Journal of Information Warfare

Volume 16, Issue 1
Winter 2017

Contents

Journal of Information Warfare

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

Print Version
ISSN 1445-3312

Online Version
ISSN 1445-3347

From the Editor

Winter 2017

We certainly live in interesting times. Just look at the United States presidential election in 2016. Which of the following were at play? Partisan politics? Russian interference? Information Warfare? An act of war? All of the above? Or, none of the above? The articles in this latest edition of the *Journal of Information Warfare* cover issues and ideas as wide-ranging as the possibilities so characteristic of our current times. We hope you enjoy these papers.

For those looking for additional opportunities to publish academic articles in these areas, the ***16th European Conference on Cyber Warfare and Security*** will be held at University College Dublin in June 2017. Following that, the ***13th International Conference on Cyber Warfare and Security*** will be hosted by the National Defense University (NDU) in Washington, D.C., in March 2018; the ***17th European Conference on Cyber Warfare and Security*** will be held in Oslo, Norway in July 2018; and the ***18th European Conference on Cyber Warfare and Security*** will be held at the University of Coimbra in Portugal in July 2019.

In addition, Edith Cowan University's Security Research Institute in Perth, Australia, is holding the ***2017 SRI Security Congress*** in the late November-early December timeframe, in Perth, Australia. All of the conferences of this federated series are co-located and offer attendees the opportunity to interact with a number of different, but inter-related cyber events and lectures:

- 17th Australian Information Warfare Conference
- 15th Australian Information Security Management Conference
- 15th Australian Digital Forensics Conference
- 10th Australian Security and Intelligence Conference
- 5th Australian e-Health Informatics and Security Conference

For more information regarding these events, visit http://conferences.secau.org/.

Finally, *JIW* is always interested in increasing our roster of subject-matter-expert reviewers. If you are interested in serving on our Editorial Review Board, please contact me (larmistead@gbpts.com) or email Michael McGill at mmcgill@gbpts.com. The researchers who serve as reviewers help us conduct our double-blind, peer-review assessment and their names appear, with our thanks, in each issue of the journal.

Until next time, cheers,

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

Authors

Nurudeen Ajayi is a lecturer in the Discipline of Information Systems and Technology at the University of KwaZulu-Natal, South Africa. His research interests include information security, risk and assurance, enterprise information systems, and technology in education.

Kenneth J. Boyte is an instructor in the Department of Business, English, and Language Arts at Cabrillo College in the San Francisco Bay Area. He has a background in teaching English as a Second Language, and writing to adults in South Korea and the United States. He also has worked extensively in educational publishing for the U.S. Army Defense Language Institute and McGraw-Hill, and in the field of journalism. He earned a bachelor's degree in journalism at Auburn University, a master's degree in journalism from Southern Illinois University, a master's degree in Teaching English to Speakers of Other Languages (TESOL) from the Middlebury Institute of International Studies, and a master's certificate in intelligence analysis from the American Military University. His current research interests include issues related to reading comprehension and memory, the computer-mediated uses of English for political purposes via social media, and cyber warfare.

Dr. Douglas C. Derrick is an associate professor of IT Innovation at the University of Nebraska at Omaha. He earned his doctorate in management information systems from the University of Arizona. He holds a master's degree in computer science from Texas A&M University and a master's degree in business of administration from San Jose State University. He is a Distinguished Graduate of the United States Air Force Academy. Prior to joining UNO, he worked as a Program Manager at MacAulay-Brown, Inc. and also served as an Air Force Officer. As a contractor and academic, he has been awarded contracts and grants from the Department of Defense totalling $36.23 million over the last 10 years (principal investigator awards total $15.37 million). His research interests include human-agent interactions, intelligent agents, data fusion, decision support systems, and persuasive technology. He has had articles published in journals and conferences including the *Journal of Management Information Systems, ACM Transactions on Management Information Systems, IEEE Intelligent Systems, AIS Transactions on Human-Computer Interactions, Group Decision and Negotiation*, *Hawaii International Conference on System Sciences, IEEE International Conference on Intelligence and Security* Informatics, and the *IEEE International Carnahan Conference on Security Technology*.

Olajide Kolawole Enigbokan is a Cloud Support Associate (Developer and Mobile Services) at Amazon Web Services. He earned a master's degree in information systems and technology from the University of KwaZulu-Natal, South Africa.

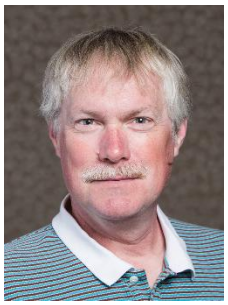
Eric Filiol is the head of the $(C+V)^{O}$ Research Lab at ESIEA, France and a senior consultant in offensive cyber security and intelligence. He spent 22 years in the French Army (Infantry/Marine Corps). He holds an Engineer diploma in cryptology, a doctorate in applied mathematics and computer science, and a Habilitation Thesis in computer science. He also graduated from NATO in InfoOps. He is the editor-in-chief of the *Journal in Computer Virology* and has been a speaker at international security events including Black Hat, CCC, CanSecWest, PacSec, Hack.lu, Brucon, and H2HC.

Cecilia Gallais joined TEVALIS in the framework of a thesis on the formalisation and establishment of an algebraic model for the cyberattack and critical infrastructure concepts. She studied mathematics and cryptography at the University of Rennes in France and completed an internship at Orange Labs in Caen, France.

Ross Gordon has served as a commissioned officer in the Australian Army for 13 years, having undertaken a broad range of communications roles within the Royal Australian Corps of Signals. He is a graduate of both the Australian Defence Force Academy and the Royal Military College-Duntroon and has deployed multiple times throughout his career. In 2015 he completed a master's degree in information technology at the Australian Centre of Cyber Security. He also holds a bachelor's degree in computer science and management, and a master's degree in business, both from the University of New South Wales.

Mackenzie Harms is a doctoral student in industrial and organizational psychology at the University of Nebraska at Omaha.

Dr. Gina Ligon is an associate professor of management at the University of Nebraska at Omaha and serves as the Director of Research and Development in the Center of Collaboration Science (CCS). She earned a doctorate in industrial and organizational psychology with a minor in measurement and statistics from the University of Oklahoma. She is a member of the National Consortium of Studies of Terrorism and Responses to Terrorism (START). Since arriving at UNO, she has been awarded more than $1 million in grants and contracts with principal investigator awards totalling approximately $400,000. She currently is the principal investigator on a grant from the Department of Homeland Security examining the leadership and performance of transnational Violent Extremist Organisations (VEOs,) and is the originator of the Leadership of the Extreme and Dangerous for Innovative Results (LEADIR) database. She has worked with Department of Defense agencies through grants and contracts focused on markers of violent ideological groups, leadership assessment, organisational innovation, and succession planning for scientific positions. Prior to joining UNO, she was a faculty member in the Department of Psychology at Villanova University. She also worked as a management consultant in St. Louis with the firm Psychological Associates, partnering with Fortune 500 organisations on the implementation of leader development and succession planning initiatives. Her research interests include violent ideological groups, expertise and leadership development, and collaboration management. She has published more than 40 peer-reviewed articles in the areas of leadership, innovation, and violent groups.



Dr. William R. Mahoney is an associate professor in the College of Information Science and Technology at the University of Nebraska at Omaha, and the Director of the Nebraska University Center for Information Assurance (NUCIA). He regularly teaches in both the information assurance and computer science areas and is a reviewer for several information warfare publications and conferences. Prior to entering academe, he worked for more than 20 years in the computer design industry, in the areas of embedded computing and real-time operating systems. He earned undergraduate degrees from Southern Illinois University, and both a master's degree and doctorate from the University of Nebraska. His primary research interests include language compilers, hardware and instruction set design, and code generation and optimisation—as those topics relate to information assurance goals.

Erick Mandt is a Department of Defense civilian employee. Prior to working with the DoD, he served for 20 years in the U.S. Navy, and retired as a Master Chief Petty Officer. He earned an undergraduate degree in Russian area studies from Excelsior College and a master's degree in cyber security from Utica College. His research interests focus on critical thinking and structured analysis processes within network defence operations.

Nicholas Shallcross is an Operations Research Analyst in the United States Army. He earned a master's degree in operations research from the Air Force Institute of Technology in 2016. His research interests include the probabilistic modelling of large systems and global conflict forecasting.

Social Media and Information Operations in the 21st Century

NJ Shallcross

Graduate School of Engineering & Management
Air Force Institute of Technology
Dayton, Ohio, U.S.A.
E-mail: Nicholas.j.shallcross.mil@mail.mil

Abstract**:** *Modern military operations continue to be extraordinarily susceptible to the effects of cyber-based Information Operations (IO). Within social media lies the ability to gain a clearer perspective of the 21st-century battlefields, enabling rapid and informed decision making and decisive action by commanders and their staffs. This paper discusses emerging trends, threats, and concepts that are being employed by numerous actors around the globe to gain positional advantage both internal and external to the cyberspace domain.*

Keywords: *Information Operations, Social Media, Information Warfare*

Introduction

We do not talk to say something, but to obtain a certain effect.
—Josef Goebbels

The mid-day sun was oppressive, as only a Tehran summer can be. Neda Agha-Soltan, a 26-year-old native of Tehran, had been waiting patiently for several minutes in a car with her sign language instructor and two friends, when they decided to seek a respite from the 120 degree heat enveloping them in the small sedan. Neda and her friends were on their way to attend a protest against the outcome of the 2009 presidential elections when they proceeded on foot to a shaded area not far away from where some intermittent protests were taking place. Within minutes of their arrival, a sniper's bullet, fired by one of the pro-government militiamen, the Basij, struck Neda in the chest; the wound proved fatal within minutes. For many of the anti-government Iranians this was a scenario, commonly repeated over the years; however Neda's death was different—it was captured on video.

Within hours of the incident, the video had spread like wildfire across the Internet, sparking ten days of violent protests in the Iranian capital. Twenty-four hours after Neda's death, there were over 6,800 references to her on the Persian Language Google site, and Twitter recorded anti-Iranian government tweets that numbered in the millions (Fathi 2009). Scores of Iranian students took to Twitter, Facebook, and Flickr to communicate, organise, and coordinate future demonstrations, much to the consternation of the Iranian government and its security forces. The aftermath of Neda's death demonstrated the potential and adaptability of Social Media as an instrument of power and as weapon; it is a tool that is neither easily wielded nor contained.

Information used as an element of warfare and national power is as old as civilization itself; however, the advent of the information age has resulted in an exponential propagation of tactics, technologies, and threats as they relate to the relatively new art and science called Information Operations (IO). The advent of Web 2.0 technologies, specifically social media, and their subsequent use in IO, represents an evolution in military affairs with the potential to level the playing field between the greater and lesser powers, as well as the non-state actors of the world. This paper provides an overview of social media as an instrument of national power, potential risks, benefits and limiting factors associated with its use, the operational environment, and a proposed national social-media strategy.

Before exploring the capabilities of social media as an instrument of national power, or as a weapon, it is appropriate to first identify and define key terms specifically associated with information conflict as it relates to social media. Social media is a subset of Web 2.0 technologies that include all social networks, Internet web-logs (blogs), wiki-sites, and mobile telecommunication device applications (Van Niekerk & Maharaj 2013). The concept of user-defined and -created content, open collaboration, information sharing and propagation, and collective intelligence are the common underpinnings for all such technologies (Van Niekerk & Maharaj 2013). While social media has benefitted modern society, such as in emergency response and social advocacy, it has also been employed for more nefarious activities and even as a weapon. This was demonstrated by Hezbollah during its 2006 conflict with Israel when it employed social media, as part of a greater IO campaign, to control and manipulate the narrative in its favour (Kalb 2007). As the following concepts are all germane to the subject of this paper, and the development of a social-media security strategy, they are considered.

- Network Warfare (NETWAR or cyber warfare): Offensive and defensive actions in relation to information, communications, and computer networks and infrastructure (Brazzoli 2007)
- Command and Control Warfare (C2W): Actions taken to manage, direct, and coordinate the movement and activities of various forces; seeks to protect this ability in friendly forces and disrupt the ability for an adversary (Brazzoli 2007)
- Intelligence-based Warfare: Actions taken to degrade an adversary's intelligence cycle while protecting one's own cycle (Brazzoli 2007)
- Psychological Operations (PSYOP): Actions taken to alter the perceptions of the target audience in support of the commander's objectives (Brazzoli 2007)
- Cyber power: The ability to use cyberspace to create advantages and influence events in all operational environments across the instruments of power (Murphy 2010)

While the above definitions are all not necessarily doctrinal, in accordance with U.S. Army or Joint doctrine, they share many similarities with doctrinal definitions used by the United States, our allies, Russia, and the People's Republic of China (PRC). Ultimately, the role of all types of media, particularly social media, in conflict is what separates the wars of the 21st century from those of the past, the real-time and near real-time reports have transformed warfare into an interactive 'spectator sport' for regular citizens; thus, media has become an integral part of warfare and the modern battlefield (Qiao & Wang 1999).

Threat Environment and Capabilities

Four hostile newspapers are more to be feared than a thousand bayonets.
—Napoleon Bonaparte

What is the threat environment of the 21st century? It is an environment of persistent conflict between not only states, but also between states and proto-governments, states and non-state actors, and even between states and super-empowered individuals. It is an environment where populations and the support they provide have increasingly become the key terrain in conflicts where some of the most decisive battles are fought in the multi-dimensional realm of cyberspace. The Department of Defense (DoD) defines cyberspace as

> A global domain within the information environment consisting of the independent network of information technology infrastructures, including the Internet, telecommunications networks, computer systems, and embedded processors and controllers. (Murphy 2010)

The Internet is the prominent domain in this definition, and for good reason. The Internet, or World Wide Web, is essentially an ungoverned state of literally billions of people. It is an environment that promotes anonymity, gives a voice to the individual that can be heard by millions; but more importantly, it is an environment without gravity. In this context, the phrase 'without gravity' means that the Internet, as an operational environment, is not constrained by the three-dimensional environment that bounds the physical world. For example, the dimensions and terrain of a battlefield may inhibit an army's ability to deploy and mass forces effectively; only so many troops and equipment can fit into a finite space. In another scenario, a person may wish to disseminate some important piece of information, but all he or she has is a telephone or a courier; either mode ultimately limits his or her ability to reach multiple recipients and requires greater quantities of time per individual contact, a resource that may otherwise be limited. By comparison, cyberspace transcends the constraints of geography or physical location. An army of ten hackers attacking a site can easily turn into 1000, converging on the objective from across the globe, as it is not subject to the dual tyrannies of space and time. A blogger with one post can reach millions of targets because, unlike a phone call, a blog post is not scalable; it takes the same amount of effort to reach one person as it does five million. What this means to the United States is that the traditional threat profile has changed; the decades-wide technological gap between the United States and its adversaries has closed significantly. It also means that the security once provided by the Atlantic and Pacific oceans is no longer as relevant.

