified from the conducted case studies that are necessary to achieve information security accountability. These were found to be: *accountability support environment*, *flexible change process*, *mechanisms*, *collaboration*, and *commitment to external criteria*.

The findings of this research contribute to the growing awareness of the importance of accountability.It provides an understanding of the importance of the accountability elements that policymakers need to address in cloud computing projects if they wish to be accountable in terms of information security. In future this research is aiming to produce a holistic and heuristic accountability framework that enables a theoretical-based description and analysis of the gap that exists between the ideal and the actuality of the institutional and technical environments of cloud computing implementation in regards to accountability. Finally, this research is seeking to find ways to strengthen the trust and confidence between organisations that have adopted the cloud and CSPs, which in turn strengthen the accountability, which also helps to reduce risks connected to the adoption of cloud computing services.

REFERENCES

Al-Rashdi, Z, Dick, M & Storey, I 2015, 'A Conceptual Framework for Accountability in Cloud Computing Service Provision', *Australasian Conference on Information Systems*.

Cayirci, E 2013, 'A joint trust and risk model for MSaaS mashups', paper presented to Proceedings of the 2013 Winter Simulation Conference: Simulation: Making Decisions in a Complex World.

Cederquist, J, Conn, R, Dekker, M, Etalle, S & den Hartog, J 2005, 'An audit logic for accountability', paper presented to Policies for Distributed Systems and Networks, 2005. Sixth IEEE International Workshop on.

Eisenhardt, KM 1989, 'Building theories from case study research', *Academy of management review*, vol. 14, no. 4, pp. 532-50.

Gellman, R 2012, 'Privacy in the clouds: risks to privacy and confidentiality from cloud computing', paper presented to Proceedings of the World privacy forum.

Guitart, J, Macias, M, Djemame, K, Kirkham, T, Jiang, M & Armstrong, D 2013, 'Risk-driven proactive fault-tolerant operation of iaas providers', paper presented to Cloud Computing Technology and Science (CloudCom), 2013 IEEE 5th International Conference on.

Ko, RK, Jagadpramana, P, Mowbray, M, Pearson, S, Kirchberg, M, Liang, Q & Lee, BS 2011, 'TrustCloud: A framework for accountability and trust in cloud computing', paper presented to Services (SERVICES), 2011 IEEE World Congress on.

Ko, RK, Lee, BS & Pearson, S 2011, 'Towards achieving accountability, auditability and trust in cloud computing', in *Advances in Computing and Communications*, Springer, pp. 432-44.

Kvale, SE 1989, *Issues of validity in qualitative research*, Studentlitteratur.

Morin, J, Aubert, J & Gateau, B 2012, 'Towards cloud computing SLA risk management: issues and challenges', paper presented to System Science (HICSS), 2012 45th Hawaii International Conference on.

Muppala, J, Shukla, D & Patil, S 2012, 'Establishing Trust in Public Clouds', *J Inform Tech Softw Eng*, vol. 2, p. e107.

Pearson, S 2011, 'Towards accountability in the cloud', *Proc. IEEE Internet Computing*, pp. 64-9.

---- 2013, 'Privacy, security and trust in cloud computing', in *Privacy and Security for Cloud Computing*, Springer, pp. 3-42.

Pearson, S & Charlesworth, A 2009, 'Accountability as a way forward for privacy protection in the cloud', in *Cloud computing*, Springer, pp. 131-44.

Pearson, S & Wainwright, N 2013, 'An interdisciplinary approach to accountability for future internet service provision', *International Journal of Trust Management in Computing and Communications*, vol. 1, no. 1, pp. 52-72.

Rajani, B, Nagasindhu, K & Saikrishna, K 2013, 'Integrity Verification & Distributed Accountability in High Performance Distributed Clouds', *International Journal Of Electronics Communication And Computer Engineering*, vol. 4, no. 6.

Rashidi, A & Movahhedinia, N 2012, 'A model for user trust in cloud computing', *International Journal on Cloud Computing: Services and Architecture (IJCCSA)*, vol. 2, no. 2, pp. 1-8.

Rush, B 2010, *The Best Practices Act of 2010 and Other Privacy Legislation,2010*, <https://www.govtrack.us/congress/bills/111/hr5777

Strauss, A & Corbin, J 1998, 'Basics of qualitative research: techniques and procedures for developing grounded theory'.

The Centre for Information Policy Leadership, T 2009, *"Galway Project Plenary session Introduction"* US.

The Centre for Information Policy Leadership , T 2010, *Demonstrating and Measuring Accountability - The Accountability Project – Phase II Paris, France*, France.

---- 2011, *Getting Accountability Right with a Privacy*

Management Program Hunton & Williams LLP, Washington, DC.

Weins, K 2017, 'Cloud Computing Trends: 2017 State of the Cloud Survey'.

Yao, J, Chen, S, Wang, C, Levy, D & Zic, J 2010, 'Accountability as a service for the cloud', paper presented to Services Computing (SCC), 2010 IEEE International Conference on.

AN INVESTIGATION INTO SOME SECURITY ISSUES IN THE DDS MESSAGING PROTOCOL

Thomas White[2], Michael N. Johnstone[1, 2], Matthew Peacock[1, 2]
[1]Security Research Institute, [2]School of Science, Edith Cowan University, Perth, Western Australia
{thomas.white, m.johnstone, m.peacock}@ecu.edu.au

Abstract

The convergence of Operational Technology and Information Technology is driving integration of the Internet of Things and Industrial Control Systems to form the Industrial Internet of Things. Due to the influence of Information Technology, security has become a high priority particularly when implementations expand into critical infrastructure. At present there appears to be minimal research addressing security considerations for industrial systems which implement application layer IoT messaging protocols such as Data Distribution Services (DDS). Simulated IoT devices in a virtual environment using the DDSI-RTPS protocol were used to demonstrate that enumeration of devices is possible by a non-authenticated client in both active and passive mode. Further, modified sequence numbers were found to be a potential denial of service attack, and malicious heartbeat messages were fashioned to be effective at denying receipt of legitimate messages.

Keywords*:* Data Distribution Services, DDS, Critical Infrastructure, Cyber-physical systems, Internet of Things, Network Security

INTRODUCTION

The Internet of Things (IoT) refers to the multitude of interconnected computers, sensors, controllers, and other devices which interact with the physical world. Ubiquitous computing devices are the driving force behind technologies such as smart electrical grids, autonomous cars, wearable health devices and home automation. Evans (2011) made the frequently cited prediction that the number of connected devices would surpass 50 billion by the year 2020, however, revised predictions have forecast the number of devices to be significantly fewer, with Gartner, Inc. (2017) forecasting 20.8 million devices by 2020. Even with a revised prediction this is still a significant number of devices reinforcing the need for robust security considerations.

Potential vulnerabilities in IoT messaging protocols could have serious repercussions if exploited. Whilst in theory industrial networks should be robust, this is not always the case, and the impact of unauthorised access or data modification within these networks could be quite severe. Interruptions or compromise of a power grid by exploiting OPC UA information transfer, intercepting personal health data through a poorly-secured CoAP-based health tracker, or attacking a DDS-based tactical control system in a military vessel are examples of potential attacks which, if successful, could have serious consequences for critical infrastructure.

Physical damage is major concern for industrial cyber-physical systems, but cyber-attacks in general are also creating a significant cost for organisations. IBM and the Ponemon Institute (2016) have stated in their Cost of Data Breach study that the average data breach in Australia comes at a cost to the breached organisation of $2.64 million, at an average cost of $142 per stolen record. Analysing and understanding vulnerabilities in IoT protocols can assist organisations in evaluating how their risk appetite may influence protocol choice when making architectural design decisions.

This research aims to test if identified vulnerabilities that appear to be present in parts of the DDS protocol are realisable. The remainder of the paper describes the security landscape for Industrial IoT systems, defines the experimental methodology used and discusses the findings of the research.

SECURITY ISSUES IN IOT SYSTEMS

Historically, protocol security has been an avenue for exploitation. For example, DNS, FTP, ICMP and EAP are protocols which have had vulnerabilities in their design, rather than programming errors in implementations of the protocols. Even recently ratified protocols such as HTTP/2 have been found to contain vulnerabilities (Imperva, 2016). In addition to common protocols in use on the Internet, continued research has revealed vulnerabilities in control systems protocols, for example BACnet (Peacock & Johnstone, 2014) and DNP3

(Crain & Bratus, 2015) demonstrating that continuing analysis of these protocols can reveal further weaknesses and reinforcing that control systems are a continued focus for security vulnerability analysis.

The security of Industrial Control Systems (ICS) has been viewed as a cause for concern in recent times (Harp & Gregory-Brown, 2016). Many legacy control systems run on standards, protocols and software designed and implemented at a time when the threat landscape was primarily physical based, due to less interconnection between devices. However, in an interconnected world, ICS are gaining attention from cyber adversaries. For example, in 2015 Ukraine's power grid was attacked (Lee, Assante & Conway, 2016) and availability severely compromised after attackers gained access to SCADA systems and shut down parts of the grid. This was one of the first known successful cyber-attacks on power infrastructure, highlighting the growing threat of sophisticated attack operations against cyber-physical infrastructure.

