Journal of Information Warfare

Volume 15, Issue 3

Contents

Authors

Dr. Jim Q. Chen is professor of Systems Management and Cybersecurity in the iCollege at the U.S. National Defense University (NDU). His expertise is in cybersecurity technology and cybersecurity strategy. He is a recognized cybersecurity expert.

Tiago Cruz is an assistant professor in the Department of Informatics Engineering at the University of Coimbra. He also serves as a senior researcher at the Centre for Informatics and Systems of the UC. His research interests include management of communications infrastructures and services, critical infrastructure security, and network function virtualization.

J. S. Hurley is the course manager for Cyberspace Strategies and co-manager of the Critical Infrastructure Protection Laboratory at the National Defense University (NDU). He has also worked as senior manager of Distributed Computing at the Boeing Company, directed three research centers, and served as the co-director of the Army Center of Excellence. He is also a 2014-2015 Seminar XXI Fellow.

Dr. Andre Karamanian is a consulting solutions architect at Cisco Systems, where he consults for Fortune 500 and enterprise clients. He is the author of *PKI Uncovered: Certificate-Based Security Solutions for Next-Generation Networks*. He periodically speaks at Networkers and Cisco Live. He completed his doctoral dissertation at Capitol College in information assurance and has earned a dozen industry certifications including the CCIE and CISSP. He enjoys finding synergies between different areas of research and fields of study.

Marc M. Kolenko is a solutions-oriented Cyber Defense and Information Systems Security Engineering Professional with more than 30 years of notable success directing a broad range of Enterprise IT initiatives in both the private sector and government. He is currently the Senior Cyber Security Solutions Architect and Strategist at Information Innovators, Inc. (Springfield, VA). He is responsible for delivering Computer Network Defense (CND), Continuous Monitoring, Security Operations, Cyber Threat Intelligence, Information Assurance (IA) & Compliance, and Systems Security Engineering solutions that aid clients with meeting federal government cybersecurity mandates (i.e., RMF/FISMA, FedRAMP, ICD 503, CNSS 1253, and the Comprehensive National Cyber Security Initiative) through technology.

Christine MacNulty, FRSA, is the CEO of Applied Futures, Inc., which specializes in strategy, strategic planning and change, and understanding cultures. For the last twenty years, she has been a consultant for the Department of Defense and NATO. She has also worked with many Fortune Global 500 companies. She is the co-author of two books and a speaker at many conferences.

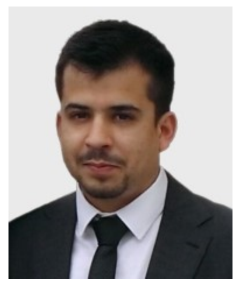
Edmundo Monteiro is a full professor at the University of Coimbra, Portugal, where he earned a doctorate in Electrical Engineering in 1996 and the Habilitation in Informatics Engineering in 2007. His research interests include computer networks, wireless communications, quality of service and experience, service oriented infrastructures, and security. He is the author of more than 200 publications including books, journals, book chapters, and has presented at international conferences. He is also co-author of nine international patents. He participated in many

European initiatives and projects. He is an editorial board member of *Computer Communications* and *Computer Networks*, and is involved in the organization of national and international conferences and workshops. He regularly serves as a reviewer of Portuguese and European projects. He is the Portuguese representative in IFIP-TC6, and senior member of IEEE Communication Society, and ACM Special Interest Group on Communications.

Njabulo Mkhonto is a researcher and software developer for the Cyber Defense team at the Council for Scientific and Industrial Research (CSIR). He has an interest in the applications of Artificial Intelligence research and techniques in solving real-world problems. He studied at the University of Johannesburg where he completed his BSc and BSc Hons in Information Technology, focusing on the use of swarming technologies for improved image processing. Since joining the CSIR, his focus has been on cyber security, where he has been involved in research efforts involving cyber threat intelligence, mobile security, and network security.

Dr. Jabu Mtsweni is a Research Group Leader for the Cyber Defense team at the Council for Industrial and Scientific Research (CSIR). He has research interests and technical expertise in Internet of services, software and firmware reverse engineering, malware analysis, threat intelligence, web security and general cyber warfare. He has more than 13 years of academic and industry experience and has published more than 38 peer-reviewed conference and journal papers/articles in both local and international forums. He has also publicly presented and actively contributed at various technology forums over the years, including the ITWeb Security Summit, TEDx, the SADC Cybersecurity Conference, IST-Africa, the South African Institute for Computer Scientists, the International Conference on Cyber Warfare and Security, and the Information Technologists and International Information Security South Africa Conference. He is a co-organizer of Random Hacks of Kindness (Pretoria) and a member of the Suganang Foundation, focusing of human capital and capacity development in the ICT space.

Muyowa Mutemwa is a Cyber Security Researcher for the Cyber Defence team at the Council for Industrial and Scientific Research (CSIR). He has research interests in Platform, Application and Network Security. He completed his master's degree in computer science at the University of the Western Cape with a specific focus in Information Communication Technologies for rural developments. He previously worked for Telkom SA as a Data Centre, Network Strategy Architect.