Social media as a weapon

It takes a special type of person to become an insurgent; the individual must possess the required ideology, the physical and mental characteristics to tolerate hardship, and the desire to fight for his or her set of beliefs (Metz 2012). Such people are exceedingly rare as a percentage of the population, and it was with great difficulty that insurgent and terrorist organisations recruited members prior to the advent of the Internet. Through the adept use of social media, an organisation can openly and anonymously recruit supporters, members, and financiers with limited risk to the organisation itself. Does this mean that social media is truly a weapon that can be wielded against an adversary? On the surface, social media resembles a command and control tool to recruit supporters and direct operations from remote locations. In this context, social

media more closely resembles a radio set than a weapon; but in military applications, a radio is one of the most effective weapons in any arsenal. A radio might not directly achieve kinetic effects, but today most munitions employed against an adversary receive terminal guidance and control through a radio set. Framed in this setting, social media becomes an integral part of the kill chain that has the capability to direct a host of both lethal and non-lethal effects onto a target. In this regard, insurgencies and terrorists employ social media to direct attacks abroad, to damage credibility, and to undermine authority. It has also been shown that the greater the degree of networking within an organisation, the greater the likelihood that social networking and information technologies are used to support the group's decision making (Arquilla, Ronfeldt & Zanini 1999). Additionally, given the huge flow of daily Internet traffic, the open use of social media as weapon can easily go undetected or be discounted as a less serious incident by state security services (Metz 2012). It should be noted, however, that terrorist networks are not the only entities that utilize social media as a weapons system; many world governments have now grasped its potential and have even used it to achieve significant operational successes.

Russian social media operations

The proliferation of social media as a weapon used by terrorist organisations is due in part to its ability to level the playing field between the organisation and the resource-rich state. However, the state can greatly benefit from using weaponised social media, as demonstrated in Russia's recent use of social media against the Ukraine and the West. In its ongoing conflict with the Ukraine, Russia has implemented a complex IO strategy, directed by several government agencies, which is actively supplanting Ukrainian media outlets to undermine the government and coordinate the actions of pro-Russian rebels in the Crimea and elsewhere (Taia Global Inc. 2015). Additionally, the Russian Federal Security Service (FSB) has developed a network of 'Trolls' or fake social media accounts to help spread propaganda. In this system, a single agent may control up to ten accounts that are skilfully designed to look like the accounts of regular people, with daily traffic common to most social media users. This so-called 'Troll Army' has consistently attacked any media outlet that has criticized the Putin regime, with some websites reporting up to 40,000 comments per day (Gregory 2014).

The FSB's use of social media is not confined to non-lethal influence operations; it is believed to be the catalyst for a series of bombings, intended to destabilize the Ukrainian government, carried out by Russian proxies operating in the Ukraine (Tucker 2015). In addition to those attacks, the Ukrainian government reported that Russian-controlled terrorist cells have been ordered to attack critical government infrastructure, transportation hubs, and security forces (Tucker 2015). Through the use of proxies, controlled through state-supported social media, Russia has achieved some of its operational aims within the Ukraine, despite a relatively limited use of conventional military forces. Additionally, its IO strategy has provided Russia with 'plausible deniability' of its involvement in many of the recent attacks inside the Ukraine, which has hindered international efforts to curb the ongoing violence.

Information and the Application of Operational Art

Throughout the entire course of history, warfare is always changing.
—Andre Beaufre

It is impossible to understand and apply military operational art without understanding the basic definition of combat power. The United States' Department of the Army (2012) doctrine defines 'combat power' as the total means of destructive, constructive, and information capabilities that a military unit or formation can apply at a given time; formations generate combat power by converting potential action into effective action (ADRP 3-0 2012). In the execution of combined operations, capabilities are conceptualized in terms of the eight elements of combat power: Mission Command, Movement and Maneuver, Intelligence, Fires, Sustainment, Protection, Leadership, and Information. In **Figure 1**, below, the elements of combat power are visualized as the six Warfighting Functions (Mission Command, Movement and Maneuver, Intelligence, Fires, Sustainment, Protection) directed and managed through the application of leadership, all of which are influenced by the overall information environment. With this in mind, the "hallmark of operational art is the integration of the temporally and spatially distributed operations into one coherent whole" (Crowell 2010).

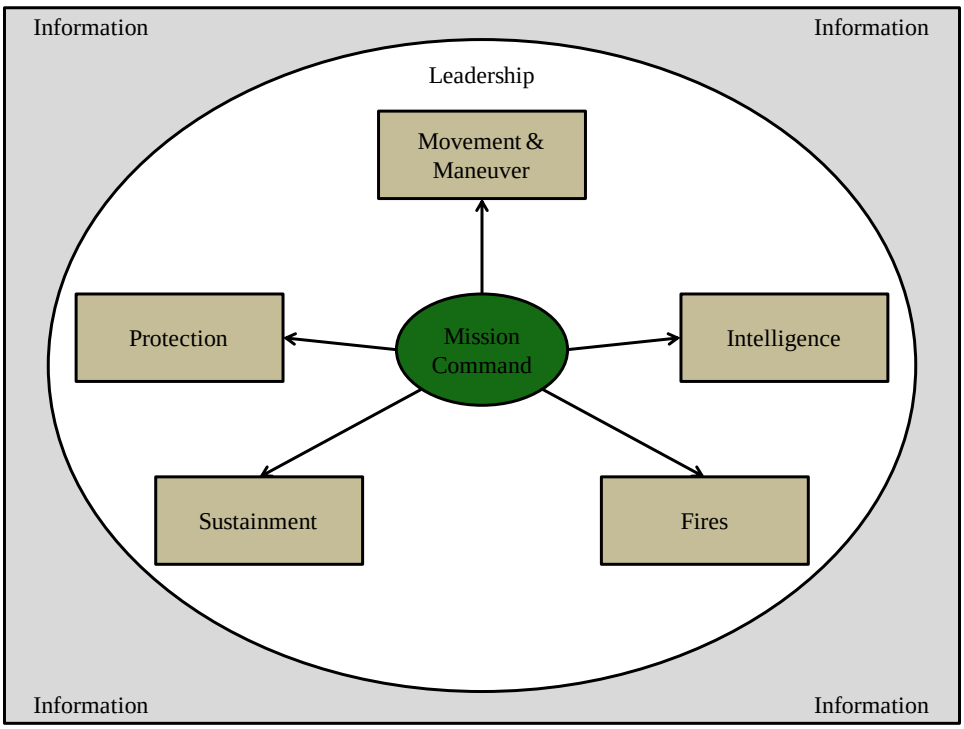


Figure 1: The elements of combat power (ADRP 3-0 2012)

The key takeaway from the doctrinal definition above is its focus on combined arms in terms of application. In the Western sense, the notion of combined arms is relatively constrained in its definition, generally referring to the combined use of infantry, armour, artillery and aviation, and other combat and supporting arms to achieve battlefield effects. While this definition has expanded over the past decade to include a host of other combat enablers, it still implies a

discrete and rather rigid doctrinal template on which to plan operations. America's adversaries, in comparison, tend to take a more fluid approach to the application of operational art, an approach that may be more suited for 21st-century warfare. Concepts such as omni-directionality, synchrony, and asymmetry are not just buzz words in Eastern military thinking, as they used to be; they now form the basis of the adversary's doctrine (Crowell 2010). To gain a better understanding of this paradigm shift, the definitions below are provided for doctrinal context.

Omni-directionality requires that commanders observe a potential battlefield without mental preconditions or blind spots. The designing of plans, employment measures, and combinations must use all war resources which can be mobilized. The commander is enjoined to make no distinction between what is or is not the battlefield. All traditional domains, (ground, sea, air, and outer space) as well as politics, economics, culture, and moral factors are to be considered battlefields (Qiao & Wang 1999).

Synchrony enjoins commanders to link the disaggregated nature of multiple battlefields in different domains with consideration of the temporal dimension. In other words, they must conduct actions in different spaces in the same period of time to achieve desired effects. Instead of phases with accumulated results of multiple battles, strategic results can now be attained rapidly by simultaneous actions or at designated times (Qiao & Wang 1999).

Asymmetry manifests itself to some extent in every aspect of warfare. However, asymmetry has been sought in operational terms within traditional military dimensions. In war beyond limits, the spectrum for overlooking the normal rules is much wider (Qiao & Wang 1999).

As a critical element within the information environment, social media provides a readily available means and conduit to attack and exploit the temporal dimensions of the modern battlefield. The Second Lebanon War fought between Israel and Hezbollah (Hizb'allah) is a classic example of the judicious use of social-media operations in hybrid war pitting a Western style military against a foe employing many of the above concepts. At the outset of hostilities, the Israeli Defense Force (IDF) assumed it was fighting the same insurgent force it had battled for decades—an assumption that would prove costly (Crowell 2010). By 2006, Hezbollah had transformed itself into an exceptionally lethal and technologically advanced hybrid force with modern weapons, reconnaissance and communications equipment, and tactics. The IDF soon realized it was combating Hezbollah not only on land, in the air, and at sea, but now in cyber space. Hezbollah's IO, which were enabled by hacking into several web sites including a Texas-based cable company, quickly disseminated their strategic messages to a worldwide audience. The result garnered significant moral, physical, and financial support, resulting in enhanced effectiveness of their tactical operations (Crowell 2010). In essence, the IDF had greatly underestimated a weaker adversary that literally beat it to the punch regarding strategic communications, amongst other things, the impacts of which were felt throughout the whole of the Israeli government and society.

Less than three years after the termination of the Second Lebanon War, Israel once again found itself in the midst of a violent conflict with Lebanese Hezbollah. On 27 December 2008 the IDF initiated OPERATION CAST LEAD in response to continued missile attacks originating in Gaza

(Caldwell, Murphy & Menning 2009). In an effort to avoid the mistakes of the previous conflict, Israel embarked on a massive public relations campaign using social media sites such as blogs, YouTube, and Facebook for both tactical and strategic communications. Following the conclusion of the 2006 conflict, the IDF created a special think tank, the Winograd Commission, which, among other things, recommended the formation of special IO units to coordinate public relations across a wide array of media outlets (Caldwell, Murphy & Menning 2009). Additionally, what was omitted from the Israelis' strategic communications was as important as what the Israelis communicated to their global audience. In 2006, the Israeli government had publically stated very rigid objectives and timelines, a strategy that would come back to haunt it. In the 2008-2009 conflict, the Israelis were far less definitive in their operational and tactical objectives; moreover, there was no timeline publically placed on the IDF in terms of mission accomplishment (Caldwell, Murphy & Menning 2009). As was the case in nearly all previous conflicts since 1948, the Israelis knew that United Nations' intervention was only a matter of time. With this in mind, IDF senior planners had designed their IO campaign around this eventuality, with the added knowledge that a new U.S. administration would take office in late January of 2009. Therefore, the goal of the IO campaign was to buy the IDF operational time by continuing to frame the Gaza incursion in a positive or at least neutral light. In the end, the campaign bought the Israelis the strategic depth their country lacked and allowed them to accomplish a majority of their tactical and operational objectives (Caldwell, Murphy & Menning 2009).

Social-Media Security Strategy

In communicating ground, I would pay strict attention to my defenses.

—Sun Tzu, 4th Century BCE

Any strategy that employs social media as a principal or supporting means must contend with the fact that by its very nature, social media is a double-edged sword. The basic concept of social media is based on information sharing and collaboration, the nature of which results in more vulnerabilities than those of traditional web pages, thereby increasing the potential for security risks (Van Niekerk & Maharaj 2013). The use of social media is fraught with risks, from both cyber-security and strategic-communications perspectives, that must be mitigated and accounted for in any operational plan. Therefore, a strategy regarding the use of the Internet to influence the information environment requires managing risk of attack, while pursuing any and all opportunities to compete (Murphy 2010). As noted earlier, success in 21st-century conflict requires that all available resources that can be mobilized be brought to bear on the adversary requiring a 'whole of government' approach, and must depend on the judicious and balanced application of the elements of national power.

Offensive operations

One objective of any offensive operation is to seize and maintain the initiative. As demonstrated in the two most recent Israeli conflicts, the first side to present its strategic communications to the world enjoyed a marked advantage over its adversary. However, this simple example portrays the social media offensive operations as discrete events rather than as a continuous, evolving process. In reality, social-media offensive operations are far more complex, and require more than being the first to get the word out. As shown in the **Figure 2**, below, social media can be employed to exploit vulnerabilities in an adversary's software and wetware, to compromise

sensitive information, to command and control remote operational teams abroad, as well as to influence a population's perception. While it is a good schematic of social media's role in operations, the figure tends to portray the system as rather stove-piped. It is imperative to the success of commanders and their staff that those modelling, planning, and managing these processes take into account that they are not only simultaneous, but are also potentially completely intermingled; for example, actions taken to disrupt communications will also, to some degree, end up influencing the population, whether that outcome is desired or not.

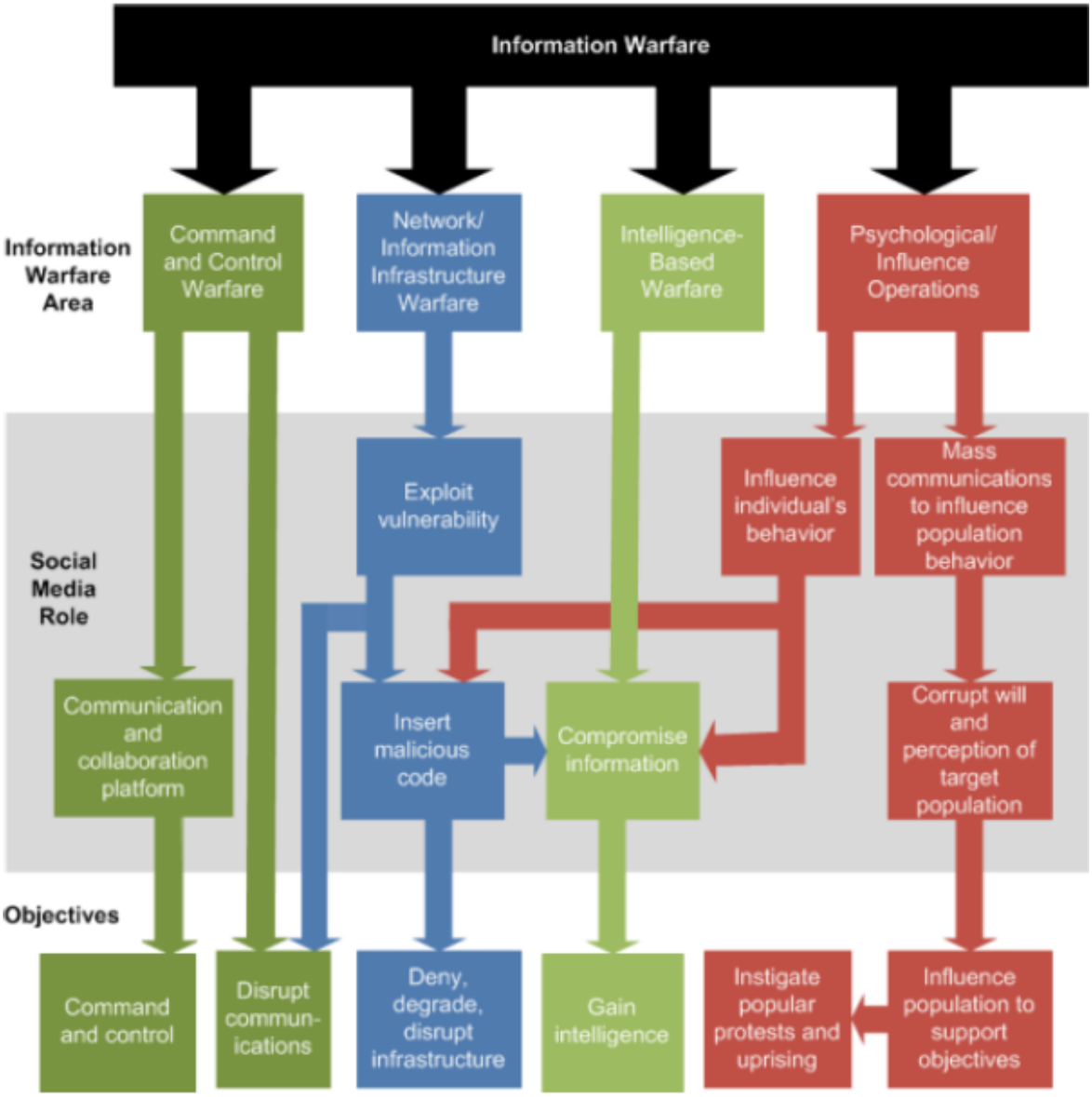


Figure 2: Social media in information warfare (Van Niekerk & Maharaj 2013)

Because social media is comprised of a wide range of technologies, communication styles, physical and computer languages, as well as a multitude of beneficial, benign, and malign actors,

it is helpful to view and model this environment as a complex adaptive system, with the potential to self-organize (Van Niekerk & Maharaj 2013). In this context, the modelling and analysis effort of any social-media operational strategy should not only focus on the measures of merit/effectiveness used within the campaign, but also on the secondary and tertiary effects of possible decisions made throughout. For example, a misstep in strategic communications, at any level of war, may result in a persistent degradation of public perception at home and abroad, which could take years to overcome. Additionally, the National Command Authorities may object to the use of a certain technology that could be vital in preventing a debilitating attack by an adversary. In this scenario, the commander must balance both the short- and long-term risks associated with either using the 'technology', which may be a closely held national secret and which might forever limit its subsequent effectiveness, or abstaining in favour of a less effective measure.