Data Distribution Services or DDS (Object Management Group, 2015) is an open standard primarily intended for peer-to-peer inter-device communications. This protocol defines a data-centric publish/subscribe model and is focussed on low latency communications between devices, rather than between a device and a server or between two servers. The specification defines DDS as:

"... a Data-Centric Publish-Subscribe (DCPS) model for distributed application communication and integration. This specification defines both the Application Interfaces (APIs) and the Communication Semantics (behaviour and quality of service) that enable the efficient delivery of information from information producers to matching consumers." (Object Management Group, 2015, p. 1)

DDS has found uses in many critical environments, such as amongst the energy and aerospace industries, as well as the military. Wang et al. (2008) explored the use of DDS in network-centric operations and warfare systems, demonstrating the increased use of these protocols in environments where security is essential. This is unsurprising as the DDS protocol has broad usage in military applications, having originally been developed by Thales (2015) for use in their TACTICOS Combat Management System. This usage has been one of the primary drivers for the high performance and resilient design requirements of DDS. DDS defers to TLS to provide the bulk of security rather than providing security at the application layer. However, reliance on TLS is clearly not sufficient, given the creation of a standardised post-protocol ratification security specification (aptly named DDS Security). This additional specification provides "authentication, authorization, non-repudiation, confidentiality and integrity" (Object Management Group, 2016) to DDS implementations. He & Liang (2015) have analysed the DDS specification for security issues and put forward a scenario where unauthorised publishers or subscribers may be able to inject data into the DDS network or receive data not intended for the legitimate recipient. They present a high-level overview of theoretical attacks on DDS and it is these types of attacks that DDS Security has been designed to mitigate. Unfortunately, at this point there appears to be limited research on the effectiveness of the DDS Security specification in mitigating the defined theoretical attacks.

Given the range of vulnerable network protocols in use in the IoT, and the associated cost of data breaches; further research is necessary to reduce the attack surface of critical infrastructure installations. The following section describes a series of laboratory experiments undertaken which aims to test a subset of vulnerabilities specific to the DDS protocol.

RESEARCH METHOD

The research was designed as a number of laboratory experiments. A combination of appropriate hardware and software resources were used to attempt to detect, capture, and then analyse specific communication used by a selection of devices using an implementation of the DDS protocol (DDSI-RTPS). The specific research questions were:

1. What risks do vulnerabilities in IoT messaging protocols introduce to IIoT networks and critical infrastructure?
    a. Are there theoretical vulnerabilities present in the Real-Time Publish Subscribe DDS Interoperability Standard protocol specification?
    b. If so, can these vulnerabilities be tested with simulated IoT devices in an isolated environment?

The hypotheses supporting the research questions and experiments designed to test the hypotheses are listed in ***Table 1*** and ***Table 2*** respectively.

Table 1: Hypotheses derived from research questions

Hypotheses
H_1: Enumeration of devices is possible by a non-authenticated client.
H_2: Sequence number and heartbeat messages can be formulated to deny receipt of messages in a DataReader.

Table 2: Experiments designed to test hypotheses

Experiment	**Description**	**Hypothesis**
E_1: Participant Enumeration (Passive**)**	To identify and enumerate RTPS participants on a network	H_1
E_2: Participant Enumeration (Active)	To identify and enumerate RTPS participants on a network	H_1
E_3: Heartbeat Spoofing	To deny receipt of messages to RTPS participants on a network	H_2

Materials:

The virtual lab consisted of four virtual machines, representing devices in the scenario connected by a virtual switch representing a DDS bus. The network topology is shown as ***Figure 1***. All simulation and data collection occurred within an isolated, controlled laboratory environment, therefore the risk of unauthorised access to systems when testing for vulnerabilities was minimised.

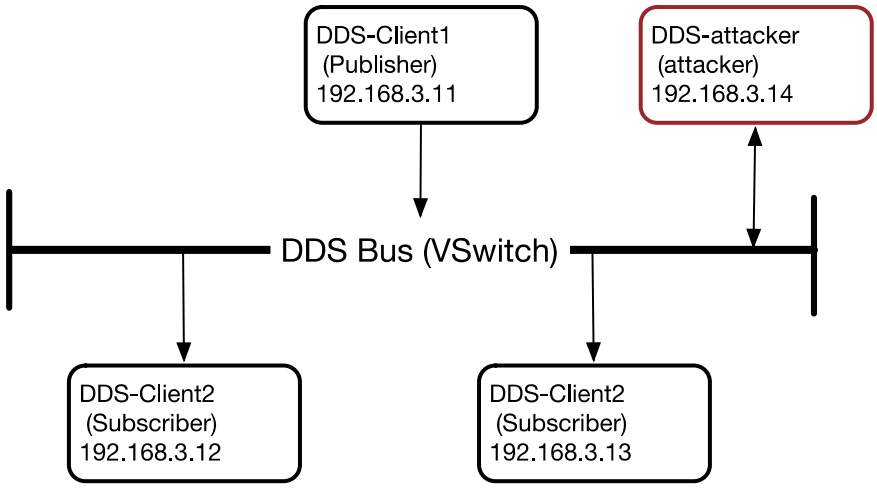


Figure 1: Scenario Network Topology

The virtual machines are listed in ***Table 3***. Note that RTPS participants may generally contain both *DataWriters* and *DataReaders*, which is the case for the attacker virtual machine in research.

Table 3: Virtual machines used in experiments

Hostname	**IPAddress**	**Operating System**	**Purpose**
DDS-client1	192.168.3.11	Ubuntu 17.04	RTPS Participant (Example Publisher)
DDS-client2	192.168.3.12	Ubuntu 17.04	RTPS Participant (Example Subscriber)
DDS-client3	192.168.3.13	Ubuntu 17.04	RTPS Participant (Example Subscriber)
DDS-attacker	192.168.3.14	Kali 2017.1	Testing remote experiments (Attacker)

ANALYSIS AND DISCUSSION

Enumeration

Information gathering is crucial for any attacker when attempting to penetrate a network, and no less so in industrial systems. DDSI-RTPS is reasonably verbose by default, providing reliably identifiable traffic. ***Figure 2*** shows the output of a Python script executed from the attacker, which successfully detects multicast RTPS SPDP packets transmitted on the local network segment as part of E_1. The information that can be obtained from a single SPDP message include: *Host IP address*, *RTPS GUID Prefix*, *RTPS Version*, *vendor ID*, *Time synchronisation information* and the Contents of *Submessages*.

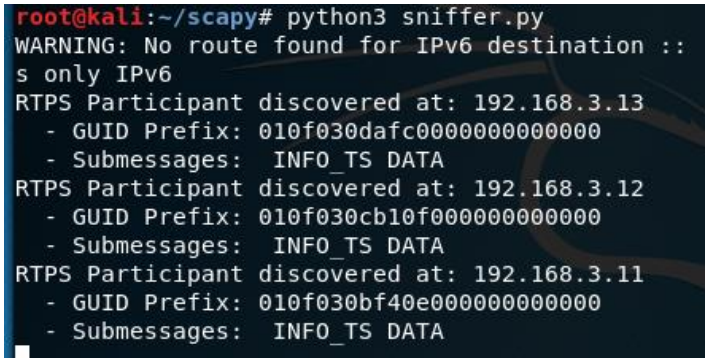


Figure 2: Passive Network Scan and Enumeration Output

In ***Figure 2*** the *Source Address*, *GUID prefix* and overall *Submessages* are displayed. The result of E_1 provides support for H_1 (Enumeration of devices is possible by a non-authenticated client).

The packet capture reconstruction in ***4*** demonstrates the DDSI-RTPS Discovery announcement from the attacker (192.168.3.14) to each scanned address and the associated response. For clarity, only a small range of the network address space was scanned in this simulation (192.168.3.10 - 192.168.3.15).

The packet capture has been colour coded as:

- Yellow indicates legitimate communication between the 3 RTPS participants;
- Red indicates traffic from the attacker; and
- Blue indicates a response to the attacker's discovery message.

Table 4: Active Network Scan and Response Capture

No.	Time	Source	Destination	Protocol	Length	Info
46	12.11818	192.168.3.13	192.168.3.11	RTPS	106	INFO_DST, ACKNACK
47	12.11818	192.168.3.12	192.168.3.11	RTPS	106	INFO_DST, ACKNACK
49	14.00157	192.168.3.11	192.168.3.12	RTPS	154	INFO_TS, DATA, HEARTBEAT
50	14.00157	192.168.3.11	192.168.3.13	RTPS	154	INFO_TS, DATA, HEARTBEAT
51	14.11858	192.168.3.13	192.168.3.11	RTPS	106	INFO_DST, ACKNACK
52	14.11859	192.168.3.12	192.168.3.11	RTPS	106	INFO_DST, ACKNACK
53	14.16125	192.168.3.14	192.168.3.10	RTPS	206	DATA(p)
54	14.1615	192.168.3.12	192.168.3.14	RTPS	270	INFO_TS, DATA(p)
55	14.1615	192.168.3.12	192.168.3.14	RTPS	106	INFO_DST, ACKNACK

Table 4 shows the result of an active scanning and response experiment (E_2) which provides further support for H_1 (Enumeration of devices is possible by a non-authenticated client). Thus H_1 is accepted, given that DDSI-RTPS can provide reliable communications over an unreliable communication medium or best-effort protocols.

Denial of Service

DDSI-RTPS uses *HEARTBEAT* messages sent from a *DataWriter* to a *DataReader* to indicate available sequence numbers on the writer so that the reader can synchronise and determine if any messages are missing. The reader may respond with an *ACKNACK* to indicate to the writer any messages which may be missing, or if the writer has specifically requested a mandatory *ACKNACK* from the reader by setting the *FINAL* flag in the *HEARTBEAT* message.

It was theorised that advancing the sequence number state on the reader may cause the reader to miss legitimate messages if the reader transitioned to a state where it is expecting a higher sequence number than the writer is currently using.
Initial experimentation was conducted through extracting the appropriate DDSI-RTPS *HEARTBEAT* message from a packet capture and modifying the *GUID Prefix*, *entity ID* and *sequence number* fields. With the altered *GUID Prefix* reference implementation, test programs stopped processing once the 'malicious' *HEARTBEAT* messages were sent. The experiment was repeated with varying sequence numbers. Once the legitimate *DataWriter* reached the sequence number provided by the *attacker*, the subscriber would recommence processing messages from the attacker provided sequence number, messages between the last real and attacker provided sequence number are not transmitted.