Jorge Proença is a PhD student in Information Science and Technology at the University of Coimbra. He earned his M.Sc. degree from the same institution in 2012. Since 2012 he has been a junior researcher in the Centre for Informatics and Systems of the University of Coimbra (CISUC), where he participates in research projects in the fields of network virtualization, security, and critical infrastructure protection.

Rui Queiroz is an M.Sc. student in the Department of Informatics Engineering at the University of Coimbra. He also serves as a research student at the Centre for Informatics and Systems of the UC. His research interests include management of communications infrastructures and critical infrastructure security.

Daniel J Ryan is a lawyer in private practice, an author, and an educator teaching cyberlaw and information security as an adjunct professor at George Washington University. He previously served as a faculty member at the National Defense University.

Julie JCH Ryan, D.Sc., is a professor at the National Defense University, teaching in the areas of cyber security and information assurance. Her service in academia follows a career in the private sector and service as a U. S. government civilian and as a U. S. Air Force officer. She has published three books – *Defending Your Digital Assets Against Hackers, Crackers, Spies and Thieves*, *Leading Issues in Information Warfare and Security*, *Detecting and Combatting Malicious Email*.

Dr. Char Sample is currently a visiting research fellow at the University of Warwick and is employed as a research fellow for ICF International at the Army Research Labs in Maryland. She has more than 20 years of experience in the information security industry, including roles as a developer, integrator, architect, product designer, and researcher. Most recently she has been researching the role of national culture in cyber security events. Presently she is continuing research on modelling cyber behaviours by culture, metrics, risk quantification, and modelling of other security issues.

Paulo Simões is a tenured assistant professor in the Department of Informatics Engineering at the University of Coimbra, Portugal, where he earned his doctorate in 2002. He regularly collaborates with Instituto Pedro Nunes as a senior consultant, leading technology transfer projects for industry partners such as telecommunications operators and energy utilities. His research interests include Future Internet, network and infrastructure management, security, critical infrastructure protection and virtualization of networking, and computing resources. He has more than 150 publications in refereed journals and conferences and is a member of the IEEE Communications Society.

Lanier Watkins is currently a Senior Professional Staff II member of the Asymmetric Operations Sector of the Johns Hopkins University Applied Physics Laboratory (JHU/APL) and an Associate Research Scientist at the JHU Information Security Institute. Prior to joining APL, he served as a senior engineer and product manager at the Ford Motor Company and AT&T**.**

Murdoch Watney is a professor in the Department of Public Law and head of the Private Law Department at the University of Johannesburg, South Africa. She holds the degrees BA, LLB, LLM (RAU), LLM (UNISA) and LLD. She previously worked as a prosecutor and is an admitted advocate of the High Court of South Africa. She has contributed to three textbooks and has published extensively in law journals – both nationally and internationally – on the law of criminal procedure, criminal law, law of evidence, and cyber law. Most of her research focuses on cyber law. She has delivered peer-reviewed papers at national and international conferences.

Shuang Xie is a software engineer at Alpine Electronics Research of America, Inc. He earned a bachelor's degree in Security Informatics from the Shanghai Jiao Tong University, Shanghai, China, in 2012, and a master's degree in Security Informatics from the Johns Hopkins University, Baltimore, Maryland, in 2014.

Tianning Yang is a security engineer at Nav Technologies, designing security schemes and protecting confidential data related to credit reports and personally identifiable information. She earned her bachelor's degree in Information Security and an LLB in Law from Nankai University in 2012, and a master's degree in Security Informatics from Johns Hopkins University in 2014.

Development of a Cyber-Threat Intelligence-Sharing Model from Big Data Sources

J Mtsweni, M Mutemwa, N Mkhonto

Council of Scientific and Industrial Research (CSIR)
Defence, Peace, Safety, and Security
Pretoria, South Africa
E-mail: jmtsweni@csir.co.za; mmutemwa@csir.co.za; nmkhonto@csir.co.za

Abstract: *As data in cyberspace continues to grow because of the ubiquity of Information Communication Technologies (ICT), it is becoming challenging to obtain context-aware, actionable information from Big Data to timely detect and respond to cyberattacks that are increasing in severity, complexity, and frequency. In fact, cybercriminals are developing and sharing advanced techniques for their cyber espionage, reconnaissance missions, and ultimately devastating attacks. In order to reduce cybersecurity risks and strengthen cyber resilience, strategic cybersecurity information-sharing is a necessity. This article discusses one way of handling large volumes of unstructured data that have been generated by multiple sources across different sectors into a cyber-threat intelligence-sharing model.*

Keywords: *Cybersecurity, Cyber-threat Intelligence, Cyber Intelligence, Crowdsourcing, Big Data, Web Security, Vulnerabilities*

Introduction

Public sectors, private sectors, and individuals continue to rely on the Internet for effective information sharing and communication. Key to the effectiveness of the Internet for these three stakeholders is the exponential growth in the ubiquity of Information Communications Technologies (ICT), with Gartner reporting that the growth in the number of ICT devices connected to the Internet will grow by 35% in 2016 and will reach a total of 6.4 billion connected devices in 2020 (Meulen 2015). The sheer large number of ICT devices on the Internet has become instrumental in enabling these key stakeholders to be prosumers—that is, both producers and consumers of large information on the Internet. This has inadvertently led to what is commonly called Big Data: “extremely massive and highly complex data sets of information” that are continuously increasing in volume, velocity, and variety (Khan *et al.* 2014; Zikopoulos & Eaton 2011). Although Big Data presents various opportunities for organisations (Kaisler *et al*. 2013), it also introduces numerous challenges, such as difficulties in collecting, storing, processing, and analysing all this data in order to act upon it or extract value in a timely manner (Katal, Wazid & Goudar 2013; Zikopoulos & Eaton 2011). The timely processing of Big Data is necessary to provide the context-aware, actionable information that is required to remedy vulnerabilities affecting a large number ICT devices connected to the Internet, remedies without which a major threat is presented to the users of these ICT devices in the form of cybercrimes.