How might the United States offensively employ social media? The Arab Spring may provide a formula for the offensive use of social media against a government or transnational group. In this scenario, the target is the government of the adversarial nation, with the aggressor being the general population of both nations, which will act as proxies in the overall strategy (Van Niekerk, Pillay & Maharaj 2011). The aggressor's motivation is to remove the targeted government from power, due to the perception (whether real or otherwise) of corruption, ineptitude, illegitimacy, and/or oppression. The United States' IO should employ a combination of diplomatic engagements, as well as psychological and exploitation operations on both the aggressor populations and the targeted government, in conjunction with 'network-centric' warfare aimed at exacerbating and exposing deficiencies and weaknesses within the target's structure. The target sets are considered to be political and social constructs of the targeted nation, along with the perceptions of the local populace and international community (Van Niekerk, Pillay & Maharaj 2011). However, any strategy that seeks to supplant a government or leadership structure must also account for the resulting power vacuum and associated risks that will inevitably follow any such event, and must include multiple, redundant branch plans and sequels for consequence management.

Defensive operations

Due to the nature of the information environment, a balanced approach is required to successfully conduct simultaneous offensive and defensive social media operations (see Information Warfare Lifecycle Model in Van Niekerk, Pillay & Maharaj 2011).

The cornerstone of any defence entails a robust and redundant information-assurance and early-warning system that can detect and deny attacks on America's information infrastructure. By their design, these systems are reactive in nature as they attempt to form a shield against attacks but fail to prevent the launch of such attacks. A comprehensive defensive strategy must also include active, non-reactionary defenses, where the objective is not to defeat an attack, but to prevent it from occurring in the first place. One method, using the whole-of-government approach, is deterrence, similar to that used during the Cold War. This method requires the defending nation to publically declare that any attack on its information infrastructure will be considered an act of war, possibly akin to a Weapon of Mass Destruction (WMD) strike, and that the nation will exercise its rights to defend itself using any method at its disposal. This strategy requires both the national and political will, as well as the physical capabilities, to present a

credible deterrence. Another method, which may be politically more palatable, may take the form of a 'cyber-spoiling attack'. In military parlance, a 'spoiling attack' is a limited defensive operation mounted against a staging-offensive force with the objective of disrupting, delaying, or preventing a future attack. In cyberspace, this tactic could also achieve the same objectives as its conventional counterpart, possibly in a covert manner, preventing the escalation of hostilities between two states or groups. A spoiling attack may take the form of an electronic attack on some piece of hardware or system, or an attack on an individual, group, or nation using one of the various forms of new media. The crux of this strategy is its heavy reliance on intelligence, as an undetected attack cannot be prevented. Ultimately, an effective defensive strategy for the United States must entail an evolving combination of passive, active, and deterrent measures to adequately defend the homeland.

Conclusions

The great masters of warfare techniques during the 21st century will be those who employ innovative methods to recombine various capabilities to attain tactical, campaign, and strategic goals.

—Yier Tierfude

While today's security outlook is decidedly terrorism-focused, and with good reason, the United States must use social media, as part of an overall security strategy, to defend itself against attacks from transnational groups and nation-states which have the potential to threaten America's interests. Today's terrorist network is a widely dispersed but connected organisation which relies heavily on swarming tactics to achieve operational effects (Metz 2012). While this aspect of terrorist organisations makes them particularly hard to defeat, let alone destroy, it also makes it unlikely that they can achieve a decisive victory, and much more likely that they will suffer a decisive cyberspace defeat (Metz 2012). In comparison, a state such as the PRC, Russia, or Iran, is able to leverage the elements of national power and is far more likely to achieve decisive victory than suffer defeat. Cyberspace, and specifically social-media operations, can and will be used by America's current and potential adversaries to achieve their desired ends. The United States' ability to interpret, act, and mitigate these threats is crucial to maintaining its status as the preeminent super power. As Clausewitz noted,

> The general unreliability of all information presents a special problem in war: all action takes place, so to speak, in a kind of twilight, which like fog or moonlight tends to make things seem grotesque and larger than they really are. (Clausewitz 1989)

Modern military operations are still susceptible to this tenet, with the added effect that tactical missteps may have far-reaching strategic implications. Within social media lies the ability to gain a better perspective than the commanders of Clasuewitz's era, and to achieve significant operational objectives at a fraction of historical costs in terms of personnel, materiel, and treasure. Conversely the same social-media factors that allow for the rapid exploitation of an advantage can quickly be turned against the unskilled, unwary, or egotistic operator. Additionally, the United States must continue to develop a cohesive IIO strategy that, in accordance with the U.S. Constitution and federal laws, allows for the defence of the homeland and international interests, while preventing states such as the PRC from using America's own laws against it.

References

Arquilla, J, Ronfeldt, D, Zanini, M 1999, *Networks, netwar, and information-age terrorism*, Naval Post Graduate School, Graduate School of Operational Information Sciences, Monterey, CA, U.S.A.

Brazzoli, M 2007, 'Future prospects of information warfare and particularly psychological operations', *South African Army Vision 2020*, Institute for Security Studies, Pretoria, ZA.

Caldwell, W, Murphy, D, Menning, A 2009, 'Learning to leverage new media: the Israeli Defense Forces in recent conflicts', *Military Review*, May-June 2009, pp. 2-9.

Clausewitz, C 1989, *On war*, Princeton University Press, New York, U.S.A.

Crowell, R 2010, *War in the information age: a primer for Cyberspace Operations in 21st century warfare*, Naval War College, Newport, RI, U.S.A.

Fathi, N 2009, 'In a death seen around the world, a symbol of Iranian protests', *The New York Times*, June 23, viewed 1 March 2017, <http://www.nytimes.com/2009/06/23/world/middleeast/23neda.html >.

Gregory, P 2014, 'Inside Putin's campaign of social media trolling and faked Ukrainian crimes', *Forbes*, 11 May, viewed 8 August 2015,<http://www.forbes.com/sites/paulroderickgregory/2014/05/11/inside-putins-campaign-of-social-media-trolling-and-faked-ukrainian-crimes/>.

Kalb, M 2007, 'The Israeli-Hezbollah War of 2006—the media as a weapon in asymmetrical conflict', *Harvard International Journal of Press/Politics*, vol. 12, no. 3, pp. 43-66.

Metz, S 2012, 'The Internet, new media, and the evolution of insurgency', *Parameters*, Autumn, pp. 80-90.

Murphy, D 2010, 'Attack or defend: leveraging information and balancing risk in cyberspace', *Military Review*, May-June, pp. 88-96.

Qiao, L & Wang, X 1999, *Unrestricted warfare*, FBIS Translation, PLA Literature and Arts Publishing House, Beijing, P.R.C.

Taia Global Inc. 2015, *Russian Federal Security Service (FSB) Internet operations against Ukraine*, viewed 8 August 2015, <https://autoblog.postblue.info/autoblogs/lamaredugoffrblog_6aa4265372739b936776738439d4ddb430f5fa2e/media/e69ef19e.FSB-IO-UKRAINE.pdf>.

Tucker, M 2015, 'Russia launches next deadly phase of hybrid warfare on Ukraine', *Newsweek*, 31 March, viewed 8 August 2015, <http://www.newsweek.com/2015/04/10/russia-launches-next-deadly-phase-hybrid-war-ukraine-318218.html>.

United States Department of the Army 2012, *ADRP 3-0: Unified Land Operations*, Washington, D.C., U.S.A.

Van Niekerk, B & Maharaj, M 2013, 'Social media and information conflict', *International Journal of Communications,* vol. 7, pp. 1162-84.

——, Pillay, K & Maharaj, M 2011, 'Analyzing the role of ICTs in the Tunisian and Egyptian unrest from the Information Warfare perspective', *International Journal of Communications,* vol. 5, pp. 1406-16.

Cyber-Sophistication Assessment Methodology for Public-Facing Terrorist Web Sites

DC Derrick[1], GS Ligon[2], M Harms[3], W Mahoney[1]

[1] *School of Interdisciplinary Informatics*
University of Nebraska at Omaha
Omaha, Nebraska, U.S.A.
E-mail: dcderrick@unomaha.edu; *wmahoney@unomaha.edu*

[2]*Department of Management*
University of Nebraska at Omaha
Omaha, Nebraska, U.S.A.
E-mail: gligon@unomaha.edu

[3]*Department of Psychology*
University of Nebraska at Omaha
Omaha, Nebraska, U.S.A.
E-mail: mharms@unomaha.edu

Abstract: *Cyber technologies are becoming an ever-increasing part of the portfolios of Violent Extremist Organisations (VEO). Terrorist groups use these technologies in a variety of ways, such as group decision-making, cyber-facilitated financing, broader recruitment, and propaganda dissemination. However, evaluating the actual cyber capabilities of covert organisations cannot be accomplished through conventional channels. In this study, a methodology is developed and piloted in order to rate the source code supporting public-facing web pages of terrorist organisations as a proxy for assessing the cyber-sophistication capabilities of those organisations. The research team applied this methodology to a sample of VEOs. First, web pages for various organisations were discovered and evaluated to ensure their credibility. Next, a list of hyperlinks reflecting each organisation's current domain was compiled with source code for each domain being evaluated. Finally, the research team used a hybrid-coding scheme, developed from work done on evaluating the dark web, to assess the cyber sophistication of each domain. This technique allowed researchers to assess each of the sampled organisation's technical capabilities and overall cyber sophistication. Using the coding scheme, it is shown that, in this sample, al-Shabaab, Jamaal Ansharut Tauhid, and al-Qassam (Hamas militant arm English site) are the most sophisticated. The article concludes by discussing implications and offering future directions.*

Keywords: *Rating Systems, Violent Extremist Organisations (VEOs), Cyber Sophistication, Web site Assessment*

Introduction

Violent Extremist Organisations (VEOs) have posed security challenges for decades. However, in the modern era, with the advent of more lethal weapons, global mobility, and improved communication media, the span and impact of these groups has grown from regional to worldwide (Ligon, Harms & Derrick 2015). These technologies have increased VEO lethality and messaging reach (Derrick *et al.* 2016). Recently, terrorist groups have turned to cyber technologies to facilitate their missions and increase their scope. In short, cyber technologies are becoming an ever-increasing part of terrorist portfolios (Denning 2010). Terrorist organisations use these technologies in a variety of ways, such as group decision-making; cyber-facilitated financing, knowledge, and skill acquisition; and overt espionage/aggressive cyber acts. However, evaluating the cyber capabilities of covert organisations cannot be accomplished through conventional channels (Zelin 2015). This challenge motivated the authors' work to develop a methodology that uses the code underlying public-facing web pages as a proxy for assessing an organisation's cyber sophistication and capabilities.

VEOs use cyber technologies as other groups do, for communication and decision support. Most of the cyber technologies used by VEOs fall into the realm of Computer-Supported Cooperative Work (CSCW). According to Wilson (1991), 'CSCW' is a generic term, which combines the understanding of the way people work in groups with the enabling technologies of computer networking, and associated hardware, software, services, and techniques. For example, the pursuit of financial goals to maintain the organisation plays a key role in a VEO leader's decision-making (Rudner 2010); and, in order to acquire funding, VEOs must either look for large state actors or fundraising activities. According to the Congressional Research Service and the U.S. Department of State (Sullivan & Beittel 2009; 2015), Cuba, Iran, Libya, North Korea, Syria, and Sudan are countries that are known to support terrorism by supplying financial support, weapons, and other resources. However, another means by which cyber may be used to obtain financing is through online fundraising. Poorly funded groups can rapidly gather donations and other income through online fundraising or online businesses that may or may not be covert in their ties to the organisations. Specifically, VEOs have leveraged existing Internet technology to conduct operations, recruit members, solicit financing, and facilitate strategic objectives during conflict (Denning 2010; Weimann 2004). These researchers discovered that several of the more sophisticated web sites had a means for financial contributions to be collected and that they often tried to invoke sympathy for their cause to encourage donations.

Besides financing, cyber technologies are used in VEO human resource functions in at least three ways: recruiting of individuals with the desired skills and knowledge, acquiring knowledge through tutorials and information search and sharing, and facilitating knowledge transfer and training (Hunter *et al.* 2017). There is substantial evidence that a variety of technologies are highly consumed during radicalisation in an online format, particularly given the unregulated nature and ease of dissemination when compared to traditional media (Heath & O'Hair 2008). A report from the International Centre for the Study of Radicalisation and Political Violence (2009) focuses on terrorist organisations using the Internet to recruit members. Edwards and Gribbon (2013) also provide some insightful case studies of terrorists who have used the Internet as a catalyst in the radicalization process. They point out that, despite the global view of jihadists, many are still strongly rooted in Arabic web sites and online forums. They also point out the role that jihadi forums play in these recruiting efforts.