The specification defines certain conditions in which a *DataReader* must treat a sequence number as invalid and thus the entire *HEARTBEAT submessage* as invalid. These conditions include:

- Negative sequence numbers (The SequenceNumber data structure is signed, however negative sequence numbers are invalid); and
- Last sequence number < first sequence number.

In the conducted experiments, sending a negative sequence number, or sending a sequence number which is lower than the sequence number most recently allocated by the legitimate *DataWriters* had no effect on the processing of messages by the *DataReaders.*

*Table **5*** shows an extract of the packet capture taken during E_3. Once the attacker (192.168.3.14) sends a malicious *HEARTBEAT Submessage* (packet 504), the *DDS-client2* acknowledges the new sequence number (packet 505), then stops responding to the *HEARTBEAT Submessages* from the legitimate *DataWriter (192.168.3.11)*. This result supports H_2 (Sequence number and heartbeat messages can be formulated to deny receipt of messages in a *DataReader*).

Table 5: Network Packet Capture of HEARTBEAT Experiment

No.	Time	Source	Destination	Protocol	Length	Info
490	59.89478	192.168.3.11	192.168.3.13	RTPS	154	INFO_TS, DATA, HEARTBEAT
491	59.89478	192.168.3.11	192.168.3.12	RTPS	154	INFO_TS, DATA, HEARTBEAT
493	59.94067	192.168.3.11	192.168.3.13	RTPS	94	HEARTBEAT
494	59.94067	192.168.3.11	192.168.3.12	RTPS	94	HEARTBEAT
495	60.01186	192.168.3.13	192.168.3.11	RTPS	106	INFO_DST, ACKNACK
496	60.01186	192.168.3.12	192.168.3.11	RTPS	106	INFO_DST, ACKNACK
504	60.82424	192.168.3.14	192.168.3.12	RTPS	94	HEARTBEAT
505	60.94139	192.168.3.12	192.168.3.11	RTPS	110	INFO_DST, ACKNACK
508	61.89489	192.168.3.11	192.168.3.13	RTPS	154	INFO_TS, DATA, HEARTBEAT
509	61.89489	192.168.3.11	192.168.3.12	RTPS	154	INFO_TS, DATA, HEARTBEAT
510	61.98716	192.168.3.11	192.168.3.13	RTPS	94	HEARTBEAT
511	61.98717	192.168.3.11	192.168.3.12	RTPS	94	HEARTBEAT
512	62.01184	192.168.3.13	192.168.3.11	RTPS	106	INFO_DST, ACKNACK

Research question one posited, "What risks do vulnerabilities in IoT messaging protocols introduce to IIoT networks and critical infrastructure?" In relation to DDSI-RTPS, the vulnerabilities introduced could cause significant risk in an industrial control network. Reconnaissance is often the first task undertaken by a cyber adversary, results from E_1 and E_2 show that an attack can passively and actively survey the DDSI-RTPS network to discover all devices running on the bus. Modification of the sequence numbers can result in loss of message transmission between devices on the DDSI-RTPS network. Given the ability to forge malicious *HEARTBEAT* messages, H_2 can be accepted, as a device which has received the malicious packet is prevented from processing further messages. Given that industrial control systems often do not directly employ network-monitoring software, but rather gain system insight via system specific data collection such as Trending or Polling, this type of attack may go unnoticed, or not identified as a cyber-attack for a duration longer than is typical of IoT based networks. With the acceptance of both H_1 and H_2, this paper argues that the introduction of vulnerable IoT messaging protocols into IIoT networks increases the ability of cyber adversaries to undertake reconnaissance of industrial control system networks, and impede the availability of critical systems operating in the network.

CONCLUSION

This research set out to examine security flaws in the DDS protocol (specifically, the Real-Time Publish Subscribe extension). There was theoretical evidence that the protocol could be suborned. The experiments undertaken suggest that the identified theoretical vulnerabilities are present in the Real-Time Publish Subscribe DDS Interoperability Standard protocol specification, answering Research Question 1a. The vulnerabilities were tested with simulated IoT devices in an isolated environment, with acceptance of both H_1 and H_2, answering Research Question 1b affirmatively. The experiments undertaken suggest that enumeration of IIoT devices communicating with DDSI-RTPS is possible by a non-authenticated client in both passive and active mode, respectively. Additionally, modified sequence numbers were found to be largely ineffective at preventing messages from reaching *DataReaders.* However, if a large enough sequence number is provided, in relation to the current sequence number, a denial of service attack is effectively achieved. Additionally, malicious heartbeat messages sent from an attacker device can be crafted to deny receipt of messages between a *DataWriter* and *DataReader*. Given these results, incorporating vulnerable IoT protocols such as DDSI-RTPS into IIoT, which manage critical infrastructure without mitigating the vulnerable protocol increases the risk of cyber adversaries conducting reconnaissance and impeding the availability of critical device-to-device network communication.

REFERENCES

Crain, J. A., & Bratus, S. (2015). Bolt-On Security Extensions for Industrial Control System Protocols: A Case Study of DNP3 SAv5. IEEE Security Privacy, 13(3), 74–79. doi:10.1109/MSP.2015.47

Evans, D. (2011). The Internet of Things: How the Next Evolution of the Internet Is Changing Everything. Retrieved from http://www.cisco.com/c/dam/en_us/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf

Gartner, Inc. (2017). Gartner Says 8.4 Billion Connected. Retrieved from http://www.gartner.com/newsroom/id/3598917Dineen, M., & Cahill, V. (2001). Towards an open architecture for Real-time Traffic Information Management. *Proceedings of the 8th World Congress on Intelligent Transport Systems*. Sydney, Australia.

Harp, D., & Gregory-Brown, B. (2016). SANS 2016 State of ICS Security Survey. SANS. Retrieved from https://www.sans.org/reading-room/whitepapers/analyst/2016-state-ics-security-survey-37067

He, Z. Y., & Liang, Y. (2015). Study on the DDS Network Information Security Technology. Applied Mechanics and Materials; Zurich, 738–739, 1213–1216. doi:10.4028/www.scientific.net/AMM.738-739.1213

IBM, & Ponemon Institute. (2016). 2016 Cost of Data Breach Study: Australia. Retrieved from https://www-03.ibm.com/security/au/data-breach/index.html

Imperva. (2016). HTTP/2: In-depth analysis of the top four flaws of the next generation web protocol. Retrieved from https://www.imperva.com/docs/Imperva_HII_HTTP2.pdf

Lee, R. M., Assante, M. J., & Conway, T. (2016). Analysis of the cyber attack on the Ukrainian power grid. SANS Industrial Control Systems.

Object Management Group. (2015). DDS. Retrieved from http://www.omg.org/spec/DDS/1.4/PDF/

Object Management Group. (2016). DDS Security. Retrieved from http://www.omg.org/spec/DDS-SECURITY/

Peacock, M., & Johnstone, M. (2014). An analysis of security issues in building automation systems. Australian Information Security Management Conference. doi:10.4225/75/57b691dfd9386

Thales. (2015). TACTICOS. Retrieved from https://www.thalesgroup.com/sites/default/files/asset/document/thales_tacticos.pdf

Wang, N., Schmidt, D. C., Hag, H. van't, & Corsaro, A. (2008). Toward an adaptive data distribution service for dynamic large-scale network-centric operation and warfare (NCOW) systems. In MILCOM 2008 - 2008 IEEE Military Communications Conference (pp. 1–7). doi:10.1109/MILCOM.2008.4753364

DECEPTIVE SECURITY BASED ON AUTHENTICATION PROFILING

Andrew Nicholson, Helge Janicke, Andrew Jones, Adeeb Alnajaar
Cyber Technology Institute, De Montfort University, United Kingdom
heljanic@dmu.ac.uk

Abstract

Passwords are broken. Multi-factor Authentication overcomes password insecurities, but its potentials are often not realised. This article presents InSight, a system to actively identify perpetrators by deceitful adaptation of the accessible system resources using Multi-factor Authentication profiles. This approach improves authentication reliability and attributes users by computing trust scores against profiles. Based on this score, certain functionality is locked, unlocked, buffered, or redirected to a deceptive honeypot, which is used for attribution. The novelty of this approach is twofold; a profile-based multi-factor authentication approach that is combined with a gradient, deceptive honeypot.

Keywords: Authentication, Multi-Factor, Deceptive Security, Trust, Honeypot

INTRODUCTION

This paper addresses an aspect of multi-factor-authentication in combining a number of behavioural indicators to provide more trust in the identity of the user. Multi-factor authentication, especially when based on behavioural metrics such as key-stroke recognition provide a level of trust in the identity of the user. Typically such systems use a threshold to either authenticate or deny access to a user. The novelty of the approach is that instead of only authenticating a user, the system is dynamically configuring a deceptive honeypot to include additional attribution techniques to ascertain the identity of the actual user. The result is the InSight system that adjusts dynamically the deployment of deceptive features and attribution techniques based on the trust-level established through the user's interaction with the system.

Weaknesses have long been identified in traditional authentication systems based on username and password (Adams 1999) (Ives 2004) (Schaffer 2011). A typical person is capable of remembering 4 to 5 passwords, however, research by (Sasse 2011) shows that at work a person is likely to need in the region of 15 to 16 different passwords. This creates bad practices, such as writing passwords on post-it notes and reusing the same password between systems. When one system is cracked, all others topple over like a line of dominoes (Ives 2004).

Worse still, people often choose memorable personal passwords, such as their date of birth or the name of their first pet. This information is often shared on social networking websites or can be obtained through social engineering. On the other hand, when people do choose complex passwords, they often are difficult to remember and are consequently recorded in notebooks or on yellow stickers. Even complex passwords remain in many cases trivial to brute force attacks.