ISSN 1445-3312 print/ISSN 1445-3347 online

In cyberspace, threats and attacks continue to increase in number and complexities (Mtsweni *et al*. 2016). As such, cybersecurity threat intelligence is gaining prominence, mainly to enable users to collect Big Data that will allow them to recognize, understand, and protect themselves against sophisticated cyber adversaries and vulnerabilities that are reported on a daily basis. However, organisations are also finding it increasingly challenging to adequately tap into security-related Big Data and implement appropriate solutions for the exposed vulnerabilities or imminent threats. The main reason for this is that most organisations still operate in isolation when it comes to gathering cybersecurity intelligence. However, it is apparent that no one organisation can act on all the security-related Big Data alone.

Notably, in the developed world, large organisations are already collecting and sharing threat intelligence in order to protect themselves from emerging threats and attacks (Brown, Gommers & Serrano 2015). Powerful governments, such as in the United States and United Kingdom, are already enacting legislation that attempts to encourage cybersecurity information sharing between the government and private sectors (Fransen, Smulders & Kerkdijk 2015; Ring 2014). At the same time, large organisations across the world already have security teams gathering large security datasets in order to understand the current threats and protect themselves from imminent cyberattacks. Computer Incidents Response Teams (CSIRTs) are also common in many countries for receiving, reviewing, and responding to computer security incident reports and activities (CERT.org 2016).

Nevertheless, cybersecurity threat intelligence or information sharing is still emerging and possibly immature; thus, many nations, particularly developing nations, are lagging behind. In addition, relevant use cases and models that could encourage cybersecurity information sharing are limited. Existing solutions for sharing cybersecurity information are mostly commercial, and most lack the necessary semantics, intelligence, and visualizations necessary for sharing actionable, reliable, context-aware, and timely information. CSIRTs are mostly reactive and do not have foresight capabilities; moreover, they tend to focus on a multitude of vulnerabilities and incidents, which are not necessarily relevant to every organisation within a specific domain. This article presents a preliminary cyber-threat intelligence-sharing model that could be used by collaborating and trusted stakeholders to share cybersecurity information that might make it possible to limit and/or prevent cyberattacks more quickly.

Background

In the fast-moving domain of cybersecurity, "receiving the right information at the right time" is vital in reducing security risks, deterring attackers, and improving the security posture of an organisation (Goodwin & Nicholas 2015). Hence, making effective use of cyber-threat intelligence is an important component of any organisation's cybersecurity strategy. In government and military environments, intelligence is a well-understood concept and involves the collection, analysis, and interpretation of information for battlespace awareness (Waltz 1998) and, eventually, for decision-making purposes (for example, whether to defend or attack). This concept is also gaining ground within the cybersecurity space, chiefly because software vulnerabilities, threats, and attacks are becoming more complex, severe, and dynamic. Threats are changing on a daily basis and so should the solutions. Indeed, intelligence and continuous awareness of software vulnerabilities, cyber threats, and attacks that face individuals and organisations on a daily basis are essential for mission accomplishment in cyberspace (Polancich 2014).

Before defining threat intelligence, this article makes a clear distinction between security information and cyber-threat intelligence. **Table 1**, below, highlights some of the differences. As may be noted in **Table 1**, ordinary security information, such as that found in common-vulnerabilities' databases, might be unstructured and might make it a challenge for organisations to act upon information in a timely manner. In essence, Big Data can be classified as a large set of raw information, whereas cybersecurity intelligence is information that is possibly structured, relevant, actionable, and well-timed *and* that enables organisations to achieve business goals (for example, to secure information assets).

Security Information	Cyber-Threat Intelligence
Is structured, unstructured, raw (general), unfiltered information.	Is structured, relevant, sorted, and processed information.
Is aggregated from virtually every source.	Is reliably aggregated and correlated for accuracy.
May be true, false, misleading, incomplete, relevant, or irrelevant.	Is accurate, timely, complete (as possible), assessed for relevancy.
Is not actionable.	Is actionable.

Table 1: Security information vs cyber-threat intelligence (Mishra 2014)

Although cybersecurity intelligence is an emerging discipline, it is fairly defined (Brown, Gommers & Serrano 2015; Ring 2014). It is often referred to as threat intelligence, intelligence-driven information security, cyber intelligence, or cyber-threat intelligence (Eom 2014; Ring 2014; Mishra 2014; Goodwin & Nicholas 2015). For the purposes of this discussion, these terminologies are used interchangeably to loosely refer to the collection and analysis of cybersecurity vulnerabilities, threats, incidents, and Indicators of Compromise (IOCs), such as malicious IPs and URLS, tactics, techniques and procedures (TTPs), and recent attacks.