Additionally, VEOs use communication technologies to provide significant opportunities for anonymous online participation (for example, private chat rooms). The mass of violent images related with the greater anonymity found on violent ideological group web sites parallels what has been observed among violent ideological groups in non-virtual, real-world settings. For example, in a field study examining how the use of anonymity relates to levels of violence in Irish Republican Army (IRA) punishment squad attacks, Silke (2003) found that when groups wore matching versus dissimilar identity-concealing masks they were more likely to commit acts of atrocious violence. Moreover, violent group web sites foster a sense of moral righteousness through feelings of group superiority. Finally, there have been cases of overt, aggressive cyber acts by terrorist groups. These acts range from cyber vandalism, to Denial of Service (DoS), to espionage. For example, Boko Haram successfully hacked into the Nigerian Secret Service personnel records, which included names, addresses, banking records, and other confidential information (Baken 2013).

Given the prevalence of cyber technology use by VEOs, it is necessary to have some consistent methodology for gauging their cyber sophistication. 'Cyber sophistication' is an ambiguous expression for a complex and indistinct concept. Essentially, as the name implies, cyber sophistication is an attempt to define the ability of an online actor to operate in cyberspace and on the Internet; it delineates what sort of capabilities are necessary to conduct various cyber operations. This paper offers an approach for estimating the cyber sophistication of a VEO using the technology on its public-facing web page and performs comparisons on a sample of actual VEO public-facing web sites.

Background

Research has suggested that recipients of online communication have perceptions about the cyber medium that has implications for their opinions and openness to the message sender, above and beyond the actual content of the communication (Cebi 2013; Heisler & Crabill 2006; Huizingh 2000). These perceptions are related to the credibility of the information contained on the web site, and may be a function of explicit or implicit design characteristics (Fogg 2003; Rains & Karmikel 2009). In non-violent ideological organisations, the credibility of their web sites has been shown to influence attitudes toward the organisations and their missions (Long & Chiagouris 2006). Therefore, it is important to note that web site credibility refers to a more abstract 'perception' of credibility, rather than an objective measure of the product, person, or process (Fogg *et al.* 2001; Spinks 2009). Perceived credibility results in the evaluation of several different aspects of an organisation simultaneously—primarily trustworthiness and expertise. However, these researchers were primarily interested in more objective metrics of cyber sophistication that can be examined from the web pages as a proxy measure of VEO cyber sophistication. Therefore, as described in the following sections, the expertise of web sites was evaluated by applying a comprehensive rating system drawing from diverse work evaluating covert or concealed web domains (Chen 2012; Patil, Manwade & Landge 2012).

Given the high amount of variability that exists across VEOs in cyber usage, rating an organisation's public-facing web page provides an indication of that organisation's cyber sophistication and capabilities that has implications for the threat it poses. For organisations whose violent activities preclude outsourcing programming via legal channels, public-facing sites provide a certain level of knowledge that indicates the level of cyber sophistication

organisations likely possess. To most visitors, a public-facing site provides an exterior view that is evaluated based on its aesthetics and the displayed content. This type of evaluation was recently highlighted in an assessment of credibility and persuasive components of ideological organisations' web pages (Dunbar *et al.* 2014). However, an analysis of the source code of each site provides a more objective indication of the capabilities of the site designers, which was the focus of this research. For instance, a site that uses only recycled, open-source scripts to create a web page indicates that the developers are at a rudimentary level of programming ability and likely possess little knowledge as to the intricacies of programming and web development. The site may look visually appealing to the visitor, but it does not require a high-level of expertise to produce.

Therefore, although an aesthetically pleasing site is more likely to attract visitors and new recruits than either a plain or visually displeasing site, it is not a robust measure of sophistication. Some organisations may benefit from an understated appearance in order to deter unwanted users from discovering covert activities and capabilities on the domain or to prevent adversaries from identifying the web page as a communication hub for the organisation. A site may look unsophisticated and poorly produced, but if the source code behind the site is complex and original, this implies the developers likely chose for their site to appear displeasing in order to prevent detection from state actors or casual users. For instance, organisations such as al-Qaeda Central frequently have their web pages suspended or shut down by state actors and other opposing individuals, and so a more basic site built with original code would be easier to replicate on another domain following the suspension of their current sites. These factors were all important considerations in the evaluation of the credibility and sophistication of a site.

With the increased use of the Internet on mobile devices, having a web page that can scale properly to be viewed via a mobile device is becoming more desirable. A dynamic web site is a good indicator of a greater understanding of web development because the web site creator would need advanced knowledge on how to shrink the content to fit the window. The other option is to have two different layouts, one mobile and one for a desktop, so that when the web site is used with a mobile browser, the site launches the mobile version of the web site. Having both layouts is indicative of a competent design team with a fairly high level of cyber sophistication.

JavaScript is used for advanced control of how the web site interacts with the user, and the new rating methodology aids in the assessment of weighted attributes, such as this one. In order to make a web site dynamic, JavaScript would be used to help control the web site as it shrinks and expands to different window sizes. Sites that are easy to use and are interactive often utilise JavaScript to make this flexibility possible.

In addition, web sites housing a user feedback system show an advanced understanding of web site creation. There are various feedback systems, each requiring different levels of expertise. A forum system requires the most experience to implement, since many users can post to multiple threads at any time and can display new posts when they are created. An in-site feedback system that allows a user to send feedback through the web site requires a moderate understanding of web site creation, since the developer does not have to account for displaying the feedback for users to see. The use of email feedback requires the lowest understanding of web site creation;

the administrator only has to have an email address for a user to send emails to, and the developer does not have to account for any user feedback. Even if a web site looks pleasing, it may have been created in a way that shows a lack of knowledge. The developer could create a web site, not by using forums and tables, but by using lists, which could allow the creation of a web site with less original code. While there is no 'wrong' way to create a web site, there are certainly preferred methods of doing so. An organisation using these methods demonstrates a higher competency at web design, and this indicates a much greater cyber expertise than an organisation using outdated and inefficient design processes.

Method

The methodological approach used in this study is based on content-analysis techniques from observational research. In order to conduct the content analysis, it was important to first identify the content/unit of analysis and then develop an objective and repeatable way to assess the content. For this research, the content/unit of analysis is the underlying source code of a web site, and the construct that researchers are trying to assess from this content is the unseen but inherent cyber sophistication of the developer of the web page. **Figure 1**, below, shows the flow of the method. Each step will be discussed in the following sections.

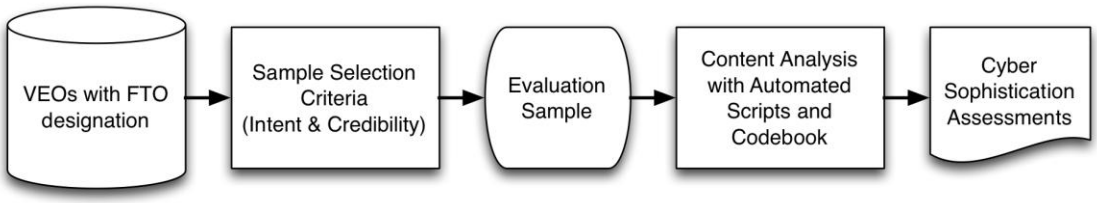


Figure 1: Method for assessing VEOs' public web pages

Sample Identification

To conduct the content analysis for assessing cyber sophistication using public-facing web pages, the researchers used stratified sampling to identify a small, representative sample of notable VEOs designated as Foreign Terrorist Organisations (FTO) by the U.S. Department of State (2015). Designation as an FTO is an indicator of a high-threat organisation because VEOs must go through a rigorous evaluation process in order to receive this designation. The process includes an analysis of VEO funding lines, endorsements by state adversaries, and demonstrated sophistication of attacks (Ilbiz & Curtis 2015; Ligon, Harms & Derrick 2015). The sample was defined based on historical activity and current and emerging threats. First, all VEOs currently designated by the Department of State as FTOs were examined in order to provide a comprehensive sample of high-threat organisations. Second, VEOs were identified that have emerged within the past ten years (for example, al-Shabaab) and have demonstrated cyber capabilities and a history of destructive performance. Finally, historic VEOs were identified that have evolved to use cyber technologies to execute their missions. This last set of historical VEOs (such as the Kurdistan Workers Party) allowed for the examination of the validity and generalisability of the cyber technologies, in the context of technological advancement over time. Based upon previous research suggesting that the web site content of violent ideological groups may differ from non-violent ideological groups despite similar belief systems (Connelly *et al.*

2015; Dunbar *et al.* 2014), two notable non-violent ideological organisations were included as a control for the proposed rating system and to test biases related to the ideological nature of these groups. Finally, due to the covert nature of their activities, some VEOs (for example, Boko Haram) choose to avoid a public web presence. Because the goal of this research was to investigate the sophistication of public-facing web pages of VEOs as a means to market their organisations to potential investors, recruits, and sympathizers, the sample did not include VEOs that did not have a public web site.

After identifying a potential sample of VEOs on which to test this methodology for content analysis, a series of steps were followed to select and to evaluate the primary web sites used by the organisation for news releases, recruitment, communication, and coordination activities. Since there are groups that often desire to mimic, imitate, or 'impersonate' the web sites of the VEOs under examination, assessing the credibility of these web sites was a critical early stage in the proposed rating process. The research team evaluated each possible web site regarding its legal standing, as declared in current U.S. anti-terrorism statutes. To incorporate these statutes into the rating process, the team created criteria for each site before it could be selected and evaluated for credibility (see **Table 1**, below). As Fogg *et al.* (2001) contend, the evaluation of a web site must consider multiple factors simultaneously to assess credibility. The second portion of **Table 1** illustrates the multiple factors the team considered when evaluating each web site. Web sites had to meet at least one of the first three criteria for 'intent', and all of the following three criteria for 'credibility', to be considered for selection.

	Criteria
Intent	Establishes and maintains Internet web sites or posts detailed information on such web sites with the specific intent to recruit persons to join terrorist organisations (as designated under Sec. 219 of the Immigration and Naturalization Act), or with the specific intent to recruit persons to engage in acts of violence against the United States or citizens of the United States
	Establishes and maintains Internet web sites or posts detailed information on such web sites with the specific intent to encourage violent attacks against the United States government or its citizens, to include, but not limited to violations of those United States Code sections set forth in 18 U.S.C. § 2339A(a),
	Establishes, maintains, or posts detailed information on Internet web sites with the specific intent to assist, encourage, or facilitate funding to designated terrorist organisations in violation of 18 U.S.C. § 2339B
Credibility	Attempts or conspires to do such acts as defined by paragraphs (1) through (3)
	Appears to be a credible and legitimate source in terms of, for example, grammatical proficiency, typographical errors, and visual appeal
	Has a deeper layer of web site links that can be assessed (hyperlinks that appear to be related to the VEO)

Table 1: Web site selection criteria

The 'official' web site location of VEOs is often difficult to locate. Using a variety of open source tools and databases, along with data received from Chen (2012), Patil, Manwade and Landge (2012), and other VEO subject matter experts, the team compiled a list of web sites linked to the VEOs under examination. Several organisations did not attempt to hide their Internet presence, and their web sites were easy to locate. For instance, the Jamaah Ansharut Tauhid (JAT)—the largest splinter cell of Jemaah Islamiyah—was overt in its web domain, in terms of intent, mission, and recruitment. JAT is clear about the intent of its site, and maintains it with a high level of sophistication, which contributes to the level of credibility and trustworthiness the site projects. Its domain name—http://www.ansharuttauhid.com—illustrates the overt nature and intent of its web presence (for example, using a simple web search so that most individuals with basic Internet capability could locate, access, and retrieve information from this site). The JAT also has Foreign Terrorist Organisation (FTO) designation, which means that it meets the criterion that it "establishes, maintains, or posts detailed information on Internet web sites with the specific intent to assist, encourage, or facilitate funding to designated terrorist organizations" (Williams 2007).

However, not all sites were easily located. For example, the process of locating an al-Qaeda domain was much more difficult. In part, this may be because the JAT, while a designated FTO, does not pose the same level of global threat that al-Qaeda has demonstrated. Therefore, it is less difficult for the JAT to maintain its site while avoiding removal by Internet Service Providers (ISP) or Domain Name Service (DNS) registrars. Given its global notoriety, al-Qaeda is a more frequent target for governments and other state actors attempting to deter activity by shutting down communication and recruitment outlets such as public web sites. Consequently, its domain is much more difficult to locate. Using Chen's database (Chen 2014), the team identified a former al-Qaeda domain and used this domain to track it to a live web domain. The URL http://www.h-alali.net was determined to be a live al-Qaeda web site and was selected for evaluation of its sophistication. The remaining web pages were found using a similar process.

Instruments for Assessment (Content Analysis)

Figure 2, below, shows the process used to analyse the cyber sophistication of each organisation based on its web site.

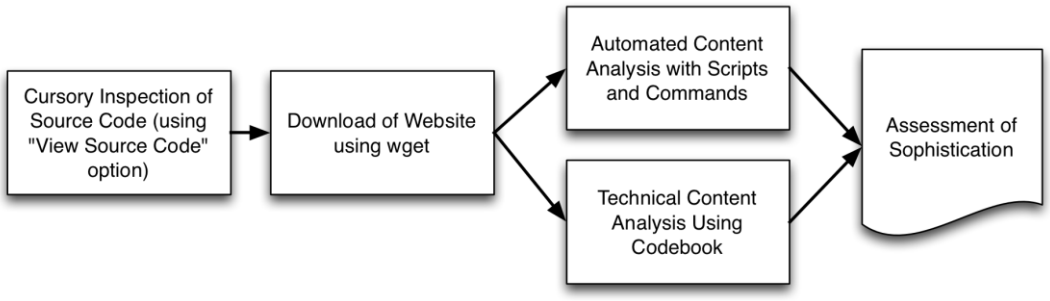


Figure 2: Process for content analysis to assess cyber sophistication

Using the source code of each target domain, the researchers examined how the organisation structured its web site. The way to access a web site's source code is to open the web site in a browser (in the context of this project, Google Chrome and Tor were used) and right click anywhere on the web page, avoiding large banners, ads, videos, and pictures. A dropdown box will appear and the user can select 'View Page Source', which will present the web page's source code in another window in the web browser. Factors evaluated included the use of tables to format the web site, the use of custom style for pages, and the use of JavaScript (whether it was original or used predefined functions). Once this information was collected, the researchers could draw conclusions about the credibility and sophistication of the web site. These conclusions were based on theoretical foundations from literature exploring web site credibility (discussed in the background section), as well as Subject-Matter-Expert (SME) assessments regarding characteristics of web sophistication and technical capabilities that more advanced programmers would know (for example, web page structure and visitor metadata). By examining how each web page was built, the researchers were able to conduct preliminary assessments regarding the expertise and capabilities of each organisation and determine the credibility of the site in question.

Although the team gained insight and important information about each organisation's capabilities, this method of obtaining a web page's source code is not sufficient to indicate the full sophistication of the organisation's web domain. For instance, in opening the source code, the ability to see the Hypertext Preprocessor (PHP) code used in the web page is lost. This is because PHP is handled on the server side of the web site and the source code only provides the result of the PHP. For example, al-Qaeda uses PHP to rapidly set up its next web site after a current site is taken down. VEO web pages are commonly targeted during deterrence efforts; relocating those web pages is important for their ability to maintain an online presence. Al-Qaeda does this by registering many domain names with different global companies. The alternate web site then lies dormant until the current domain hosting the al-Qaeda web site is taken down. The organisation separates the data from the web page by having a database containing all of the images, video, metadata, and other files it wants displayed; then in the PHP, it is able to query for the needed file and use it to rebuild the web site. Since the database is separate from the web site, the web site can be taken down without affecting the database; all that is needed to switch the web page to the new domain is to move the PHP files to the new web hosting service. This allows the organisation to move only 30 or 40 files, as opposed to 1,000 or 2,000 files, which results in a faster and more discrete establishment of the new web site. Furthermore, a single PHP file can be used with different parameters to yield a different look.