Security breaches at large organisations have, in the past, led to password credentials being posted publicly to websites such as pastebin.com or through bittorrent sharing websites. An analysis of these credentials often shows easy to crack login credentials. For example, in one corpus of 32 million credentials, it was found that the average password length was between 6 and 9 characters. Further analysis of the data set showed that the most popular password was ``123456" with 290731 unique occurrences[1].

Adversaries typically use three well known methods to compromise passwords; dictionary attacks, brute force attacks and rainbow tables. Simple passwords such as 'Password' and 'princess' can be compromised by an adversary using a dictionary based attack. A dictionary file contains a list of words which is used to sequentially attempt authentication. Dictionary attacks are fast; using a typical home desktop computer, dictionary words are trivially compromised in a matter of seconds. Dictionaries even exist that replace characters with common substituted character variations (e.g. `p4s\$w0rd`).

[1] http://www.imperva.com/docs/WP_Consumer_Password_Worst_Practices.pdf

Alternatively an adversary may perform a brute force attack to compromise a password by using all possible permutations of available characters. This is a slower attack, but has the potential to crack any password. If an adversary knows something about the password (e.g. it is greater than 8 characters and it ends with a number) then this can be incorporated into the brute force attempt to reduce the effective search space and time required to compromise the password. This is also known as a mask attack.

Adversaries can use rainbow tables, a pre-computed file containing all possible one-way hashes for a given dictionary. However, it is considered unrealistic to generate rainbow tables for all possible salt + password combinations. Recent advancements in graphical processing units (GPU) shows us that passwords can be cracked at a rate of 3.3 billion passwords per second. A common practice in the security of user passwords stored in databases is to use a cryptographic salt. This should be a complex, difficult to guess, string that is added to the user's password before it is encrypted. Therefore a standard dictionary attack or use of rainbow tables is likely to fail since the encrypted one-way hash is a representation of the salt + the password. However, if the adversary knows the salt then they can generate a rainbow table which includes the salt.

A solution to this problem that is in widespread use is multi-factor authentication, which combines the username and password with an additional authentication metric such as a biometric fingerprint scan. The three widely accepted authentication principles for the identification of a user are shown in Table 1.

Table 1: Widely accepted Authentication Principles

Authentication Principle	*Example*
i) Something the User has	Authentication Token, RFID Card
ii) Something the User knows	Password, PIN, Passphrase
iii) Something the User is	Fingerprint, Iris Scan, Voice Pattern

Multi-Factor Authentication

Multi-factor authentication typically uses a combination of authentication principles to establish a user's identity (Nag 2015). For example, the credit card payment system (Kumar 2008) with biometric authentication employs fingerprint verification with a credit card, combining principles i) the card and iii) the fingerprint. Such an approach requires the installation of additional equipment, increasing the cost of the approach and preventing a more widespread adoption. The use of additional devices such as fingerprint readers typically adds to the time taken for authentication which also affects the user acceptance of the technology. Depending on the technology used, fingerprints can also be spoofed (Ihmaidi 2006). Worse, usable fingerprints may readily be available on the credit card itself. Most current approaches to multi-factor authentication are expensive, difficult to deploy and directly affect the usability of the system as they prolong the authentication process (Naji 2011).

A well recognised alternative to fingerprints are keystroke dynamics which can be used as behavioural biometrics for users. It is an analysis technique in which the typing behaviour of users whilst inputting through a keyboard input is monitored (Oppliger 2011). However, if a keystroke is not combined with particular keystroke keys such as the password, it is insufficient to be an objective authentication factor (Teh 2010). Another challenge in this area is that the way users type is very much dependent on the devices they use to enter their credentials. This led to problems in accuracy with this authentication approach when users are able to use a variety of hardware configurations, such as laptops, tablet PCs or smart-phones for data entry. Indeed the use of hardware for authentication has been used since the 1980s. The idea was that users register their devices, e.g. based on their MAC address, so that the devices are authenticated rather than their users.

The authentication module of InSight combines hardware authentication with keystroke recognition to overcome some of their respective problems. To be compatible with traditional authentication approaches InSight extends a simple password mechanism with additional profiling techniques to create a form of multi-factor authentication that is based on hardware and behaviour profiles. Both of these additional factors do not require the user to memorise or otherwise keep any additional secret information.

This also means that InSight does not require special devices to be deployed to end-users, avoiding the impact of additional authentication procedures on usability. InSight integrates profiling information with the established username/password authentication thus discriminating the valid use of password credentials against misuse by establishing a level of trust in the authenticity of the user. This level of trust is then used to drive a deceptive back-end for attribution purposes that adapts the level of deception corresponding to the trust in the authenticity

of the user. The authentication in InSight is based on the trust in the multi-factor assessment, whereas other approaches (Koved 2015) take a risk-based approach.

InSight determines the user behaviour with respect to entering usernames and passwords in correlation with users' hardware. First InSight determines the assumed identity of the user through normal password authentication, gathering in additional information about the hardware configuration and the keystroke behaviour of the user when typing the username and password. The username is then matched against the hardware configurations stored in an associated hardware profile. This compares the ownership and usage patterns of the hardware. Based on the hardware profile the login procedure discriminates between keystroke profiles against which the current login request is evaluated.

Deceptive Security with Honeypots

Honeypots are specially crafted systems that lure adversaries by imitating vulnerable systems, services and software. Honeypots monitor the interaction between the adversary and the system so that the collected data can be analysed by an investigator or automated process. Honeypots are capable of misleading adversaries into revealing information about themselves, by e.g. inadvertently revealing their preferred tools and techniques, coding mistakes and hours of operation. Honeypots are often classified by their fidelity; low or high. Low interaction honeypots, such as Dionaea or Nepenthes, generally simulate a single service and are effective when facing automated scripts such as worms. They are easy to deploy and manage, however they are quickly identified as honeypots by human adversaries, since they offer only limited interactivity. High interaction honeypots are fully fledged operating systems hosted on physical equipment or in a virtual environment. They require high levels of human monitoring and there is an increased risk that they may be compromised and used by an adversary e.g. as a node in a botnet of machines. However, they offer much higher levels of fidelity, such that there is less chance of them being identified as a honeypot. Tools such as Sebek[2] are used to monitor high interaction honeypots and use rootkit techniques to hide deep inside the operating system to avoid detection.

A recent trend in honeypot research has been to combine honeypots with other security technologies. Honeypots have been stitched together with intrusion detection systems, firewalls and host-based security technologies such as anti-virus. While network-based services were once the only strand of honeypot research; a wider variety in the form of wireless, USB, bluetooth, client-based honeypots and honeypots in non-IP networks such as industrial control systems characterise this area of research.

InSight uses the level of trust that is established by the authentication module, to adapt the underlying system's functionality. This approach has the result that the user, or adversary, cannot distinguish the real system functionality from the one that is provided by InSight's adaptive honeypot. The novelty here is to not simply to redirect the user to a honeypot, but to have a continuous gradient between the real system and functions that are not influencing the system with direct and immediate effect.

InSight's Honeypot Trust Model

Deceptive technologies are an unusual but suitable partner to multi-factor authentication systems. This partnership can be justified by challengeing one of the primary principles of authentication: *Upon failing any part of an authentication challenge, a user should be denied access to the system.*

Traditionally, when failing one or more checks in a multi-factor authentication, the actor is not authenticated and consequently denied access to the resources. InSight relaxes this authentication principle using indicators of malice (IOM) to determine a trust-level in the actor's authenticity using a weighted average of the IOMs derived from the Hardware and the Behaviour profile. When an actor attempts to login using a legitimate username/password combination but the device signature and/or the behavioural biometric does not match the profile in the database, then these are considered IOM. Currently InSight only supports Hardware Profiles and Keystroke Recognition as a biometric, but this can easily be extended to include other IOMs as indicated in Figure 1.

As a result of the computed trust score, the actor is placed within a system that employs a gradient from high trust, in which all system functionality is enabled, and with few characteristics of a honeypot, to low trust, in which actions are buffered, and can be rolled-back as well as profiling, traceback and other deceptions enabled.

Figure 1 shows how the honeypot functionality increases as trust is decreased. Trust is decreased as users trigger IOMs indicating that they are potential adversaries that succeeded in obtaining username/password pairs, but

[2] https://projects.honeynet.org/sebek/

without matching the behaviour profiles adequately. To equal proportions, access to the functionality of the real system is minimised as trust is decreased. The partnering of multi-factor authentication with deceptive security in InSight, results in the authentication matrix shown in Table 2.