According to Eom (2014), "cyber intelligence refers to the collection, processing, analysis, integration, evaluation, and interpretation of data concerning hostile cyber organisation, cyber forces capabilities, network systems, hardware, software, threats and vulnerabilities". In addition, Eom also defines threat intelligence as

> evidence-based knowledge, including context mechanisms, indicators and actionable advice about an existing or emerging menace or hazard to assets that can be used to inform decisions regarding the subject's response to that menace or hazard. (2014)

It is important to highlight the fact that threat intelligence is about the discovery of vulnerabilities and threats before attacks are performed. Hence, Farnham (2013) maintains that intelligence is actionable information. Threat intelligence could also be classified as strategic, operational, or tactical (Eom 2014). For the current discussion, the focus is on tactical intelligence, which deals with information such as incidents, threats, vulnerabilities, TTPs, and IOCs (Goodwin & Nicholas 2015).

According to Ring (2014), threat intelligence is an expensive exercise, which only a few big organisations can afford to invest in on their own. As such, other small organisations are priced out of the threat-intelligence market. Nevertheless, different types of threat-intelligence collection mechanisms are emerging, including Open Source Intelligence, Cyber Space Intelligence, and Human Intelligence.

At the same time, there are a variety of cybersecurity threat-intelligence sources, mainly classified as internal or external and categorized as public or private. Presently, the manner in which threat intelligence is gathered and acted upon is solely based on individual organisations. The challenge with this approach (the lack of collaborations amongst organisations and governments) is well recognized (Ring 2014). The insights shared in this paper contribute to this space of collaborative and coordinated tactical cybersecurity threat intelligence. The authors posit that securing oneself against complex cybersecurity threats and attacks calls for a shared and a coordinated approach that involves a number of stakeholders, such as the owners, developers, and users of the systems that are susceptible to cyberattacks.

There are different approaches that could be adopted to exchange cybersecurity information. According to Goodwin & Nicholas (2015), some of them could be 1) voluntary exchange, 2) mandatory disclosure, 3) formalized exchange, 4) security-clearance-based exchange, 5) trust-based exchange, and 6) ad-hoc exchange. For the purposes of this article, trust-based, ad-hoc, and voluntary exchange approaches are adopted. These approaches will be further elaborated upon in subsequent sections.

The Rise of Big Data

The importance of Big Data in the cybersecurity space cannot be overemphasized. In countries such as the United States, it is even viewed as a national challenge and priority, along with health care and national security. Big Data presents a number of interesting challenges for cyber intelligence—for example, quality of data, privacy, and security (Katal, Wazid & Goudar 2013). These challenges are compounded by the fact that Big Data is generated on a daily basis through various media. This fact makes it impossible for one organisation to store and to process in a timely manner such data for decision-making purposes. Indicators of Compromise (IOCs) (for example, cyber threats, vulnerabilities, viruses, and malicious sites) contribute to the scale of Big Data. These IOCs are released on a daily basis by individuals, as well as by private and public organisations. However, organisations are unable to act on these IOCs in a timely manner for a number of reasons. One of these reasons is that organisations still rely heavily on traditional systems. Traditional systems suffer when it comes to dealing with Big Data; as such, new ways of dealing with Big Data, especially within the context of cybersecurity threat intelligence, are needed.

Cybersecurity Threat-intelligence Exchange Platforms

As touched on in the introduction, there are a couple of commercial and open-source threat-intelligence platforms that exist on the Web and in other closed environments. These platforms vary in terms of their features and target market. The following sections present a systematic literature review of these platforms and provide context for the combined platform developed by the authors as a Proof of Concept (POC).

- **Malware Information Sharing Platform (MISP)**
  MISP is a platform for sharing, storing, and correlating Indicators of Compromises of targeted attacks by allowing organisations to share information about malware and their indicators (CIRCL 2016). MISP is a web-based tool using a REST API to send and receive data. The MISP allows for storing of technical and non-technical information about malware and attacks, for tracing correlations between malwares, for storing data in a structured format, for exporting and importing in various formats, and for data-sharing with other parties and trust groups using MISP and STIX support to export data in STIX format (Barnum 2014).

- **AbuseHelper**
  This is an open-source project that is used to automatically process incident notifications. This tool is developed for Computer Emergency Response Teams (CERTs) and Internet Service Providers (ISPs) to help them in their daily jobs of following and treating a wide range of high-volume information sources (AbuseHelper 2011).

- **IntelMQ**
  IntelMQ is a solution that was designed for CERTS to collect and process security feeds, pastebins, and tweets using a message queue protocol. Its main goal is to give incident responders an easy way to collect and process threat intelligence to improve the incident-handling processes of CERTS (IntelMQ 2015).

- **Cyber Threat XChange (CTX)**
  The Cyber Threat XChange (CTX) is a component of the HITRUST Alliance Cyber Threat Intelligence and Incident Coordination Centre (C3), which was created to detect and respond to cyber threats that are targeting the healthcare industry (HITRUSTAlliance 2016). CTX collects and analyses the cyber threats and distributes actionable indicators in electronically consumable formats that organisations can utilize to improve their cyber defences (HITRUSTAlliance 2016).