A normal web page URL appears like this: **www.domainname.com/index.php/.** A web page using the same PHP but with an extra parameter looks like this: **www.domainname.com/index.php?page=2/.**

The second example is setting a variable named page in the PHP to two, which, in the code, would prompt the same appearance for the web page, but would display the desired content for page two instead. This results in the need for fewer files to transfer following the change of web domains, which substantially decreases the time a web site is inoperable for dispersed members or sympathizers. As VEOs rely heavily on the Internet to conduct several processes critical to organisational survival (such as communication, planning, and funding), having the expertise to

manipulate web domains leads to a higher chance of success. The same process used to transfer a web page to a new domain can be used with multiple aspects of the web domain, such as the aesthetics and the images or videos displayed.

In order to obtain more information on what attributes were utilised in the creation of the web page, a different method of data extraction was necessary. To do this, the researchers used a program called "wget" to extract the code from the web domains. This program downloads all of the directed HTML text of a web site and copies through HTTP, HTTPS, or FTP protocols. Using the data downloaded by "wget", it is possible to create the web site on another server up to the level of data downloaded. Therefore, instead of using screenshots to capture the images of the selected web sites that were evaluated, the entire site could be recreated using the information from "wget". This method of data collection simplifies the analysis of the site data as the querying for terms, links, and other content can be accomplished offline and in an automated manner.

After the initial evaluation of web site source code, "wget" was used to determine several quantities of analysis for each site. This data was used primarily in two ways. The most critical use of the program was to directly download the source code of each web site analysed. As stated above, the raw source code was required in order to evaluate the sites without the limits imposed by page source data when viewing the site in a browser. All of the code was uploaded to a local server. The team used this new local copy for further analysis. Because web content was being downloaded from potential adversaries, an Information Assurance (IA) lab was used for this stage of the methodology. Consequently, the team was able to protect its resources from any computer malware that might have been included in the download. These special facilities are safely isolated from other resources so the web pages can be studied without harm to the university or to external entities on the Internet.

During the analysis and rating process, the research team captured and saved the first two levels of each web site to evaluate. These levels were selected as the most appropriate and informative for the purposes of this analysis because web pages tend to house the majority of their own original content within the first two levels of page links (Spinks 2009).

Once a web site is in the third or subsequent levels, the links tend to lead the searcher to external pages. Consequently, the material housed on these pages would not reflect the sophistication of the host server. Once the pages were downloaded from the site, the team used a suite of standard Linux tools such as "find" (which searches a directory and subdirectory for matching files), "grep" (which searches files for matching text), and "wc" (which identifies word count and which will also count matching lines) to analyse the files. For example, to determine the number of HTML links embedded within a collection of downloaded web pages, the following command can be used:

find . –name '*htm*' -exec grep href= {} \; | wc –l

The commands will start from the top-level directory, match any filename that includes the string "htm", and will then look within each of these files for the string "href=". Finally, the number of

matching strings is displayed through the word-count program. The team also used "find" to count the number of hyperlinks in the pages by using the following command:

find . -type f | wc –l

In addition, the number of file/software downloads available from the site were counted. As an example, the following command counts the number of images in certain common formats:

find . –name '*.jpg' –o –name '*.jpeg' –o –name '*.gif' | wc –l

Using similar samples commands, the number of video/audio files can be counted:

find . –name'*.mp4' –o -name '*.wav' –o -name '*.mpg' | wc –l

This information was added to each site's cyber-rating codebook and contributed to the final cyber rating. Each of these specific quantities was selected due to its contributions to the sophistication ranking of a site. Sites with few hyperlinks, file downloads, images, and audio/video files are typically more rudimentary and would not be rated on the same level of complexity and richness. While the rest of the codebook relies on weighted counts with zero or the assigned value, the data counts extracted from the "wget" download add quantifiable, comparable data to strengthen the accuracy of each sophistication rating. The analysed data pulled from "wget" does not provide a unitary rating on its own, but rather adds content richness to the greater evaluation of each target site.

Assessment

In order to complete the content analysis of the cyber-sophistication coding of each organisation, two prior methods for rating technical capabilities were performed. The two methods used to rate the technical capabilities of the organisations originate with Chen (2012), and with Patil, Manwade and Landge (2012). Both rating methods are similar in that they examine the web page's sources to make an estimation of the technical prowess of the organisation that created the pages, but different in the attributes collected and rated. The proposed rating method uses the work from both Chen (2012) and Patil, Manwade and Landge (2012) to devise a more comprehensive rating system. Various weights are assigned for features present on the web pages, and comparisons are made between organisations by the accumulated weights determined by the page content.

The team incorporated the Dark Web Attribute System (DWAS), a method for gaining insight into the technical sophistication of extremist organisations (Chen 2012). DWAS is used to analyse and compare the web sites of various groups by examining downloaded web content and looking for features, which are then scored accordingly. The method employs nine high-level attributes, each of which contains several more specific traits. For example, a high-level attribute is 'Advanced technical sophistication', which encompasses the scripting languages that might be used in the web page content. **Table 2**, below, shows the rating methodology that was developed. The high-level attributes have been omitted, but the low-level attributes have been spelled out since the high levels do not contribute to the cyber-sophistication scoring. Following the work of Chen (2012) and Patil, Manwade and Landge (2012), who also use link-count analysis to measure the relative complexity and interactivity of the site, the current analysis uses the

commands described above to determine the link count within the downloaded files. For example, HTML pages including a 'form' are more sophisticated than plain HTML, and pages containing JavaScript are more advanced than plain pages. **Table 2**, below, provides the entire sophistication rating codebook the team created using specific attributes from both Chen (2012) and Patil, Manwade and Landge (2012), including the weight assigned to each attribute. Following the development of the codebook, the team applied the rating method to the sample of VEO web domains described previously. The following section outlines the results of this application

High-Level Attributes	Low-Level Attributes	Description	Weight
Technical-Sophistication Attribute	Menu	The use of menu tag for designing the web sites	2
	Meta	The use of meta tag for designing the web sites	2.5
	Style	The use of style tag for designing the web sites	1
	Label	The use of label tag for designing the web sites	2.5
Fundamental Attribute	Form	The use of form tag for designing the web sites	1.5
	Frame	The use of frame tag for designing the web sites	2
	Table	The use of table tag for designing the web sites	2
	List	The use of lists	1
Advanced Technical-Sophistication Attribute	Java script	The use of java script language	4
	Script	The use of self-defined script language	4.5
		The use of predefined script functions	2
	Advanced HTML	The use of DHTML/SHTML	2.5
Dynamic Web Programming	Java	The use of Java language	2.5
	PHP	The use of scripting language designed for web development to produce dynamic web pages	5
	ASP	The use of Active Server Pages (ASP)	5.5
Content Richness	Flash	Banner depicting representative figure, graphical symbol or seal	1
	Image	Banner depicting representative figure, graphical symbol or seal	1
	Audio	Short phrase with religious or ideological connotation	1

	Video	Video on religion, attack, for example	1
	Music	Background music	2
Communications (User-generated content)	List	List with leader name, address, for example	2.3
	Contact	Telephone number	1.2
	Email	Email address	2.5
		Email feedback	1.75
	Guestbook	Option for users to leave information in a guestbook	1.5
Online Organisational Attribute	Comment	User is able to give feedback or ask questions to the site owner or maintainer	2.4
	Videoconference	Video clip of bombings, game, animated picture, for example	3.3
	Online forum	User is able to leave information in a forum	4.25
	Online chat	User is able to live chat	4.75
Transaction-level Interactivity		Online shop	4
		Online payment	4
		Online application form	4
Web Interactivity Attribute	Online recruitment	Invitation to join or attend meeting, interview, for example	4.5
	E-tendering attributes	Invitation & publishing the E-tendering information	4.5
Content Richness (variety and amount of information)		Hyperlinks	Raw count
		File/software download	Raw count
		Image	Raw count
		Video/audio file	Raw count

Table 2: Combined cyber-sophistication codebook

Results

The results of the analyses conducted on this sample supported the rating methodology developed, and illustrated that web sites displaying a higher level of sophistication all had similarities among them that the weighted rating system revealed. These included being dynamic, using JavaScript, accepting user feedback, having a visually appealing appearance, and organising the tags used to create the layout of the web site. According to the assessment ratings, the top three web sites ranked in order of sophistication were Jamaah Ansharut Tauhid (Jemaah Islamiyah), the control, non-violent ideological group (Hizb ut Tahrir), and al-Qassam (Hamas militant arm, English web site). It is important to note that two non-violent ideological organisations were used as a control group in this research. Both of these organisations were assessed in previous research (Ligon, Harms & Harris 2014) to have above average

sophistication relative to other non-violent ideological organisations. Consequently, they were selected as comparison web domains to assess whether the violent organisations in this study's sample were less, comparably, or more sophisticated than their non-violent counterparts.

Jemaah Islamiyah's domain was rated highest because it is dynamic (for example, it could shrink for mobile phones), everything on the page was organised and clearly delineated (it was easy to find search tools and the menu bar, for example), and it was not full of the ads that plague other VEO sites. Visually, it was also the most pleasing site. The "wget" information shows that it has the most images used from a local location, which could be because the pictures it uses for its web site are all taken and uploaded by its organisation. Jemaah Islamiyah has more hyperlinks than files downloaded, which is informative because it shows how often it uses materials gathered from external sources. It also has videos on the web site, though the "wget" ratings assessed this count as lower than the actual number. This discrepancy is a result of the video files being embedded into the web domain; the query for video count was not set up to check for embedded videos. The number of videos is important to the overall sophistication and appeal of the domain because it is a method of communication between the organisation and casual visitors. The videos were counted separately as embedded and non-embedded files, and this distinction remained consistent across the analyses; therefore, it does not adversely impact the accuracy of the ratings. The main area in which this web domain was rated lower than other organisations was in the lack of diverse language options available. Specifically, while many organisations in the sample offer the option to view their page in several languages, this domain is only available in Arabic. This does not necessarily indicate a lack of cyber sophistication; there are several reasons an organisation might only offer one language for its web site. For instance, the organisation's target audience may speak a certain language. In addition, excluding other languages from the web domain may prevent some foreign enemies or governments from locating the domain using a keyword search or search algorithm.

The control group Hizb Ut Tahir received the second highest rating because it was dynamic (viewable on a mobile device), and it offered support for two languages (English and Arabic). The site is organized in a simple and straightforward manner (for example, the viewer can easily find information). In addition, the web site allows the viewer to sort the number and type of articles and web content to view. The "wget" data indicates that it has no images or videos; this suggests that any images viewable on the site are either pulled from different web sites, saved in a database that is accessed server-side only, or no images are used. Any videos on the web site are embedded. The site also has more hyperlinks than files housed on its domain. This is expected, given that it has no images housed on its web page. When the site is viewed in English, it links to a separate, English web site. In other words, rather than offering the same web page translated, the domain sends the viewer to an entirely new web site. There are several reasons a web site designer may choose this option. First, linking to a separate web page sponsored by the organisation offers more control than a translating service. Second, he or she may lack the expertise to host two languages on the same site. It should be noted that the English site was unavailable to load at the time the ratings were conducted (for example, it was cross-listed as being under review or edits). This reinforces the previous assertion that an organisation may choose to omit web pages in other languages so it is more difficult to trace or find using a search program.

Al-Qassam English site was rated third because it also is dynamic (can shrink to be viewed on a mobile device), it has support for multiple languages (there are seven languages selectable at the time of the ratings), a scrolling banner is available on all of the site's associated pages, and it is easy to navigate and find specific content. The "wget" shows that it has the most hyperlinks (38,452) and only 25 of those hyperlinks are images. This indicates that the majority of the images originated on pages external to the web domain itself. The web site had no videos, but offered a location showing live events. This function was inoperable at the time of ratings, so it was not included in their overall score. (It is unclear whether this option was down at the time of ratings or is simply a simulated function.) While the web page did host several images itself, there are significantly fewer images hosted on-site than what the web page offers, indicating that the images housed within the domain were most likely associated with content that does not change or update.

The ratings provided information about the comparison between the aesthetic 'face' appeal of the web page relative to its content. Specifically, some web pages offer more options (such as multiple languages and a dynamic interface) that contribute to the overall credibility and appeal of the web page, but may indicate less information about the diversity and sophistication of the content. The following describes the results of the selected, combined ratings Chen (2012) and Patil, Manwade and Landge (2012) proposed as a proxy method to estimate cyber sophistication beyond web site credibility. The graph in **Figure 3**, below, illustrates these ratings on the organisations and web pages sampled.

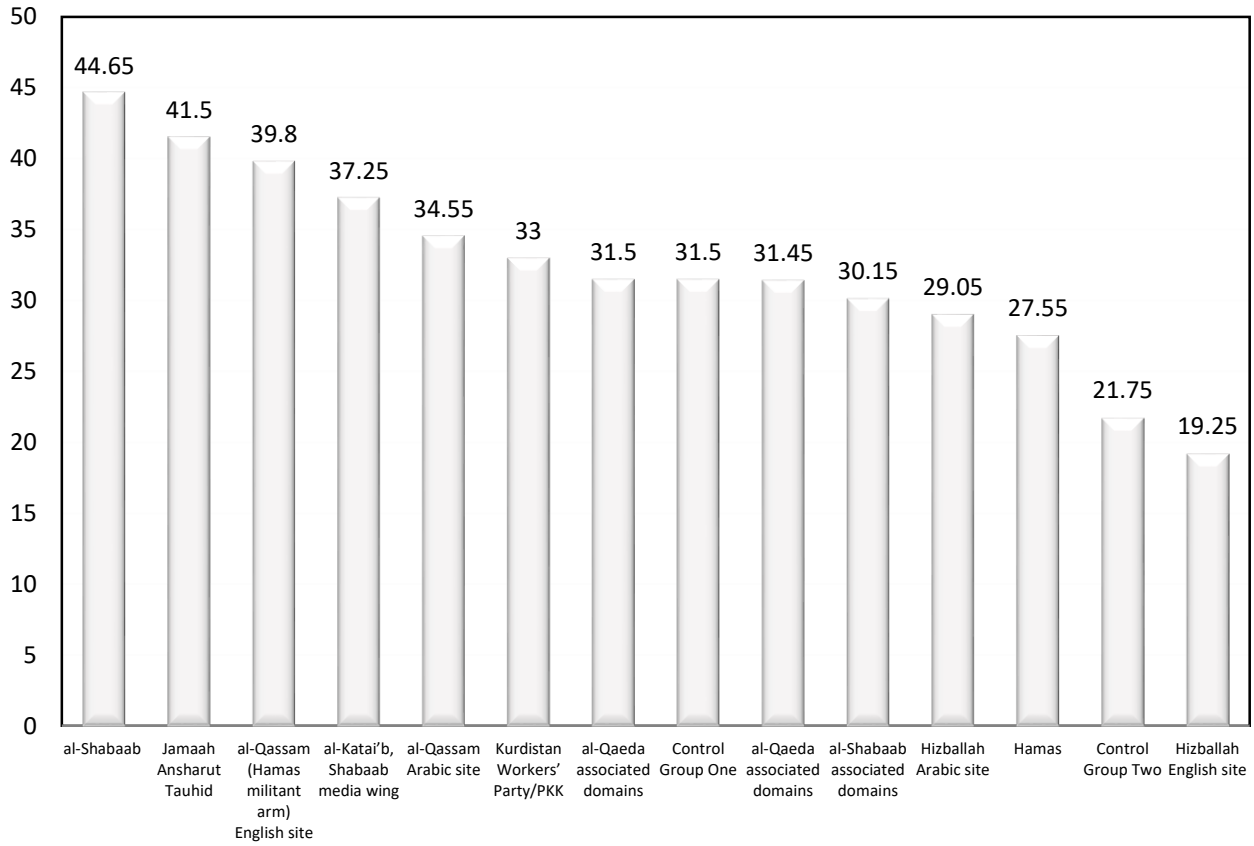


Figure 3: Bar graph showing cyber-sophistication ratings

Discussion

Based on the results of this study, the sample of VEOs varies in terms of cyber sophistication; public-facing web sites provide a compelling artefact of such differentiation. Using a multi-source method and an empirically validated theoretical framework, the researchers were able to identify important differences among the source code, credibility, and dynamic interfaces of a variety of English and Arabic language VEO web sites. The results revealed strong support for the utility of this rating method as an index of covert capabilities not easily measured through other, conventional means. Accordingly, several implications for research and practice have been identified based on the results of this study.