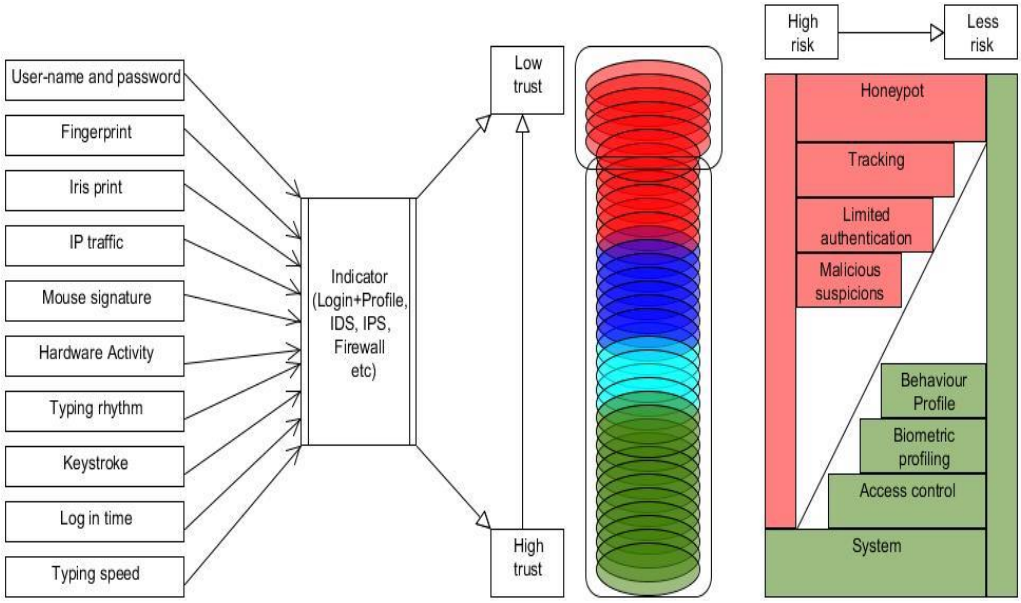


Fig 1: Honeypot Trust Model

Table 2: Multi-factor Authentication with and without InSight

Username Password	Hardware Profile	Behaviour Profile	Without InSight	With InSight
✓	✓	✓	Success, user logs in	Success, user logs in
✓	✓	✗	Denied	Partial success, user logs in to InSight system with a computed trust score
✓	✗	✓	Denied	Partial success, user logs in to InSight system with a computed trust score
✗	✓	✓	Denied	Denied
✗	✗	✗	Denied	Denied

Experiments

Our experiment is based on an online (e-)banking system. E-banking allows for customers to conduct online financial matters, such as transactions and mortgage applications through the Internet. The mid-90s saw the birth of online banking; now its use is widespread. Recent studies identified that 30% of the UK population use online banking. Banks benefit from cost savings as their branches and telephone call centres receive less interaction, customers benefit from convenient and fast access to key financial actions such as transferring funds and paying bills. However, financial fraud is a pressing issue and solutions have been proposed by security researchers (Oppliger 2011). The IC3 reported losses of hundreds of millions of dollars due to online account takeovers and unauthorised funds transfers between 2005-2009.[3]

User authentication plays a critical role as customers typically login using a web or smartphone application. Banks have adopted multi-factor authentication on a widespread scale. One-time PIN hardware devices have been sent to millions of customers free of charge by major UK banks.

In the experiments described in this paper the login server belongs to a fictional high street bank that employs numerous security methods to identify cybercriminals. The experiment is based on synthetic datasets, that were obtained through a 12 student volunteers logging into the system and processing pre-defined tasks. In subsequent experiments this will be expanded to use field data and a real-world application. The prototype experiment s

[3] http://www.ic3.gov/media/2010/100312.aspx

assume that both legitimate and illegitimate users have access to valid username and password credentials. The illegitimate users are assumed to have acquired the credentials via, e.g. compromised machines, by installing a key-logger or rootkit.

InSight grants access to the banking resources provided that they have correctly entered the username and password credentials. Unlike traditional multi-factor authentication, the user is also granted access when the hardware and behavioural profiles are not matched. This results in a lower trust score, representing an Indicator of Malice. Insight will consequently monitor their activities and deploy attribution mechanisms such as honeypots. To show the practicality of the Insight approach the honeypot features allow for delayed execution mechanisms and confirmations through independent channels. For example when money is moved from one account to another, InSight lets them appear to the illegitimate user as having been processed legitimately, thus increasing potentially the interaction with the honeypot. Two alternatives to achieve this behaviour are:

1. Carry out and then later roll-back transactions - Any actions that take place on a compromised account would be rolled back to their original state. In our e-banking scenario, this technique is problematic when funds are transferred to external accounts.

2. Buffer transactions - Actions are buffered for a given amount of time or until a certain condition is met. An example of such is that enough attribution data has been collected. Based on the user's trust-score adjusted over a period of time, Insight can then choose to either process or drop the action. This approach will produce a lag in transactions which a nefarious user may become suspicious of. This approach is deemed to be acceptable as banking transactions already have similar systems in place, causing a similar lag.

Figure 2 shows the simplified banking website from the observation point of a logged in user.

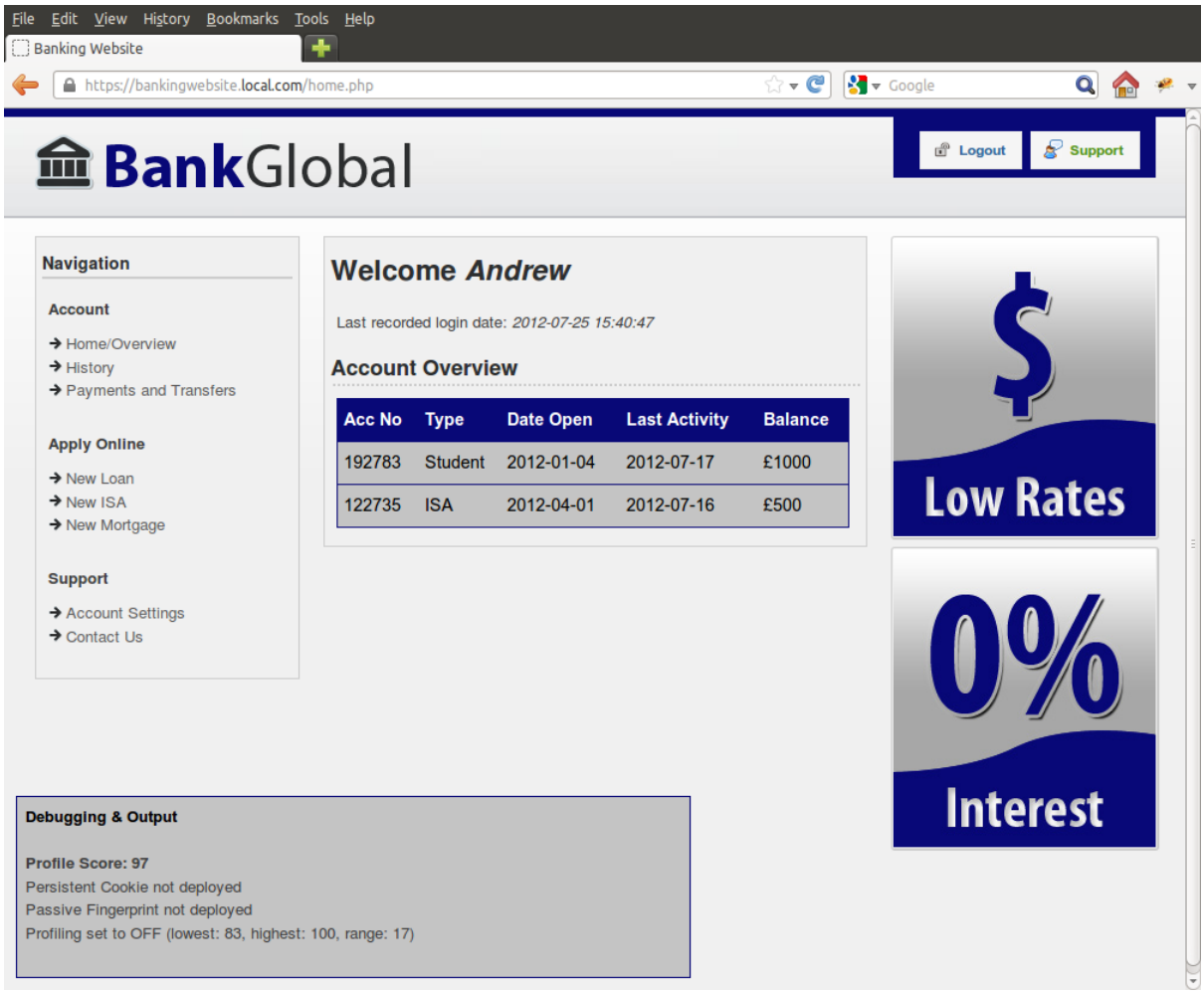


Figure 2: Simplified Banking Website for Experiment

The experiment consists of a number of realistic use cases that represent legitimate and illegitimate users who have obtained valid username/password combinations. The use cases for the experiment areas follows:

Legitimate User:

- Prerequisites: none
- Interaction Flow: Logs In Using Correct Credentials, Common machine and Correct Behavioural Biometric → InSight assesses credentials, common machine and correct behavioural biometric. User is logged in, → User is not monitored, makes required transactions → InSight accepts → User logs out.

Illegitimate User 1 (Local Nefarious User)

- Prerequisites: Acquired correct credentials by installing a keylogger onto a local shared machine (e.g. library). The illegitimate user uses the same machine to login with the stolen credentials. This user has physical access to the machine and therefore has a correct common machine and thus correct device signature.
- Interaction Flow: Logs In Using Correct Credentials, Common machine and Incorrect Behavioural Biometric → InSight assesses credentials are correct, common machine is correct and biometric is incorrect. User is logged in, → User is monitored, makes required transactions → InSight buffers → User logs out.

Illegitimate User 2 (Remote Nefarious User)

- Prerequisites: Acquired correct credentials by compromising a legitimate user's remote machine. The illegitimate user compromised the machine with a phishing email and installed a rootkit which contains a keylogger, which is able to record all interactions (e.g. keystrokes) and covertly send them to a remote server which the illegitimate user controls.
- Interaction Flow: Logs In Using Correct Credentials, Non-Common machine and Incorrect Behavioural Biometric → InSight assesses credentials → correct, common machine → incorrect and biometric → incorrect. User is logged in, → User is monitored, makes required transactions → InSight buffers → User logs out.

RESULTS

We find that within a controlled environment InSight operates as expected. The login screen provides controlled test accounts for each of the three outlined users (legitimate, local nefarious and remote nefarious). Figure 3 showed a successful login to the system by a legitimate user. The debugging and output panel shown in the bottom left corner shows the current trust-level of the user and the deployed deception mechanisms. It is only present during debugging mode and is described further below.