- **Open Threat Exchange (OTX)**
  This platform is an open-threat information-sharing and information-analysis network that is created to put effective security measures within the reach of all organisations (AlienVault 2014). OTX provides real-time, actionable information. The information shared is anonymized and shared with the AlienVault community.

- **Soltra**
  Soltra is a commercial cyber-threat intelligence-sharing platform. It integrates well with various other systems and is capable of pulling security data from disparate sources. It de-duplicates the data and routes intelligence to users, devices, or communities in real time (Soltra 2016).

- **Collaborative Research Into Threats (CRITS)**
  CRITS is an open-source malware and threat repository that uses other available open-source software to enable users to create incidents and share them with others (Goffin 2014). It can also be used by analysts and security experts to defend against malware and

cyber threats. CRITS data is converted to CybOX objects, packaged within STIX documents. As such, CRITS uses STIX as a common standard to convey the full range of cyber-threat information.

- **Trusted Automated eXchange of Indicator Information (TAXII)**
  TAXII is the preferred method of exchanging information represented using the STIX language (MITRECorporation 2016). CRITS can also be accessed locally, remotely, or via custom APIs. This article explores all the different access mechanisms as part of testing the feasibility of the sharing model within a distributed environment. Since CRITS is open source and allows for integration with other systems using open-source APIs, it can also be extended, which is essential for adapting the sharing model to the needs of different stakeholders.

Cybersecurity-threat Intelligence-sharing Model

Today, there is continuous access to instant information about virtually everything in the always-connected world. Personal and organisational information is easily accessible online using multiple-platforms, including Open Source Intelligence (OSINT). A common maxim states that 'information is power'; and in the defence environment, information is a critical element of power. As previously mentioned, receiving the right information on time can be the difference between a successful and unsuccessful cyberattack (Goodwin & Nicholas 2015).

In cyberwarfare and information-warfare environments, information operations in their various forms can be used to gain information and decision superiority over an adversary. In cybersecurity, the maleficent actors mostly rely on systems' vulnerabilities and threat information to breach or attack the target of interest. Thus, to mount a defence against malicious actors, cybercrimes, or cyberattacks, actionable information about the adversaries' systems' vulnerabilities, emerging cybersecurity threats, and current cyberattacks cannot be ignored.

Because most cyber ills are conducted using publicly available information by collaborating malicious actors, cyber defence stakeholders cannot continue to operate in isolation in their efforts to defend and protect cyberspace. The stakeholders in the defence environment need to start sharing cybersecurity threat intelligence in an inherently collaborative endeavour requiring cooperation among members of the cybersecurity community. Since sharing of threat intelligence can also occur at strategic, tactical, and operational levels, it is important that an appropriate sharing model is formulated and agreed upon by those involved.

First, the model needs to focus on the involved stakeholders' understanding of the operating environments. According to ThreatView (2016), if an organisation does not understand its assets, infrastructure, personnel, and business operations, it cannot know if it is presenting opportunities to malicious actors. Moreover, the information that needs to be shared amongst the stakeholders in some environments must be at a level higher than cyber information, which is mostly useful in reactive operations (for example, cyber incidents, observables, and Indicators of Compromise). The cybersecurity threat intelligence involved should be contextual and actionable, *and* should enable proactive and predictive responses to cyberattacks and cybercrimes. Furthermore, it should be possible to provide operational information about threat actors, their campaigns, cyber behaviours, and TTPs (Tools, Techniques, and Procedures). As such, information sharing between various stakeholders within an environment could prove vital in threat intelligence.

In some environments, there is a need to seek to share information and coordinate with industry partners in an integrated manner to promote situational and battlespace awareness. For example, if a third-party partner learns of malicious cyber activities that could affect important networks and systems, the sharing model could provide guidelines for the manner by which this threat information can be shared quickly enough to avoid the attack or reduce the risk. Above the sharing of threat information, lessons learned and cybersecurity best-practices can also be shared by the stakeholders to improve the resiliency of cyberspace. The following section discusses the research approach used to answer the questions put forward in the introduction section.

Research Approach

The research presented in this article followed an experimental approach as described by the Design Science Research (Vaishnavi & Kuechler 2014). This was preceded by a systematic literature review, in which an extensive study was conducted on the existing threat-intelligence exchange platforms, including *de facto* and *de jure* threat-exchange standards. From the literature, it was determined that a number of threat-intelligence sharing platforms exist; however, most of these platforms are isolated and do not necessarily adopt any sharing model or use semantic knowledge to ensure that actionable information is shared.

In addressing the main research problem raised in the introduction, the authors followed a conceptual modelling and practical approach to realize a Proof of Concept (PoC) for the proposed semantic-enabled sharing model. The PoC cyber-threat intelligence platform was implemented by using a series of integrated components, including an existing platform (namely, CRITS) as a foundation. In addition, Twitter, a popular social media platform, was selected as one of the data sources. Unlike the other data sources, it is discussed in this article mainly because it is capable of generating Big Data on a daily basis. As a matter of fact, to date, there are over 400 million tweets recorded per day (Tsukayama 2013), and this fact makes it a relevant case as it is highly improbable that one person or organisation can go through all these tweets or even share them with relevant stakeholders using traditional approaches. The following section discusses the technical platform prototyped for evaluating how sharing could occur using a secured web-based portal.