First, this study demonstrated a repeatable methodology for classifying cyber sophistication of an organisation using public-facing web sites. Although this is only a proxy measure, it does provide insight into the cyber capabilities of various VEOs. The methodology for classification is a useful tool for researchers and policy makers for evaluation of complex data. Researchers can use this method as a benchmark by which to compare various organisations that offer limited access. This analysis is useful for those involved in information warfare to the extent that it comprehensively and succinctly describes a set of important criteria for VEO sophistication and it provides a repeatable process. This paper represents the first time that a combined codebook from various expert researchers and deeper analysis tools have been brought to bear to form a comprehensive stratification of VEO cyber sophistication. It offers a way to account for the variability in web site construction and to compare organisations. While the FTO designation is based on the threat an organisation poses to the U.S. homeland, these threats are determined based upon physical actions the organisations take, such as attacks or statements made by their leaders. However, as technology advancements allow for global engagement via cyber pathways, the threat posed by these organisations advances from physical to one that is more virtual and far-reaching. The methodology proposed and tested here uses the credibility and code supporting public-facing web pages as a proxy for measuring the cyber-sophistication capabilities of VEOs.

Second, this methodology has been applied to VEO organisations, which demonstrates that these comparisons across multiple VEOs give some indication of sophistication. This offers insights into the similarities and differences between VEOs, their sophistication, and cyber-presence strategies (such as JAT vis-à-vis Al Qaeda). Some clarity into the characteristics of VEOs and their persistent cyber presences was also provided.

Third, this research methodology can be applied to other organisations and can offer insights into strategy, sophistication, and cyber ability. As cyber is an increasingly used resource among extremist organisations, having a mechanism to assess varying levels of expertise in using such CSCW is critical to early identification of threats.

Conclusion

In this research, a stratified sample of VEOs reflecting a range of sustainability and organisational sophistication was used as a pilot sample to test the rating methodology developed through this research. The first step in the methodology was to identify the public-facing web sites of the target organisations. After the initial research was concluded to locate target sites, the research team verified that each selected site was credible and could be linked back to the target organisation. Once the credibility of each site was verified, the team used the software program

"wget" to retrieve the source code two levels deep for each site. This allowed the team, if further analysis was necessary at a future time, to recreate the entire web site in an identical state that was rated initially. Using the source code of each web site, the team ranked each site per the specific attributes selected from the rating schemes of Chen (2012) and Patil, Manwade and Landge (2012). These attributes allowed the team to rank each site on a scale of cyber sophistication, which provides easy comparisons between each site. To further refine the cyber-sophistication ratings, the team employed several analytic functions to determine content richness of each site. Each of these factors contributed to the overall cyber-sophistication rating of each web site. This methodology holds promise to assist in evaluating the cyber capabilities of emerging VEOs and can be used to estimate the potential cyber threats and activities of organisations of interest.

While the methodology provides insight into an organisation's cyber sophistication, it is only a proxy measure. The next step of analysis should be the pairing of the sophistication measure with actual real-world cyber actions. For example, it would be interesting to study the correlation of cyber sophistication with success in fundraising activities, lethality, or the recruitment of new followers. In addition, the methodology is currently content-agnostic (that is, it measures the technical structure, not the content). It might prove useful to evaluate the content (in terms of persuasive messaging, quality of images and videos, and grammar) in conjunction with the cyber sophistication of the site. These factors almost certainly contribute to success in recruitment/radicalization, and may provide an overall picture of organisational sophistication. Finally, it might be interesting to conduct a longitudinal study of various web sites in order to watch the development of cyber sophistication to determine how organisational cyber capabilities evolve and if they relate to organisation metrics of ideological success.

Acknowledgements

Two research assistants greatly assisted this research project by gathering web sites, coding data, and analysing results. Jeremiah Wilt and James Sadowski are acknowledged for their contributions to this effort.

References

Baken, D 2013, 'Cyber warfare and Nigeria's vulnerability', *E-International Relations*, 3 November 2013.

Cebi, S 2013, 'Determining importance degrees of website design parameters based on interactions and types of websites', *Decision Support Systems*, vol. 54, pp. 1030-43.

Chen, H 2012, *Dark Web: exploring and data mining the dark side of the web*. Springer, New York, NY, U.S.A.

——2014, 'The Dark Web Project and Forum Portal', viewed 28 November 2013, <http://ai.arizona.edu/research/terror/>.

Connelly, S, Dunbar, N, Jensen, M, Griffith, J, Taylor, W, Johnson, G, Hughes, M & Mumford, M 2015, 'Social categorization, moral disengagement, and credibility of ideological group websites', *Journal of Media Psychology*, vol. 28, pp. 16-31.

Denning, D 2010, 'Terror's web: how the Internet is transforming terrorism', *Handbook of Internet crime*, eds. Y Jewkes & M Yar, Willan Publishing, Devon, UK, pp. 194-213.

Derrick, D, Church, S, Sporer, K & Ligon, G 2016, 'Social media, open architectures and ideological rationality', *Hawai'i International Conference on System Sciences (HICSS)*, 5-8 January 2016, Kauai, HI, U.S.A.

Dunbar, N, Connelly, S, Jensen, M, Adame, B, Rozzell, B, Griffith, J & O'Hair, H 2014, 'Fear appeals, message processing cues, and credibility in the websites of violent ideological, and nonideological groups', *Journal of Computer-Mediated Communication*, vol. 19, pp. 871-89.

Edwards, C & Gribbon, L 2013, 'Pathways to violent extremism in the digital era', *The RUSI Journal*, vol. 158, pp. 40-7.

Fogg, B 2003, *Persuasive technology: using computers to change what we think and do*, Morgan Kaufmann Publishers, San Francisco, CA, U.S.A.

——, Marshall, J, Laraki, O, Osipovich, A, Varma, C, Fang, N, Paul, J, Rangnekar, A, Shon, J, Swani, P & Treinen, M, 2001, 'What makes web sites credible? A report on a large quantitative study', *Proceedings of the SIGCHI conference on human factors in computing systems, April 2001, Seattle*, eds. M. Beaudouin-Lafon & R. Jacob, Seattle, WA, pp. 61-8.

Heath, R & O'Hair, H 2008, 'Terrorism: from the eyes of the beholder', *Terrorism: communication and rhetorical perspectives*, eds. D O'Hair, R Heath, K Ayotte & G Ledlow, Hampton Press, Cresskill, NJ, U.S.A, pp. 43-66.

Heisler, J & Crabill, S 2006, 'Who are "stinkybug" and "Packerfan4"? Email pseudonyms and participants' perceptions of demography, productivity, and personality', *Journal of Computer-Mediated Communication*, vol. 12, pp. 114-35.

Huizingh, E, 2000 'The content and design of web sites: an empirical study', *Information & Management*, vol. 37, pp.123-34.

Hunter, S, Crayne, M, Shortland, N & Ligon, G, 2017, 'Recruitment for terrorism', *American Psychologist*, in press.

Ilbiz, E & Cutris, B 2015, 'Trendsetters, trend followers, and individual players: obtaining global counterterror actor types from proscribed terror lists', *Studies in Conflict and Terrorism*, vol. 38, pp. 39-61.

International Centre for the Study of Radicalisation and Political Violence 2009, 'Countering online radicalisation: a strategy for action', viewed 28 November 2013, <http://www.icsr.info>.

Ligon, G, Harms, M & Derrick, D 2015, 'Lethal brands: how VEOs build reputations', *Journal of Strategic Security*, vol. 8, pp. 27-42.

——, Harms, M & Harris, D 2014, *Organizational determinants of violence: introducing the L.E.A.D.I.R. database and codebook*, National Consortium for the Study of Terrorism and Responses to Terrorism (START), College Park, MD, U.S.A.

Long, M & Chiagouris, L 2006, 'The role of credibility in shaping attitudes toward nonprofit websites', *International Journal of Nonprofit and Voluntary Sector Marketing,* vol. 11, pp. 239-49.

Patil, G, Manwade, K & Landge, P 2012, 'A novel approach for social network analysis & web mining for counter terrorism', *International Journal on Computer Science & Engineering*, vol. 4, pp. 1816-25.

Rains, S & Karmikel, C 2009, 'Health information-seeking and perceptions of website credibility: examining web-use orientation, message characteristics, and structural features of websites', *Computers in Human Behavior,* vol. 25, pp. 544-53.

Rudner, M 2010, 'Hizbullah terrorism finance: fund-raising and money-laundering', *Studies in Conflict & Terrorism*, vol. 33, pp. 700-15.

Silke, A 2003, *Terrorists, victims and society: psychological perspectives on terrorism and its consequences*, John Wiley & Sons Ltd, London, UK.

Spinks, B 2009, 'Assessing perceived credibility of web sites in a terrorism context: the PFLP, Tamil Tigers, Hamas, and Hezbollah', PhD Dissertation, University of North Texas.

Sullivan, M & Beittel, J 2009, *Latin America: terrorism issues*, Congressional Research Service Washington, D.C, U.S.A.

U.S. Department of State 2015, 'Foreign terrorist organizations', press release, retrieved from <http://www.state.gov/j/ct/rls/other/des/123085.htm>.

Weimann, G 2004, 'www.terror.net: how modern terrorism uses the Internet', *United States Institute of Peace*, Special Report 116, Washington, D.C., U.S.A.

Williams, A 2007, 'Prosecuting website development under the material support to terrorism statutes: time to fix what's broken', *NYU Journal of Legislation and Public Policy*, vol. 11, pp. 365-403.

Wilson, P 1991, *Computer-supported cooperative work: an introduction,* Kluwer, Oxford, UK.

Zelin, A 2015, 'Picture or it didn't happen: a snapshot of the Islamic State's official media output', *Perspectives on Terrorism,* vol. 9, pp. 85-97.

Integrating Cyber-Intelligence Analysis and Active Cyber-Defence Operations

EJ Mandt

Utica College
Utica, New York, U.S.A.
E-mail: ermandt@utica.edu

Abstract: *The world is experiencing a continuous state of cyber insecurity. Despite continual development of cyber-security technology, the power balance between attacker and network defender has remained largely unchanged. While the cyber-security community is attempting to change this stalemate by developing active cyber-defence tactics and emphasizing cyber-threat intelligence, these efforts remain incomplete. A synthesis of the Diamond Model of Intrusion Analysis and Robert Lee's Active Cyber Defense Cycle will demonstrate that integrating structured intelligence-analysis techniques into active cyber-defence operations has the potential to alter the power balance between attacker and defender.*

Keywords: *Cyber Security, Cyber Intelligence, Active Cyber Defense, Diamond Model of Intrusion Analysis, Cyber*

Introduction

Beginning with the Morris worm in 1988 (Eichin & Rochlis 1989), the information technology community has experienced a series of increasingly advanced attacks on computers and computer networks. The last decade, in particular, has witnessed a series of cyber security incidents resulting in major data breach after major data breach. All sectors of cyberspace have been affected, ranging from the 2008 compromise of Department of Defense computers by a suspected foreign intelligence organisation (Lynn 2010), to the 2013 breach of the payment system used by the retail store Target (Target n.d.), to the 2015 breach of U.S. health insurer Anthem Inc. (Krebs 2015), and the breach of the Democratic National Committee during the 2016 election cycle (U.S. Department of Homeland Security 2016).

This same period of time witnessed significant research and development of cyber-security technologies and a dramatic growth in commercial cyber security products on the market. Beginning with a proposal for a real-time Intrusion Detection System (IDS) in the mid-1980s (Denning & Neumann 1985) and Symantec's release of Norton anti-virus software in 1991 (Krebs 2003), the cyber security community has conceived and produced an increasingly broad array of sophisticated technologies to combat malicious cyber activity. The last 25 years have seen impressive strides in firewalls, anti-virus software, IDSs, Intrusion Prevention Systems (IPS), Security Information and Event Management (SIEM) software, and other security products.

What the cyber-security community has not seen, however, is a change in the power balance between attacker and defender. The technologies and technical capabilities of both attacker and defender have evolved rapidly and in parallel. The balance of power on the cyber playing field has been left largely unchanged. A continuous state of insecurity continues to exist in cyberspace. This has led the cyber-security community to recognize the need to evolve beyond a passive approach to cyber-defence and begin embracing the concept of an intelligence-driven active cyber defence. However, the cyber-security community faces challenges in realizing the transition from passive cyber defence to an intelligence-driven active cyber-defence posture.

One challenge facing the cyber-security community is developing a deeper understanding of what cyber intelligence is, how it is produced, and how it can best be used. Current approaches to cyber-threat intelligence rely heavily on automated processes that seek to minimize reliance on human involvement. Undoubtedly, cyberspace generates large volumes of data at speeds that necessitate the use of advanced technology to automate certain intelligence-related processes. However, the experience of the IT community over the last 25 years demonstrates that technology alone has not solved the problem of cyber-insecurity. Studies and surveys by the Intelligence and National Security Alliance and the SANS Institute suggest that the cyber-security community possesses a much deeper reservoir of knowledge and appreciation for technology than for intelligence-analysis processes. While the cyber-security community has begun to embrace the use of intelligence, much work remains to be done for cyber intelligence to mature into a fully developed discipline.

A second challenge is the unresolved nature of what precisely constitutes active cyber defence and what the legal basis is for an active cyber defence. The idea of active cyber defence has existed for several years and is relatively well defined within the U.S. government and Intelligence Community. However, consensus on the meaning and appropriate use of active defence does not yet exist in the broader cyber-security community. Frustration over continuing lawlessness in cyberspace has reportedly led some to adopt active 'hack back' cyber defences (Timberg, Nakashima & Douglas-Gabriel 2014; McFarlin 2015). Other voices argue that hacking back represents an ill-advised and "irresponsible" approach (McGraw 2013).