Figure 3: Simplified Backing Website, Experimental Results

Figure 3 shows the web-based debugging and output panel that were used during experiments. The honeypot in this prototype features three features used for attribution: i) persistent cookies; ii) passive fingerprinting; iii) behaviour interaction profiling. The profile score is computed through the use of time-of-use and place-of-use profile combined with a key-stroke metrics to create a biometric. 3A shows the results of a legitimate user login. In this case the output states that no honeypot functionality is deployed. 3B shows the results of the first illegitimate user login. In this case a persistent cookie has been deployed to track the user, however, passive fingerprinting has not been deployed. At this level and below transactions are buffered. Finally 3C shows a low profile score, using correct credentials, but the hardware profile and behavioural characteristics have failed. In this case the user has been assigned a low profile score and all methods of profiling and tracking are enabled.

Figure 3 also demonstrates that the range of the profiling technique can be defined by the operator. This e-banking scenario defined three distinct categories and the window for a trusted user is small; the trust score must be between 83 and 100. For an e-banking website this seems to be appropriate. However, adequate testing and base-lining using real customer data would help to identify a suitable range to minimise false positives and false negatives. The categories and ranges are customisable.

DISCUSSION

The previous section showed that the InSight approach performs well and delivered promising results. This section critiques the approach with respect to a real world implementation. The following are considered: i) costs, ii) false positives, iii) human factors, vi) performance factors and compromise.

Costs

Current multi-factor authentication systems require physical devices, e.g. one-time password (OTP). Newer solutions place the OTP device within the payment card itself, which is convenient for the user, but simply shift the costs into the production of the card.

Other systems such as iris scanners, fingerprint readers and RFID readers also require physical hardware to be present that is unlikely to be owned by a home user. When considering online banking, a high-street bank would need to send millions of these devices to home-users. To place this cost into perspective, personal USB device fingerprint readers currently range from GBP30 to GBP100, while eye scanners are not available for home-use. Eye scanners, such as retina scanners are costly and are invasive to individuals. The use of these devices is common in high security facilities but has not transcended to home use.

The multi-factor authentication system presented in this work does not require any new hardware to be present and therefore avoids the associated costs. Instead, the approach uses something the user already has across multiple devices; a keyboard and a device signature. So a software solution would need to be developed which implements the approach and is available on multiple device platforms.

With regard to the deceptive system, the maintenance costs can be better understood by examining the honeypots. A critical aspect is that honeypots should be maintained by a skilled team; the higher the interaction level the greater the maintenance requirements. The honeypot system for this approach need not have high interactivity, since in fact the user has full use of functionality, but their requests are buffered. The costs of honeypot maintenance may be a barrier to entry for smaller organisations. The key is the integration into the organisations' business processes. If a user is only partially trusted, the additional delay or additional effort in establishing his/her identity in the case of high-value transactions is allowing to dynamically adjust the risks to the organisation against experience of the user.

False Positives

If a user is having a stressful day and their behaviour profile is unusual, they may slip into the honeypot side of the gradient; this is a false positive. This may result in their activities being monitored and their requests, such as banking transactions, buffered. Depending on the system, this may be highly inconvenient. In fact, this result is not problematic two reasons. One, it is anticipated that the user would be contacted by an additional channel (e.g. telephone), as is already standard practice for unusual banking transactions. Two, this increased monitoring is not wasted; it can be used as part of a feedback loop which further improves the system and the novel approach, creating a better baseline of the correct user. With regard to false positives, the approach is similar to intrusion detection systems. Tuning is required and depending on the system in question, an appropriate number of false positives and false negatives should be known from the offset.

Human Factors

The multi-factor authentication technique offers significant benefits in the real world. Humans and passwords are not a particularly good combination (Kovet 2015) (Nag 2015) (Charab 2007) (Kumar 2008). Using InSight the authentication technique is based on a behavioural pattern that is difficult to replicate and is inherent to the individual. Also, while behavioural patterns do change slowly over time, such as a user becoming adept at touch typing, the algorithm can actually account for this.

Performance Factors and Compromise

In our approach there are opportunities for compromise at the client and at the server. At the client side a nefarious user with access to a compromised machine could potentially record a user's keyboard timing data and their hardware details. These could be replayed along with the correct username and password.

The InSight server collects a corpus of behavioural data. If the system is compromised then the unique behavioural records could be exposed. Therefore it is of importance that the system is hardened to deter and prevent compromise. It is also important that behavioural data is encrypted when stored.

Regarding performance of a system that were to implement such an approach, it is likely that there would be an increase in processing and that this would increase the further into the honeypot the user is. Therefore, if the adversary was aware of the system, they might be able to identify the depth at which they are located based on timing data. One solution to this problem is to normalise output to a pre-determined and acceptable speed or duration and add randomness so that output is not 'too similar'.

CONCLUSION

This paper discussed the inherent weaknesses in single-factor authentication systems and current multi-factor authentication systems. InSight was presented and demonstrated to be a novel system that uses behavioural aspects and hardware profiles for multi-factor authentication to compute a trust score. This score is used to determine the users position in a deceptive environment. Insight offers the attractive capability of being able to create a corpus of adversary activities. While the experiment demonstrated the capability within a simulated e-banking environment, the approach could be used in any other system that requires authentication. This data is useful for detecting ongoing attacks using signature matching and is especially useful for attribution, traceback and profiling purposes.

REFERENCES

Adams A. and Sasse M. A. (1999). Users are not the enemy. Communications of the ACM, 42(12): 40–46.

Chakrabarti S. and Singbal M. (2007). Password-based authentication: Preventing dictionary attacks. Computer, 40(6):68–74.

Ihmaidi H., AlJaber A, and Hudaib A. (2006). Securing online shopping using biometric personal authentication and steganography. In Information and Communication Technologies, pages 233–238,.

Ives B., Walsh K. R., and Schneider H. (2004). The domino effect of password reuse. Communications of the ACM, 47(4):75–78.

Koved L. (2015) Usable multi-factor authentication and risk-based authorization. Technical report, DTIC Document.

Kumar D and Kwon D. (2008). A survey on biometric fingerprints: The cardless payment system. In Biometrics and Security Technologies, page 16.

Nag A. K., Roy A., and Dasgupta D (2015). An adaptive approach towards the selection of multi-factor authentication. In Computational Intelligence, 2015 IEEE Symposium Series on, pages 463–472. IEEE.

Naji A. W. (2011). Security improvement of credit card online purchasing system. Scientific Research and Essays, 6(16):3357–3370.

Obaidat M. S. and Sadoun B. (2008). Keystroke dynamics based authentication. Biometrics, pages 213–229.

Oppliger R, Rytz R, and Holderegger T (2009). Internet banking: Client-side attacks and protection mechanisms. Computer, 42(6):27–33.

Sasse M. A., Brostoff S., and Weirich D. (2001). Transforming the 'weakest link'—a human/computer interaction approach to usable and effective security. BT technology journal, 19(3):122–131.

Schaffer K (2011). Are password requirements too difficult? Computer, pages 90–92.

The P.S. (2010). Keystroke dynamics in password authentication enhancement. Expert Systems with Applications, 37(12):8618–8627.

THE CONVERGENCE OF IT AND OT IN CRITICAL INFRASTRUCTURE

Glenn Murray, Michael N. Johnstone and Craig Valli
Security Research Institute, School of Science, Edith Cowan University, Perth, Western Australia
{g.murray, m.johnstone, c.valli}@ecu.edu.au

Abstract

Automation and control systems, such as SCADA (Supervisory Control and Data Acquisition), DCS (Distributed Control Systems) and are often referred to as Operational Technology (OT). These systems are used to monitor and control critical infrastructures such as power, pipelines, water distribution, sewage systems and production control,). Traditionally, these OT systems have had a degree of physical separation from Information Technology (IT) infrastructures. With changing technologies and a drive towards data-driven and remote operations the two technology environments are starting to converge. With this convergence, what was a relatively standalone secure and isolated environment is now connected and accessible via the Internet/cloud. With this interconnection comes the cyber security challenges that are typically associated with only with IT infrastructures. OT data that is then accessible from these environments could include critical information such as pressures, temperatures, proximity levels, control signals and other sensor signals. Due to the aforementioned convergence, OT data and associated control mechanisms are now significantly vulnerable to cyber-attacks. This paper provides an understanding of cyber security in an operational technology context (rather than traditional IT environments) and discusses the underlying causes, vulnerabilities, and the risks that are created by convergence and interconnection. We report on evidence of convergence between IT and OT, and use Hofstede's model of organisational culture to explain the different attitudes and value drivers in IT and OT.

Keywords: Operational Technology, Critical Infrastructure, Cyber-physical systems, Internet of Things, Network Security

INTRODUCTION

Operation Technology (OT) refers to the hardware and software used with the automation controls systems within infrastructure. OT networks and systems including, Industrial Control Systems (ICS) or Supervisory Control and Data Acquisition (SCADA) are used in multiple industries such as power, oil & gas, water treatment, transportation, defence, traffic control and even within private facilities to monitor and control functions such as heating and cooling (Shahzad et al., 2015). These industries form part of our national critical infrastructures without which society and economy would fail. OT systems were designed to integrate data acquisition systems, data collection/transmission systems and Human Machine Interface (HMI) systems to create a centralised control and monitoring solution. Thus, allowing an operator to visually interpret the state of the plant for control and monitoring purposes (Shahzad et al., 2015).

The majority of the OT systems in use are decades old. These were the forerunners of today's 'smart' solutions. These systems are often highly engineered and use proprietary protocols that are specific to project requirements. There are a limited number of people capable of supporting these systems therefore such targets are relatively easy to exploit.