Cyber-threat Intelligence Platform

To combat the impact of cyber-crime and cyberattacks, organisations need to share the known threats with other relevant and trusted organisations as quickly as possible. Enormous amounts of data exist from various sources, such as Twitter. This section explores how the Big Data from a social network such as Twitter could be used to achieve the objective of sharing semantic-enabled, threat-intelligence information. Data can be pulled from Twitter, then filtered and cleaned as it arrives on the system by using API designed by the researchers. It can then be semantically filtered using stakeholder profiles and text-analysis APIs, such as TextRazor (Textrazor 2016).

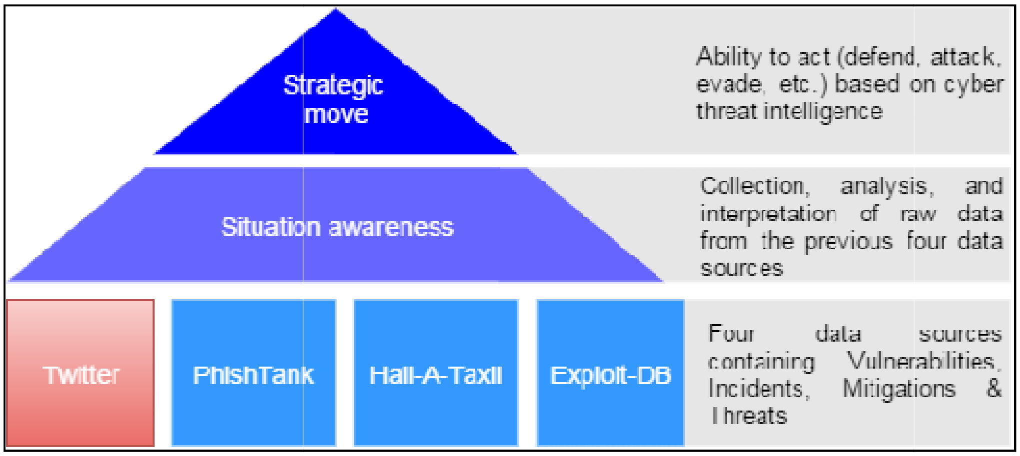


Figure 1: Platform high level architecture

At present, the in-house developed API has four data sources which contain vulnerabilities, incidents, mitigations, and threats. (See **Figure 1**, above.) The feeds from the four data sources are collected together into the database of the API where they are analysed, classified, and interpreted. The cyber-threat intelligence platform then polls from the API feeds using the TAXII poll service in order to generate cyber-threat intelligence information that can be actionable for attack on, defence against, or evasion of a cyber threat.

Figure 2, below, shows a raw Twitter feed, tweeted by the Information Security Hotspot about the SAP software download application that exposes cybersecurity vulnerability. This feed is pulled and converted into STIX format using an API developed for the platform. It should be noted that the Twitter Search API puts restrictions on the number of tweets that can be streamed by external applications; however, from the authors' experiments conducted over a few days, over 300,000 tweets were acquired. For demonstration purposes, Python was used to implement the streaming in-house API. The input to the API is a set of keywords that would assist in sourcing the relevant tweets.

Figure 2: Raw Twitter feed

Once all tweets are cleaned and tagged with the semantic data for the different stakeholders or keywords that a particular organisation is interested in, the tweet can then be pushed to the transport service in a STIX format as illustrated in **Figure 3**, below, from which it can be pulled by any other exchange platform that utilizes the TAXII service. This makes the authors' approach platform-independent and loosely coupled.

```xml
<stix:STIX_Package
        xmlns:URIObj="http://cybox.mitre.org/objects#URIObject-2"
        xmlns:cybox="http://cybox.mitre.org/cybox-2"
        xmlns:cyboxCommon="http://cybox.mitre.org/common-2"
        xmlns:cyboxVocabs="http://cybox.mitre.org/default_vocabularies-2"
        xmlns:example="http://example.com"
        xmlns:indicator="http://stix.mitre.org/Indicator-2"
        xmlns:stix="http://stix.mitre.org/stix-1"
        xmlns:stixCommon="http://stix.mitre.org/common-1"
        xmlns:stixVocabs="http://stix.mitre.org/default_vocabularies-1"
        xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" id="example:Package-2fe66f39-1ada-493c-b69e-a4c0db447834" version="1.2">
    <stix:STIX_Header>
        <stix:Title>Thread Exchange Platform Data</stix:Title>
        <stix:Package_Intent xsi:type="stixVocabs:PackageIntentVocab-1.0">Threat Report</stix:Package_Intent>
        <stix:Description>Twitter collected data</stix:Description>
    </stix:STIX_Header>
    <stix:Indicators>
        <stix:Indicator id="3950020" timestamp="2015-12-08T16:49:28.233000+00:00" xsi:type='indicator:IndicatorType'>
            <indicator:Title>Google Play Malware Can Infect Your PC Via Mobile Device    Tech </indicator:Title>
            <indicator:Type xsi:type="stixVocabs:IndicatorTypeVocab-1.1">URL Watchlist</indicator:Type>
            <indicator:Description>Google Play Malware Can Infect Your PC Via Mobile Device
http://thetechjournal.com/electronics/computer/security-computer-electronics/google-play-malware-can-infect-your-pc-via-mobile-device.xhtml
  Tech https://twitter.com/TheTechJournal/status/623406340212457472/photo/1</indicator:Description>
            <indicator:Short_Description> 5f7b910c-4f18-4bb1-b6e1-b10fee62d4ec</indicator:Short_Description>
            <indicator:Observable id="example:Observable-9777fa06-6f8c-4b05-9f21-de135778cc1f">
                <cybox:Observable_Composition operator="OR">
                    <cybox:Observable id="example:Observable-104c1f77-edb1-455f-932c-6366d4ddb36c">
                        <cybox:Object id="example:URI-525d6d5a-68f9-492d-9f16-534b02fb8bc2">
                            <cybox:Properties xsi:type="URIObj:URIObjectType" type="URL">
                                <URIObj:Value>http://thetechjournal.com/electronics/computer/security-computer-electronics/google-play-
malware-can-infect-your-pc-via-mobile-device.xhtml</URIObj:Value>
                            </cybox:Properties>
                        </cybox:Object>
                    </cybox:Observable>
                    <cybox:Observable id="example:Observable-b7dc89e5-0576-4c80-832f-93f8062c184e">
                        <cybox:Object id="example:URI-77bb240d-6af3-4839-be18-4dd5b39cea96">
                            <cybox:Properties xsi:type="URIObj:URIObjectType" type="URL">
                                <URIObj:Value>https://twitter.com/TheTechJournal/status/623406340212457472/photo/1</URIObj:Value>
                            </cybox:Properties>
                        </cybox:Object>
```