As a result of the emerging nature of cyber-threat intelligence and the unresolved issues related to active cyber defence, academic work has tended to focus internally on concerns specific to each area of study. This article seeks to begin to bridge the gap between these areas of study by offering a conceptual framework for understanding the relationship between cyber-threat intelligence and active cyber-defence activity. The conceptual framework integrates the Diamond Model of Intrusion Analysis (Caltagirone, Pendergast & Betz 2013) with the Active Cyber Defense Cycle (Lee 2015a, 2015b). For the purposes of this paper, active cyber-defence activities will be limited to actions taken on one's own network or a network on which one is authorized to operate, a definition which is consistent with Defensive Cyber Operations-Internal Defensive Measures (DCO-IDM) as described in the U.S. Department of Defense's *Joint Publication 3-12 (R) Cyberspace Operations* (2013a).

This paper is organised into sections that address the following topics: a brief discussion of intelligence-analysis basics built on the knowledge and experience of the intelligence community; a survey of the recent history of intelligence analysis; an examination of existing

cyber-threat intelligence technologies and practices; an overview of both the Diamond Model of Intrusion Analysis and the Active Cyber Defense Cycle; a proposed integration of the Diamond Model of Intrusion Analysis and the Active Cyber Defense Cycle; and suggested additional work to further the study of intelligence in active cyber-defence activities.

Intelligence-Analysis Basics

Although practiced for millennia, intelligence analysis emerged as an academic discipline only relatively recently. The following paragraphs provide a basic examination of intelligence as understood within the U.S. Intelligence Community (IC), and a brief history of how the practice of intelligence analysis has evolved in recent years. An overview is provided of the Intelligence Community's five-step intelligence cycle, the Director of National Intelligence's eight-step analysis process, and the Department of Defense's understanding of intelligence as the product of transforming data into information that is then analysed within the context of a commander's operational goals.

The IC employs a standard intelligence cycle, as shown in **Figure 1**, below. The five elements are Planning and Direction, Collection, Processing, Analysis and Production, and Dissemination (Central Intelligence Agency n.d.). Planning and Direction identifies the information an individual or organisation needs to accomplish its goals. Collection is gathering the data required to meet the information need identified during the planning and direction stage. Processing is transforming collected data into a format useable for analysis. Analysis and Production is the stage during which the collected and processed data is placed in context and given meaning to answer information needs identified during the Planning and Direction phase. Dissemination is the process of delivering a useable intelligence product to the individual or organization needing the information.

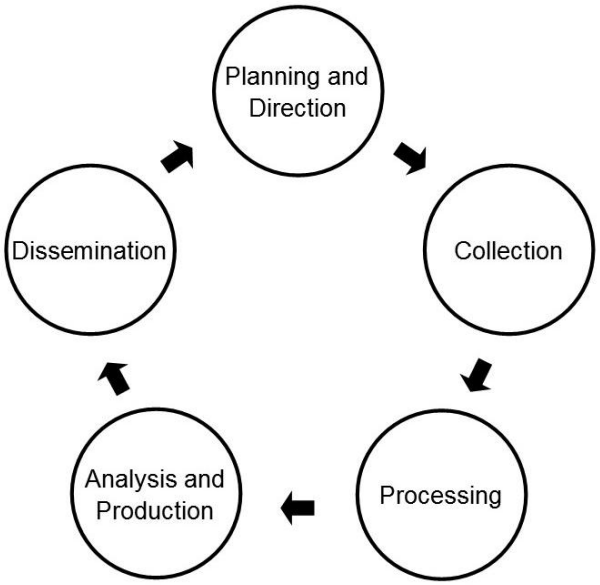


Figure 1: Intelligence cycle (Central Intelligence Agency n.d.)

The Office of the Director of National Intelligence refined the final two steps of the intelligence cycle into an eight-step process, as shown in **Figure 2**, below: Establish Context, Define the

Problem, Generate/Refine Hypotheses, Acquire Information, Structure Information, Test Hypotheses, Make Assessment, and Communicate Assessments (Director of National Intelligence n.d.). Establishing context is defining the time and space in which intelligence is needed. Defining the problem is identifying a consumer's unmet information need. The next four steps form a cycle of hypothesis generation, data/information collection, data/information structuring, and hypotheses testing, that is repeated until an analyst is able to build a coherent story supportable by validated facts and information. This coherent story is then presented in an appropriate format that effectively communicates the story to the intended customer.

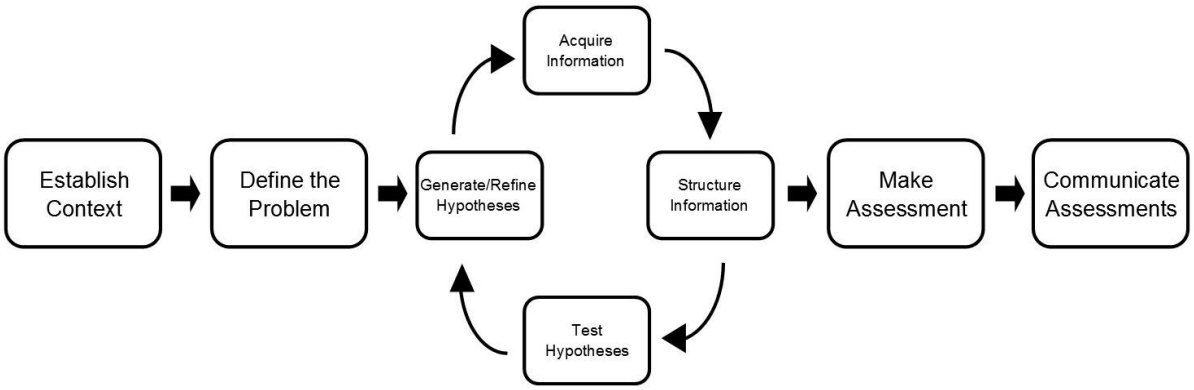


Figure 2: Eight-step problem solving process (Office of Director of National Intelligence n.d.)

The U.S. Department of Defense *Joint Publication 2-0 Joint Intelligence* (2013b) discusses intelligence as the result of processes, such as those described above, that collect data in a defined operational environment, and process the data to transform it into information that can then be analysed to meet a commander's operational needs (see **Figure 3** below).

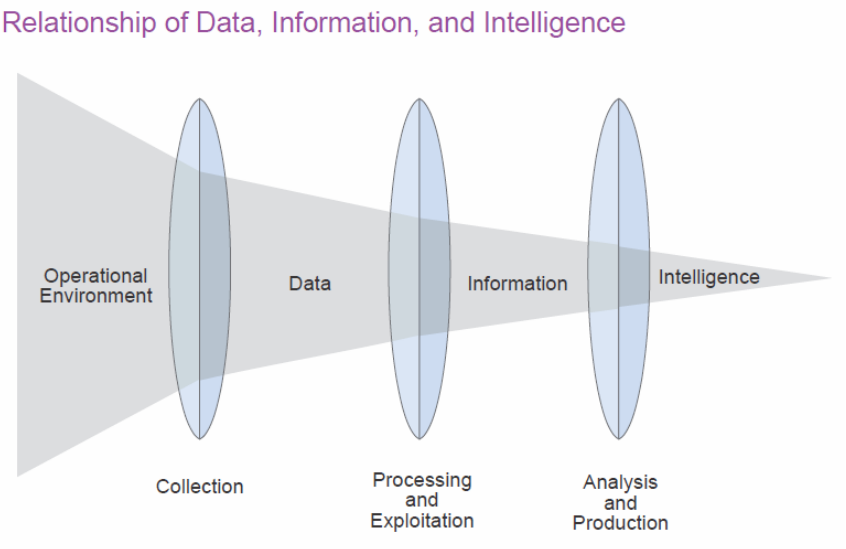


Figure 3: Transforming data into intelligence (U.S. Department of Defense 2013b)

Brief History of Modern Intelligence Analysis

The modern U.S. Intelligence Community came into existence with the passage of the National Security Act of 1947 (Federation of American Scientists n.d.). For much of the post-war period, intelligence analysis was conducted in a largely ad hoc manner that relied on individual abilities and the willingness of experienced analysts to train less experienced analysts in analytic tradecraft (Marrin 2005). The amount of emphasis placed on formalised analytic tradecraft fluctuated over time (Marchio 2013); however, this began to change in the late 1990s when intelligence analysis began to emerge as a formal academic area of study. In 1999, Richard Heuer published the book *Psychology of intelligence analysis*, a foundational work on intelligence analysis that examined how the human mind unavoidably introduces cognitive biases into its thought processes. To counter these biases, Heuer (1999) introduced the now well-known structured analytic technique known as Analysis of Competing Hypotheses, in which an analyst evaluates a set of written hypotheses according to a common set of criteria to determine which hypothesis is least likely to be incorrect.

After Heuer (1999) published his book, the intelligence community came under heavy criticism for failing to anticipate the September 11 terrorist attacks on the World Trade Center and for providing poor analytic support to policy makers prior to the invasion of Iraq in 2003 (Robb *et al.* 2005). The criticism led the intelligence community to move aggressively toward formalised and structured analysis processes, the result of which can be seen in documents such as the Central Intelligence Agency's 'Tradecraft primer: structured analytic techniques for improving intelligence analysis' (2009), which identified and explained twelve structured analytic techniques. The emphasis on formal, structured analytic techniques continues today and is expected to remain a point of emphasis for the foreseeable future. James Marchio (2013), a senior evaluator in the Analytic Integrity and Standards Group within the Office of the Director of National Intelligence, asserted that the need for strong analytic tradecraft, such as the techniques described in the Central Intelligence Agency's Tradecraft Primer, will endure.

These very brief examinations of intelligence analysis and its recent history are relevant to this discussion for multiple reasons. First, they illustrate that intelligence is the result of clearly defined processes that transform data into information and ultimately intelligence through the application of human intellect. Second, the evolution of analysis within the intelligence community can serve as an example for the cyber-intelligence community. Just as the intelligence community adopted a more structured and formalised approach to intelligence analysis after high-profile failures, so the cyber-security community has the opportunity today to deepen its understanding and use of intelligence in the wake of myriad recent high-profile data breaches.

Cyber-Threat Intelligence Today

With the emergence of cyber-threat intelligence as a point of emphasis, the cyber-security community has seen significant growth and innovation in the number and type of cyber threat intelligence products on the market (Caltagirone 2015). However, the growth has not been uniformly positive. The following paragraphs include brief summaries of studies and surveys that provide a useful picture of the state of cyber-threat intelligence today. These summaries illustrate both the significant effort dedicated to cyber threat intelligence and the areas in which cyber-threat intelligence can be improved.

In a SANS Institute whitepaper, Greg Farnham (2013) discussed the cyber-security community's approach to cyber-threat intelligence, which he defined as "threat intelligence related to computers, networks, and information technology" (p. 8). Farnham's discussion of intelligence identified two key elements deemed applicable to the cyber realm. First, Farnham stated that intelligence is not simply data, but information that has been analysed. Second, he asserted that "intelligence must be actionable" (p. 8). Additionally, Farnham noted that indicators of compromise, such as IP addresses, domain names, and file hashes, are often the focus of cyber-threat intelligence because they are easily actionable.

Building on the desired actionable nature of cyber-threat intelligence, Farnham provided an overview of the many standards and tools available for exchanging cyber-threat intelligence information. Among the standards, formats, and protocols for sharing cyber-threat intelligence are the Traffic Light Protocol, Incident Object Description and Exchange Format, Real-Time Inter-Network Defense, Mandiant's Open Indicators of Compromise framework, Verizon's Vocabulary for Event Recording and Incident Sharing (VERIS), MITRE's Cyber Observable Expression (CybOX), Structured Threat Information Expression (STIX), and Trusted Automated eXchange of Indicator Information (TAXII) standards (pp. 10-20).

Farnham's discussion illustrates an important point of existing cyber-threat intelligence practices. Namely, cyber-threat intelligence tends heavily towards technical matters at the expense of substantive discussions of intelligence production and consumption. Farnham dedicated less than two pages to conceptually discussing intelligence and approximately ten pages to examining tools, standards, and protocols for sharing threat information.

This is not a new phenomenon. In a 2011 report, the Cyber Intelligence Task Force of the Intelligence and National Security Alliance described existing cyber security practices as expensive, inefficient, and incapable of effecting a significant change in the environment. The Cyber Intelligence Task Force asserted that intelligence is a vital element of any security effort, including cyber security, but recognised that cyber intelligence is an emerging discipline needed to "systematically define and establish effective cyber intelligence approaches" (Cyber Intelligence Task Force 2011, p. 3) and develop strategies that move beyond "patch and pray processes" (p. 17).

While acknowledging the significant emphasis and effort already dedicated to meeting cyber security needs, the 2011 Cyber Intelligence Task Force emphasized a number of important issues needing to be addressed. First, the cyber security community had placed little focus on "truly defining and exploring the cyber threat environment at a higher level, its unique dynamics, and the potential impact on our economy and national security" (2011, p. 4). Second, cyber-intelligence-analysis expertise was disproportionately concentrated in the Intelligence Community. Pockets of deep technical and analytic cyber expertise exist within the intelligence community, but more than 90 percent of threat data resides in the unclassified realm. Third, incorporating analytic methodologies developed in the intelligence community through decades of accumulated experience is necessary for cyber intelligence to mature into a discipline capable of mitigating threats at tactical, operational, and strategic levels.

More recently, the SANS Institute (Shackleford 2014) surveyed 350 information technology professionals from a broad range of industries and work roles to examine the state of security analytics and intelligence capabilities within the cyber-security community. Survey results showed a decrease in the use of intelligence tools and services from 38 percent of respondents in 2013 to 29 percent of respondents in 2014. However, survey respondents who did use intelligence tools and services reported improved visibility into system and network activity. Based on survey results, SANS concluded that deploying automated tools for processing threat data improved visibility into computer networks, reduced the time needed to detect security incidents, and enabled more rapid incident response actions. However, SANS also concluded that cyber-security personnel need to expand their abilities beyond the 'buttonology' of simply running data collection and correlation tools to an understanding of how to analyse data and provide richer cyber-threat intelligence.

In 2015, the SANS Institute conducted a second related survey of 326 information technology professionals to examine who was using cyber-threat intelligence and how they were using it (Shackleford). For the survey, cyber-threat intelligence was defined as "the set of data collected, assessed, and applied regarding security threats, threat actors, vulnerabilities and compromise indicators" (Shackleford 2015, p. 1). Survey results indicated that 75 percent of respondents found cyber-threat intelligence important to security, and 69 percent of respondents reported having implemented cyber threat intelligence to some degree (Shackleford 2015). Among organisations that had at least partially integrated cyber-threat intelligence into their operations, the most common elements implemented were the use of raw, unfiltered feeds of cyber-threat intelligence data and visualization tools, as well as the integration of a variety of aggregated data. Respondents whose organisations had implemented such measures reported improved ability to rapidly and accurately detect and respond to attacks and to see attacks in context (Shackleford 2015). Despite the relatively high percentage of respondents using cyber-threat intelligence, only 27 percent reported using it extensively (Shackleford 2015). The survey results indicated that many organisations have significant room for improvement in how they gather and use cyber-threat intelligence.