Threats to OT systems are evolving daily with cyber-attacks increasing in both frequency, sophistication and impact. As recently as July 2017, cyber-attackers gained access to a Kansas nuclear power network and other energy companies. The consequences of a cyber-attack on critical infrastructure OT systems goes further than a financial loss to include prolonged outages of critical services (e.g., electricity), possible environmental impacts, and even loss of life.

This research seeks to examine if the differences between IT and OT management structures, cultures and values explain the different attitudes to cyber security. The remainder of the paper describes the security landscape for Industrial IoT (IIoT) systems, describes the key differences between IT and OT systems and discusses the findings of the research.

SECURITY ISSUES IN IIOT (INDUSTRIAL IOT) SYSTEMS

The lack of security in IIoT systems has been a cause for concern recently, as noted by Harp & Gregory-Brown (2016). Many control systems run on standards, protocols and software designed and implemented at a time when the attack surface was small, due to limited interconnection between devices and networks. However, given the (reasonable from a facility manager's point of view) drive for interconnected systems, IIoT systems are gaining attention from cyber adversaries.

Lee, Assante & Conway (2016) point out that in 2015, Ukraine's power grid was attacked and availability severely compromised after attackers gained access to OT systems and shut down parts of the grid. This cyber-attack on power infrastructure, allegedly by a nation-state, indicates the level of sophistication of attacks against critical infrastructure. The attack did not only directly target OT systems controlling the electrical grid but also systems that owners would rely on to respond to the attack e.g., disabling the telephony systems. This outcome is, perhaps, unsurprising as other protocols have also been found to be vulnerable, e.g., BACnet, used in building management systems (Peacock & Johnstone, 2014). Table 1 describes a range of attacks on critical infrastructure that have occurred over the last 35 years.

The vast majority of OT systems are operated in relative isolation from IT systems and infrastructure, which is commonly referred to as 'air gapped', and are simply not designed with a cyber-attack in mind, neither from a detection or defence perspective. OT operating companies use the air gap theory as a barrier, believing that their respective sites are not vulnerable to cyber-attacks. This complacent mindset represents a significant vulnerability as during maintenance periods, where contractors use multimedia devices such as laptops, portable hard drives and USB flash drive as maintenance aids within the industrial infrastructure. The Stuxnet attack was an example of malware was transmitted through an infected USB flash drive allowing the virus to spread resulting in the centrifuges speeding up to the point they self-destructed (Falliere, Murchu, & Chien, 2011).

Contractors may also have set up full-time Internet access, through a cloud infrastructure, to workstations and equipment to enable remote management, including monitoring, technical support and software updates. On December 2014, ICS CERT identified that malware campaign ongoing from 2011, a variant of BlackEnergy malware known as BlackEnergy3, was delivered through Internet connected devices, which compromised industrial Human Machine Interfaces (HMIs). The alert "Ongoing Sophisticated Malware Campaign Compromising ICS" was issued in 2014 and later updated in 2016. The Ukraine power grid outage in 2015, initiated through a spear-phishing campaign targeting IT staff, was attributed to the BlackEnergy malware (Ongoing Sophisticated Malware Campaign Compromising ICS (Update E) | ICS-CERT, 2016).

Both the Stuxnet and Ukraine power grid attacks are examples where human intervention, either deliberate malicious intent or inadvertent, provided the opportunity for cyber-criminals to exploit OT environments. Hence, organisations must consider deploying multi-layer cyber-security defence measures, including implementing cyber technology, educating OT personnel in cyber risk reducing behaviours, introducing cyber resilient policies and physical security.

As OT systems evolve so does the need to apply updates, i.e., software updates and software patches. As explained, within OT systems production is the highest priority or in the CIA structure availability is pivotal. This juxtaposition on prioritisation comparative to traditional IT security has led to some OT systems operating for years without a patch being applied, leaving them highly vulnerable. Furthermore what exacerbates the issues around updates is that many OT equipment suppliers "freeze" the operating systems and subsequent application for elements in an OT system. This practice demands use of arcane and outdated systems.

Finally, when the opportunity does present itself to apply the update/patch, the air gap still exists and requires direct connection to these systems via vendor computers or USB drives, both of which are potential entry points for motivated, capable cyber criminals. The concept, or approach, of using the air gap as a means to avoid/stop a cyber-attack is now flawed and must be addressed in a meaningful way. This *modus operandi* is exactly how the zero day Stuxnet malware was introduced to spread through OT systems.

Date	Cyber-Attack Name	Industry	Location	Description	Effect
Oct 1982	Siberian Pipeline Explosion	Natural Gas	Siberia	Pipeline software programmed to reset pump speed and valve settings above the specifications of the pipeline joints and welds. Allegedly conducted by the CIA.	Explosion visible from space. Vapourised part of the Soviet Union's Trans-Siberian pipeline.
1992	Chevron Emergency System	Oil & Gas	USA	Chevron employee disabled the emergency alert system for 22 states in the USA.	Emergency occurred and no alert was issued.
2000	Maroochy Shire Sewage Spill	Sewage	Australia	Disgruntled employee spoofed controllers opening valves of the sewage system	264,000 raw sewage flooded into hotel and the surrounding parks and river.
2002	Venezuela Pipeline control system	Oil	Venezuela	Allegedly cyber criminals penetrated the SCADA system responsible for tanker loading at a marine terminal.	PLCs operating systems erased. Tankers couldn't be loaded for 8 hours.
2003	Israel Electric Corporation DoS	Electric	Israel	DoS attacks originating from Iran penetrate the Israel Electric Corporation	DoS attacks penetrated however failed to shut down the power grid
Nov 2007	Stuxnet Worm	Nuclear	Iran	Supposedly created by American/Israeli Governments to attack Iran' Nuclear Facilities.	Centrifuges and valves were sabotaged/destroyed
Nov 2011	Pump SCADA	Water	USA	Destroyed a pump remotely from gaining access through a SCADA network. Allegedly the access from gaining user names and passwords from manufacturer's customers.	Chemicals in treatment plant were changed. 2.5 million customers had data exposed to the internet.
Jan 2015	German Steel Mill	Steel Mill	Germany	Cyber criminals used Phishing emails to gain access and prevented the blast-furnace from shutting down.	Catastrophic damage to the steel mill.
Mar 2016	Ukraine Power Grid	Power	Ukraine	Cyber criminals gained access remotely and cut power to 30 substations through the installation of customer firmware.	225,000 customers without power. Deleted files from the master boot records and also shutdown telecommunications.

Table 1: Selected Disclosed OT Cyber Attacks

Possible results that could occur from an OT cyber-attack include:

- A delay in the information that is being communicated to an ICS/DCS/SCADA Master from a Remote Terminal Unit (RTU). These could ultimately lead to a catastrophic event as the information could be turbine speed, level sensor, alarm sensor or actuators.
- Interrupting a connection between an ICS/DCS/SCADA Master to engage an event through a safety system.
- Changing the values received from an ICS/DCS/SCADA Master. This could either have an automatic response, e.g., shutdown a section of a plant or it could lead to a human response which could lead to an inappropriate action.
- Set point values of the RTUs. An example would be changing the HH set point for an alarm for a level sensor for a vessel. Hence causing the vessel to overflow.
- Change/modify the operation of equipment protection systems e.g., speeding up a turbine in a plant causing the blades to be destroyed.

Considering the wide range of vulnerable protocols used in the IIoT, and the high cost of data breaches (e.g., plant re-start costs, declining stock price), further research is necessary to reduce the risk to critical infrastructure. The following section describes the convergence between the IT and OT worlds and highlights the discontinuity caused by widely differing priorities. We do not assert that OT priorities are correct and that IT priorities are faulty in some way, just that they are different and that this is the root cause of a security issue.

THE CONVERGENCE OF IT AND OT

Figure 1 highlights the trend in industry to converge the IT and OT systems in order to access real time data and to interconnect facilities. The convergence is being driven by the need for quantitative management reporting, assisted by 'big data' and sensor technology, artificial intelligence, physical automation, remote operations, cloud computing, analytics. All of which have the potential to enhance productivity/production. To facilitate all of this requires operators to increase network connectivity and access to both IT and OT systems using Ethernet, WI-FI and TCP/IP standards (Shahzad et al., 2016).

By converging IT and OT, systems that were previously closed in many respects are now linked and exposed to all of the risks that have existed in the IT space for years. With this exposure, and not unexpectedly, individuals or groups looking to exploit these new-found vulnerabilities have emerged.

Convergence of itself is not the problem, although figure 1 highlights some obvious differences between the IT and OT worlds. What is a significant problem is the profoundly different priorities between IT and OT that cause a discontinuity in the security space.

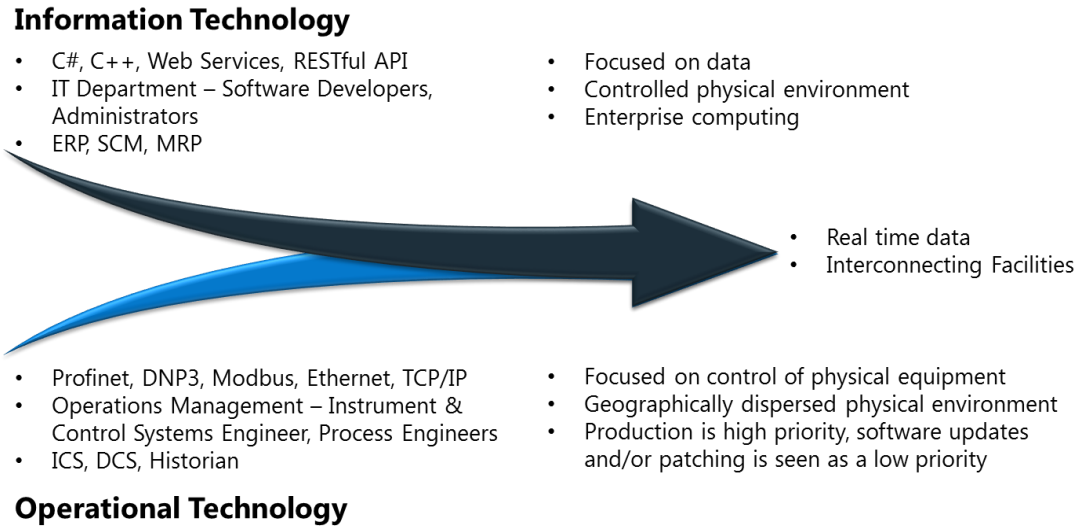


Figure 1: Convergence of IT and OT

As shown in table 2, IT security focuses where the concerns are most often associated with financial integrity, denial of service or loss of information, properties can be grouped and prioritised into confidentiality, integrity and availability, as is common in information security.