Figure 3: Tweet converted into STIX format

Once the above mentioned Twitter feed is STIXified by the in-house API, it can be given a security classification using the Traffic Light Protocol (TLP). The TLP classification is based on the possible impact the security threat would cause. This impact is measured on the estimated severity and length of the attack. **Table 2**, below, shows the threat-level classification of the Twitter feeds. Once the feeds are classified, they can then be pushed to the YETI server. The YETI server then provides three services which are Pushing, Polling, and Discovery. These three services are then consumed by the exchange platform. **Table 2**, below, also shows which cyber-threat intelligence information a public or private user has access to, based on the TLP of the cyber-threat feed. Private stakeholders can view all TLP feeds irrespective of the tag colour, whereas public stakeholders can only view TLP feeds that are tagged with a green colour.

TLP Colour	Threat Level	Description	Public Stakeholder	Private Stakeholder
RED	CRITICAL	Cyber-threat information that cannot be shared outside the organisation		✓
AMBER	HIGH	Cyber-threat information that can be shared outside the organisation		✓
GREEN	MEDIUM	Cyber-threat information that can be shared with anyone	✓	✓

Table 2: Classification of tweets

Shared cyber-threat intelligence can either be viewed in a public or private mode. The public stakeholder is able to see a list of cyber-threat information that is non-classified. This means that the information can be shared with anyone and has limited impact. The private stakeholder is able to see a table list of cyber-threat information that is classified and is only sharable within a group of affected stakeholders. In order to determine if the user can view public or private information, the platform queries the Lightweight Directory Access Protocol (LDAP) server for the type of TLP that the user is allowed to view and also the sharing rights the user has. The Twitter feeds are already classified when they are stored in the platform's local database.

Once the user has successfully logged onto the platform, he or she is immediately presented with a summary dashboard of all types of cyber-threat intelligence information; see **Figure 4**, below. From the main dashboard, the user can see the overview of each category: backdoor, malicious domain, malicious email, malicious indicators, malicious IP, and malicious samples. Each item has its own sub-dashboard with links to new items. The platform is not discussed in detail here due to space limitations.

Figure 4: Platform main dashboard

Conclusion and Further Research

The threats to cybersecurity are on the rise from different sources and for different reasons. This article discusses the need for a collaborative tool that can be used to analyse Big Data related to cybersecurity using rights management to separate publicly and privately accessible intelligence. Herein, intelligence is defined as analysed information that enables actionable reactions by stakeholders to events—in some cases, even before these events occur. For the conceptual model, the authors have discussed several options for the exchange platforms and exchange standards. By putting together the different exchange platforms and exchange standards, a conceptual model was selected and implemented using the experimental and practical research approach. The data sources fed into the conceptual model were raw Twitter feeds, which were STIXified and pulled into CRITS using a TAXII server as the Transport Service. The analysis of this data is displayed on the CRITS dashboard. Further research might include the integration of a mathematical model to measure the value and impact of shared cybersecurity threat intelligence over a period of time.

References

AbuseHelper 2011, *AbuseHelper: automatically process your incident notifications,* viewed 29 May 2016, <http://abusehelper.be/>.

AlienVault 2014, *Alien Vault: The Value of Crowd-Sourced Threat Intelligence,* viewed 29 May 2016, <https://www.alienvault.com/docs/whitepapers/AlienVault_The-Value-of-Crowd-Sourced-Threat-Intelligence.pdf>.

Barnum, S 2014, *Standardizing cyber threat intelligence information with the structured threat information eXpression (STIX™),* viewed 29 May 2016, <http://stixproject.github.io/getting-started/whitepaper/>.

Brown, S, Gommers, J & Serrano, O 2015. *From cyber security information sharing to threat management,* ACM, New York, U.S.A., pp. 43-9.

CERT.org 2016, *CSIRT frequently asked questions (FAQ),* viewed 29 May 2016, <https://www.cert.org/incident-management/csirt-development/csirt-faq.cfm?>.

CIRCL 2016, *Malware Information Sharing Platform MISP - a threat sharing platform,* viewed 29 May 2016, <https://www.circl.lu/services/misp-malware-information-sharing-platform/>.

Eom, JH 2014, 'Roles and responsibilities of cyber intelligence for cyber operations in cyberspace', *International Journal of Software Engineering and Its Applications,* vol. 8, no. 9, pp. 137-46.

Farnham, G 2013, *Tools and Standards for Cyber Threat Intelligence Projects,* viewed 29 May 2016, <http://www.sans.org/reading-room/whitepapers/warfare/tools-standards-cyber-threat-intelligence-projects-34375>.

Fransen, F, Smulders, A & Kerkdijk, R 2015, 'Cyber security information exchange to gain insight into the effects of cyber threats and incidents', *e & i Elektrotechnik und Informationstechnik,* vol. 132, no. 2, pp. 106-12.

Goffin, M 2014, *Collaborative Research Into Threats (CRITs),* viewed 29 May 2016, <https://www.mitre.org/capabilities/cybersecurity/overview/cybersecurity-blog/collaborative-research-into-threats-crits>.

Goodwin, C & Nicholas, JP 2015, *A framework for cybersecurity information sharing and risk reduction,* viewed 2 January 2016, <http://download.microsoft.com/download/8/0/1/801358EC-2A0A-4675-A2E796C2E7B93E73/Framework_for_Cybersecurity_Info_Sharing.pdf>.

HITRUSTAlliance 2016, *Cyber Threat XChange (CTX),* viewed 29 May 2016, <https://hitrustalliance.net/cyber-threat-xchange/>.

IntelMQ 2015, *Welcome to IntelMQ!,* viewed 29 May 2016, <https://github.com/certtools/intelmq>.

Kaisler, S, Armour, F, Espinosa, JA & Money, W 2013, *Big data: issues and challenges moving forward*, IEEE, pp. 995-1004.

Katal, A, Wazid, M & Goudar, RH 2013, *Big data: issues, challenges, tools and good practices,* Noida, IEEE, pp. 404-09.

Khan, N, Yaqoob, I, Hashem, IAT, Inayat, Z, Ali, WKM, Alam, M, Shiraz, M, & Gani, A, 2014, ‘Big Data: survey, technologies, opportunities, and challenges’, *The Scientific World Journal,* vol. 2014, p. 18.

Mishra, P 2014, *Cyber Threat Intelligence,* viewed 29 May 2016, <http://www.slideshare.net/prachimishra31/cyber-threat-intelligence>.

MITRECorporation 2016, *About TAXII,* viewed 29 May 2016, <http://taxiiproject.github.io/about/>.

Mtsweni, J, Shozi, NA, Matenche, K, Mutemwa, M, Mkhonto, N, & van Vuuren, JJ, 2016, *Development of a semantic-enabled cybersecurity threat intelligence sharing model,* Academic Conferences and Publishing International Limited, Boston, MA, U.S.A., pp. 244-52.

Polancich, J 2014, *Cyber risk intelligence: what you don’t know is most definitely hurting you,* viewed 29 May 2016, <http://www.securityweek.com/cyber-risk-intelligence-what-you-don’t-know-most-definitely-hurting-you>.

Ring, T 2014, *Threat intelligence: why people don't share,* viewed 04 January 2016, <http://www.sciencedirect.com/science/article/pii/S1361372314704695>.

Soltra 2016, *Soltra edgethreat intelligence solution,* viewed 29 May 2016, <https://soltra.com/>.

TextRazor 2016, *Extract meaning from your text,* viewed 29 May 2016, <https://www.textrazor.com/>.

ThreatView 2016, *Cyber threat & reputation intelligence,* viewed 24 March 2016, <https://www.threatview.ca/>.

Tsukayama, H 2013, *Twitter turns 7: users send over 400 million Tweets per day,* viewed 1 January 2016, <https://www.washingtonpost.com/business/technology/twitter-turns-7-users-send-over-400-million-tweets-per-day/2013/03/21/2925ef60-9222-11e2-bdea e32ad90da239_ story.html>.

Vaishnavi, V & Kuechler, W 2014, *Design Science Research in Information Systems,* viewed 1 January 2016, <http://www.desrist.org/design-research-in-information-systems/>.

Van der Meulen, R 2015, *Gartner Says 6.4 billion connected “things” will be in use in 2016, up 30 percent from 2015,* viewed 8 March 2016, <http://www.gartner.com/newsroom/id/3165317>.

Waltz, E 1998. *Information warfare principles and operations,* 1st ed., Artech House, Inc., Norwood, MA, U.S.A.

Zikopoulos, P & Eaton, C 2011. *Understanding big data: Analytics for enterprise class hadoop and streaming data,* McGraw-Hill Osborne Media, New York, NY, U.S.A.