1. The confidentiality of data is of paramount importance,
2. The integrity of the data, and;
3. The availability of the data (Zhu, Joseph & Sastry, 2011).

In an OT control system environment, safety and operational risk are critical, which changes the order of security properties to:

1. The availability of the data is paramount to safely maintain production,
2. The integrity of the data, and
3. The confidentiality of the data (Zhu, Joseph & Sastry, 2011).

Table 2: Differences in IT and OT priorities

	Protecting	Priority	Update Frequency	Operating System	Protocols	Cyber Criminal Motivation	Cyber Attack Mission
IT	Data	Confidentiality Integrity Availability	High	Standardised	Standardised	Monetisation	IT Stack Specific
OT	Asset	Availability Integrity Confidentiality	Low	Proprietary	Proprietary	Disruption	Industry Specific

As explained, the priority of IT is to protect data, therefore the IT evolution has seen tools, practices and procedures put in place to protect IT systems from cyber threats.

In the OT space the main priority is to protect the asset base and its associated production. This has translated to minimal effort or changes being undertaken in the OT cyber security space, as production almost certainly would have to be taken offline to accomplish this goal. This production loss, and the associated loss of revenue combined with the cost of designing and implementing the necessary solutions has resulted in OT systems significantly lagging behind IT systems in addressing cyber security threats.

ANALYSIS AND DISCUSSION

In this initial work, based on our combined experience in the IT/OT spheres of at least (conservatively) 60 years, we posited that there is a convergence occurring between IT and OT. There is evidence of that convergence in the IoT area identified by Baig et al. (2017) in terms of the range of disparate IoT networks and systems that are being connected in smart cities-of course critical infrastructure, as managed by OT systems, is one aspect of a smart city. Further, we suggested that security problems might arise from the different priorities encountered in the IT and OT spheres. We assert that these differences are due to elements of organisational culture that have not yet been examined and compared across these two intersecting spheres.

In this section, we use Hofstede's theory of organisational culture as a lens through which to view and explain the differences between IT and OT. Hofstede (1998) perceives culture as 'programming of the mind' where members of one culture can be distinguished from members of another culture. Pertinent to our discussion, Ahmed, et al., (2012) note that when teams from different cultures interact, the complexity of the work relationship can be challenging.

Hofstede's work has been widely cited by many subsequent researchers, but acceptance of his theory and constructs is by no means universal. Probably the most often-cited criticism is that Hofstede studied a single firm, as pointed out by Soares et al. (2007). Nonetheless, a single multi-national firm can exhibit a myriad of cultures. As noted by Jones (2007), Hofstede's survey covered 60,000 employees of a multi-national firm over 50 countries and "the research framework used by Hofstede was based on rigorous design with systematic data collection and coherent theory".

Hofstede's model comprises five dimensions or variables, viz.: Power Distance Index, Individualism/Collectivism, Masculinity/Femininity, Uncertainty Avoidance Index and Long Term Orientation. Of the five, Masculinity/Femininity deals with social gender roles and is therefore not relevant to this analysis. The remaining dimensions are discussed briefly below.

A high power distance index culture accepts the decisions of superiors in a hierarchy without question. Such a culture accepts inequality in the power relationship. Conversely, a low power index culture may operate by consensus and discussion of the decisions of superiors is the acceptable norm. Interestingly, with respect to the deployment of system development methodologies, Iivari and Huisman (2007) found in their study of organisation culture, that IS managers promoted a hierarchical culture that was oriented toward security, order, and routinsation- a stark contrast to managers' take-up and acceptance of agile methodologies.

An individual culture values individual authority and achievement, the right to make self-decision, to hold self-opinion and to exercise some degree of autonomy. In contrast, a collectivist culture values a group's well-being over any individual's desires.

Uncertainty avoidance index is a measure of the extent to which members of a culture feel vulnerable or endangered by uncertainty. In uncertainty avoiding cultures, members are seen to be expressive, and in uncertainty tolerating cultures the expression of thought is repressed. Thus, in the former there is a necessity for consensus.

Long-term orientation is perhaps self-explanatory and refers to a cultural orientation that is not focussed on short-term goals.

Table 3: Core cultural dimensions related to IT vs. OT teams.

Cultural dimensions	**IT**	**OT**
Individualism / Collectivism	Collectivist	Collectivist
Power Distance Index	Low	High
Uncertainty Avoidance Index	Low	High
Long-term Orientation	No	Yes

Our preliminary analysis, summarised in Table 3, shows that IT and OT exhibit widely-differing cultural values across several dimensions. Both IT and OT are classed as collectivist as they both value a team outcome (product, in the case of IT, and production in the case of OT). IT teams are also likely to be using agile development methods, which tend to promote values resulting in a low power differential within a team. Similarly, in IT, changes in technology are embraced swiftly, leading to a low uncertainty avoidance index. This same dimension is paired to a lack of long-term orientation in IT, which is a marked contrast to OT, where the operating technology has not changed for decades (because the focus is availability, as noted in the previous section). This contrast in values, coupled with the difference in priorities described in table 2 leads to the inevitable conclusion that, rather than a peaceful co-existence or a smooth transition to convergence, the clash of cultures will lead to less secure OT systems and the benefits accrued from experiences in the IT world will likely not be realised in the OT world without a substantial readjustment and realignment by both parties.

CONCLUSION

This research set out to examine the state of cyber security in operation technology (OT). As was outlined previously, there is evidence that OT systems have been attacked. We showed that the convergence of IT and OT, whilst inevitable due to other forces, is problematic in terms of cyber security. Additionally, we used Hofstede's organisational culture theory to show the contrast in values between the IT and OT worlds. This variance in cultural values explains the difference in importance placed by each group on information security properties, viz., that OT values availability and that IT values confidentiality.

In future work, we intend to explore both the technical and social dimensions of the problem that we have identified. In the former, we will examine unknown-unknown vulnerabilities of physical and virtual OT systems using a machine learning approach. To address the latter, we will embark on an action research programme to surface and examine the cultural differences in detail and provide a framework for successfully fusing the IT and OT worlds whilst maintaining both availability and confidentiality.

REFERENCES

Ahmed, F., Capretz, L., Bouktif, S., & Campbell, P. (2012). Soft Skills Requirements in Software Development Jobs: a Cross-cultural empirical Study. *Journal of Systems and Information Technology, 14*(1), 58 - 81.

Baig, Z., Szewczyk, P., Valli, C., Rabadia, P., Hannay, P., Chernyshev, M., Johnstone, M., Kerai, P., Ibrahim, A., Sansurooah, K., Syed, N., Peacock, M., (2017). Future Challenges for Smart Cities: Cyber-Security and Digital Forensics. *Digital Investigation*, 22(1), pp. 3-13.

Falliere, N., Murchu, L. O., & Chien, E. (2011). *W32.Stuxnet Dossier*, 1-68. Retrieved from http://www.symantec.com/content/en/us/enterprise/media/security_response/whitepapers/w32_stuxnet_dossier.pdf

Harp, D., & Gregory-Brown, B. (2016). SANS 2016 State of ICS Security Survey. SANS. Retrieved from https://www.sans.org/reading-room/whitepapers/analyst/2016-state-ics-security-survey-37067

Hofstede, G. (1998). Attitudes, Values and Organisational Culture: Disentangling the concepts. *Organisational studies, 19*(3), 477.

Iivari, J. and Huisman, M. (2007). The Relationship between Organizational Culture and the Deployment of Systems Development Methodologies. *MIS Quarterly*, 31(1), pp. 35-58.

Jones, M (2007). Hofstede – Culturally questionable?, Oxford Business & Economics Conference. Oxford, UK, 24-26 June, 2007.

Lee, R. M., Assante, M. J., & Conway, T. (2016). Analysis of the cyber attack on the Ukrainian power grid. SANS Industrial Control Systems.

Ongoing Sophisticated Malware Campaign Compromising ICS (Update E) | ICS-CERT. (2016). Retrieved from https://ics-cert.us-cert.gov/alerts/ICS-ALERT-14-281-01B

Peacock, M., & Johnstone, M. (2014). An analysis of security issues in building automation systems. Australian Information Security Management Conference. doi:10.4225/75/57b691dfd9386

Shahzad, A., Lee, M., Xiong, N., Jeong, G., Lee, Y., Choi, J., … Ahmad, I. (2016). A Secure, Intelligent, and Smart-Sensing Approach for Industrial System Automation and Transmission over Unsecured Wireless Networks. *Sensors*, *16*(3), 322. doi:10.3390/s16030322

Soares, A. M., Farhangmehr, M., & Shoham, A. (2007). Hofstede's dimensions of culture in international marketing studies. *Journal of Business Research, 60*(3), 277.

Zhu, B., Joseph, A., & Sastry, S. (2011). A Taxonomy of Cyber Attacks on SCADA Systems. *2011 International Conference on Internet of Things and 4th International Conference on Cyber, Physical and Social Computing*. doi:10.1109/ithings/cpscom.2011.34