Two points are important to note regarding these surveys. First, they defined cyber-threat intelligence as a set of data, not as a product of the data being processed and analysed. Second, the surveys considered cyber-threat intelligence almost exclusively through a technical lens focused on tactical cyber-defence activities. Fifty-nine percent of respondents indicated they were gathering intelligence from their internal systems, and 76 percent from the security community at large (Shackleford 2015). Among those receiving cyber threat intelligence from vendors, respondents identified the following sources of cyber threat intelligence: endpoint security vendors; unified threat management, firewall, and intrusion detection system vendors; cyber threat intelligence platform providers; vulnerability management vendors; SIEM vendors; application security vendors; log management vendors; whitelisting vendors; forensics vendors; and others (Shackleford 2015). This definition of cyber-threat intelligence and the technical lens focused on tactical cyber-defence activities both reflect a narrow view and limited understanding of intelligence analysis.

Poirier and Lotspeich (2013) surveyed the development of cyberspace from a U.S. Air Force perspective. Recognizing that computer networks generate data at a rate that exceeds the ability

of humans to process, Poirier and Lotspeich (2013) called for expanding the presence of network sensors and improving automated data correlation capabilities in order to reduce reliance on human analysts. Analyst involvement in defensive cyber operations was discussed in terms of "intuition and experience" (Poirier & Lotspeich 2013, p. 88). The need for conscious reasoning, a foundational process of intelligence production, was left unaddressed.

Folker and Bressette (2012) argue that the use of automation within intelligence analysis to reduce manpower requirements and compensate for human limitations is not without its downsides. Massachusetts Institute of Technology research revealed that, while such an approach can reduce some common biases in human thought processes, it may introduce an "automation bias" in which humans tend to uncritically accept answers generated from automated systems (Folker & Bressette 2012, p. 132). A second downside is the "lack of analytical agility" (Folker & Bressette 2012, p. 132) inherent in coded algorithms. Development and modification of algorithms is an expensive and labour-intensive process that presents a significant challenge when confronted with a dynamic opponent. While recognising the speed and data-volume advantages of analytics, Folker and Bressette (2012) argue that a well-trained cadre of intelligence analysts will be key to optimising the capabilities analytics present. Combining the computing power of analytics with the analytical agility of the human mind will maximize intelligence capabilities in a way neither could achieve alone.

Mattern *et al.* (2014) agree with both the results of the SANS surveys on the use of cyber-threat intelligence and Poirier and Lotspeich's (2013) understanding of cyber intelligence. Current cyber-security activities are largely reactive in nature and focus on identifying and eliminating on-network intrusion activity in order to minimize disruption of network operations. Cyber intelligence relies largely on visible, on-network data generated after an adversary is "already inside the wire" (Mattern *et al.* 2014, p. 705). However, Mattern *et al*. (2014) argue that such a reactive approach is insufficient. Cyber-security must transform into a proactive effort driven by cyber intelligence derived from both on-network data and off-network information regarding the capabilities, intentions, and activities of adversarial cyber actors. Cyber intelligence must address not only technical issues of network operations, but also issues related to the motivations and capabilities of adversaries. Use of these expanded sources of data and information will move cyber-security professionals beyond tactical cyber-defence activities and into the realm of operational and strategic decision making.

Townsend *et al.* (2013) examined the cyber-intelligence practices of six government agencies and twenty private sector organisations from academia and industry and found that the organizations used a wide variety of approaches to cyber intelligence. The current state of cyber intelligence was described as an effort to understand internal and external environments, gather data, and analyse technical threats (Townsend *et al.* 2013). However, cyber intelligence in many organizations lacked strategic analysis and failed to adequately inform executive level decision-makers, which the study authors attributed to a lack of standardised education and training requirements for analysts engaged in cyber intelligence work. Townsend *et al.* (2013) found that the cyber-intelligence workforces of the organisations studied consisted of a mixture of technically-oriented experts and traditional intelligence analysts. However, neither group fully understood the capabilities and limitations of the other group. Townsend *et al.* (2013) also found that, despite a claimed preference among respondents for training traditional intelligence analysts

in the technical aspects of cyber, when pressed to describe the qualities of an ideal candidate for cyber-intelligence work, expertise in the cyber realm was most emphasized. This bias towards cyber experience and expertise among those currently involved in cyber intelligence, most of whom have very deep roots in the technical aspects of cyberspace, suggests that the potential of intelligence analysis thought processes and methodologies is not yet sufficiently understood.

A 2013 study conducted by the Cyber Intelligence Task Force of the Intelligence and National Security Alliance explored cyber intelligence "as a disciplined methodology with understandable frames of reference in the form of operational levels" (p. 1). While conceding that the cyber-security community has not agreed upon a definition for cyber intelligence, the Cyber Intelligence Task Force asserted that a definition should not be limited only to considering data observed through network operations and activities. The cyber-security community's incomplete understanding of the entire series of actions required to execute malicious actions in cyberspace together with the natural tendency of very adept system administrators and network defenders to focus on technically-oriented activity directly observable on the network have led to a one-dimensional understanding of network defence and cyber intelligence. Overlooked is the fact that "all operations in cyberspace begin with a human being" (Cyber Intelligence Task Force 2013, p. 1). The 2013 study argued that synthesizing information from both human and technical elements will produce a more comprehensive understanding of the cyber domain and provide the potential to "get ahead" (Cyber Intelligence Task Force 2013, p. 3) by "integrating sound and time-tested intelligence thinking and methodology in the equation" (p. 11).

While the topic of cyber intelligence has entered the cyber-security arena, these studies demonstrate that it has not yet matured enough to begin altering the cyber-security landscape. The cyber-security community does not yet possess a sufficiently deep understanding of intelligence processes. However, this is not to say that the cyber-security community completely lacks an understanding of the value of formal, structured intelligence-analysis practices. Recent years have seen the emergence of the Diamond Model of Intrusion Analysis (Caltagirone, Pendergast, & Betz 2013), a proven, structured, analytic technique specifically designed for generating cyber-threat intelligence.

Diamond Model of Intrusion Analysis

Caltagirone, Pendergast, and Betz (2013) developed a structured analytic approach called 'The Diamond Model of Intrusion Analysis' (**Figure 4**, below) based on four core features: the adversary, capability, infrastructure, and victim. The four features are presented as a diamond, which represents the basic atomic-level element of intrusion activity. Each core feature of the diamond is edge-connected to the other elements, which allows analysts to add metadata to contextualise intrusion events. Six meta-features (timestamp, phase, result, direction, methodology, and resources) are offered; however, Caltagirone Pendergast, and Betz (2013) emphasized that meta-features can be added or deleted based on unique operational environment requirements.

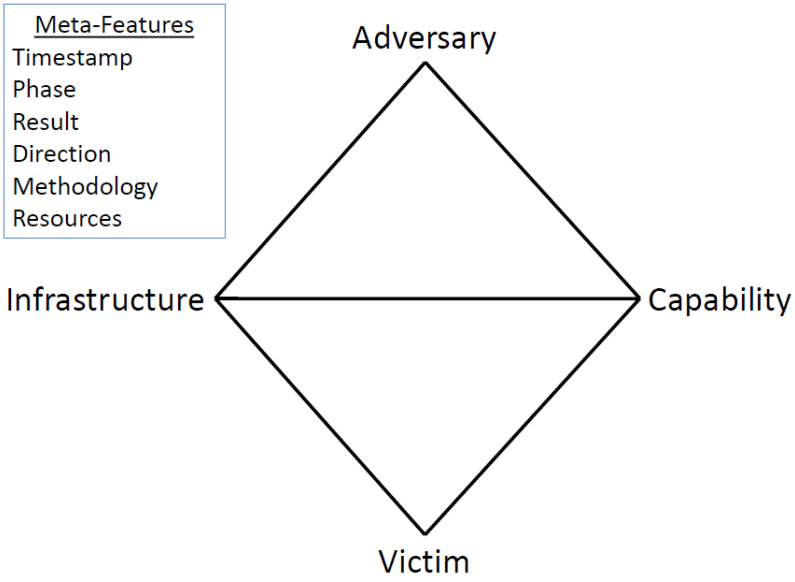


Figure 4: The core features and meta-features of the Diamond Model of Intrusion Analysis (Caltagirone, Pendergast, & Betz 2013, p. 9)

The Diamond Model's use of meta-features allows an analyst to identify a chain, or thread, of related intrusion events. Even without data on every phase of intrusion activity, meta-features allow analysts to identify commonalities between intrusion events and hypothesize the overall sequence of events. This technique can be further developed to cluster intrusion activity in activity groups by identifying sufficient commonalities in the meta-features of intrusion events and threads, which enables the Diamond Model to scale to meet tactical, operational, or strategic needs of an organisation. Caltagirone, Pendergast, & Betz (2013) also extended the Diamond Model to incorporate socio-political and technology factors, such adversary-victim relationships, victimology, and a common operational threat space.

The Diamond Model of Intrusion Analysis is relevant to this discussion for two important reasons. First, it illustrates the value having an analytic framework for consuming cyber-threat intelligence information and understanding data gathered through network security monitoring. Second, the Diamond Model provides a proven, but flexible analytic approach that can be customized to each operating environment. This will become particularly important for integrating cyber-intelligence analysis into active cyber-defence activities.

Active Cyber Defense Cycle

The Active Cyber Defense Cycle (Lee 2015a, 2015b) is a unified defensive cyber operations strategy implemented on one's own network (**Figure 5**, below). It was designed to enable the cyber-security community to move beyond current practices in which technically competent and talented individuals focus on isolated facets of cyber security without a sufficient understanding of the broader context. The Active Cyber Defense Cycle conceptualizes an ongoing cycle of active cyber-defence activity that consists of four phases: asset identification and network

security monitoring; incident response; threat and environment manipulation; and threat intelligence consumption.

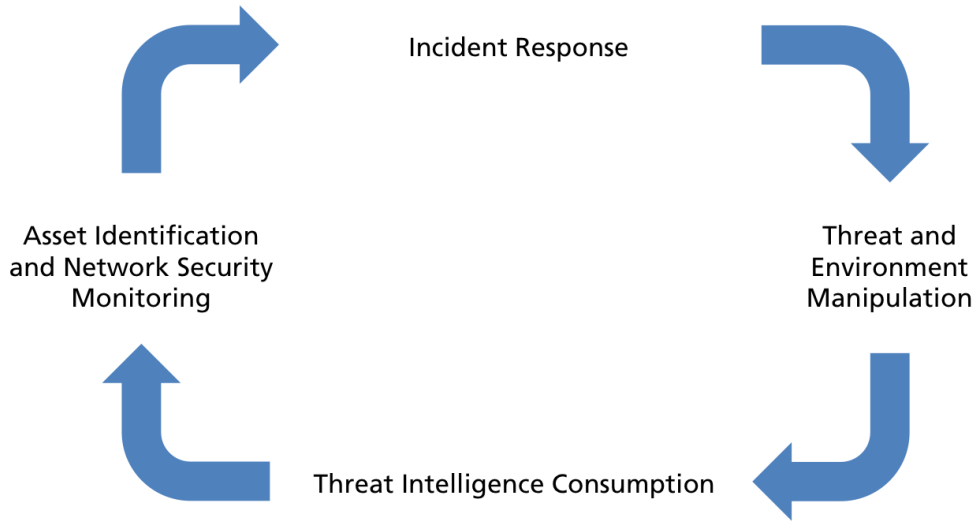


Figure 5: Active Cyber Defense Cycle (Lee 2015b)

The Active Cyber Defense Cycle is a continuous cycle with no end state. Although presented as a cycle of sequential events, in practice the Active Cyber Defense Cycle phases represent continuous processes that occur simultaneously and which are interrelated. Asset identification and network security monitoring seeks to maintain a strong situational awareness through an intimate knowledge of an organisation's environment, including an accurate accounting of network devices, an in-depth understanding of network architecture, and effective monitoring of network activity. Threat intelligence consumption is the identification and use of threat information tailored to an organisation's operating environment, network assets, and architecture. Incident response is action taken to mitigate an identified threat to an organisation's network. Threat and environment manipulation reveal how an organisation chooses to interact with a threat to derive additional intelligence information or alter the environment to mitigate the threat. This may include actions such as static or dynamic malware analysis and physical or logical changes to network architecture.

The Active Cyber Defense Cycle is relevant to this discussion for several reasons. First, the Active Cyber Defense Cycle provides a comprehensive strategy that unifies often fragmented cyber security practices. It accommodates and integrates existing cyber-security capabilities and approaches, such as network security monitoring technologies, incident response practices, and malware analysis, and also allows sufficient flexibility to integrate emerging defensive technologies. Second, the Active Cyber Defense Cycle explicitly rejects activity on adversarial networks. This has two important implications. First, the Active Cyber Defense Cycle provides a workable cyber-defence strategy unencumbered by legal and national security complications of hack back strategies. Second, restricting active cyber-defence activity to one's own network clearly defines the operational environment, which directly affects the intelligence required to support active cyber-defence activities.

The Active Cyber Defense Cycle considers cyber-threat intelligence based on established intelligence-analysis thought as described in *Joint Publication 2-0 Joint Intelligence* (U.S. DoD 2013b). As understood by the U.S. Department of Defense, intelligence is the result of a process in which data that has been collected, processed into information, and then analysed within a specific operational context in order to give it value. While valuing raw feeds of indicators of compromise and other technical details, the Active Cyber Defense Cycle does not consider such data true threat intelligence. The raw technical data provided in these feeds may contribute to cyber-threat intelligence, but only when evaluated within the context of the operational environment of a given organisation. Absent context, raw technical details remain nothing more than isolated data points lacking inherent value.

Within the construct of the Active Cyber Defense Cycle, the underlying context for cyber-threat intelligence is its applicability to active defensive actions implemented on one's own network. Cyber-threat intelligence may come from incident response data, malware analysis details, correlated sets of data on known intrusion campaigns, or indicator of compromise information shared via any number of data sharing protocols and mechanisms.

Integrating the Diamond Model and the Active Cyber Defense Cycle

An examination of the Diamond Model of Intrusion Analysis (Caltagirone, Pendergast & Betz 2013) and the Active Cyber Defense Cycle (Lee 2015a; 2015b) reveals the potential of integrating a structured cyber-intelligence analysis process into a unified defensive-cyber-operations strategy. The Diamond Model provides a simple analytic model based on just four core features, yet allows significant flexibility in using meta-features to capture important contextual information. The Active Cyber Defense Cycle provides a comprehensive strategy that incorporates existing cyber-security capabilities and practices, such as network security monitoring, incident response, malware analysis, and cyber-threat intelligence. Additionally, the Active Cyber Defense Cycle confines itself to activity on the defended network, thereby avoiding complicated and unresolved legal and national security issues associated with unauthorized activity on an adversary's network.

Integrating the Diamond Model of Intrusion Analysis and the Active Cyber Defense Cycle requires one to synthesize two distinct viewpoints on cyber security, one derived from an intelligence point of view focused on understanding the actions of an adversary, and the other derived from an operational point of view based on the goals of network defenders, as shown in **Figure 6**, below. As two core features of the Diamond Model are adversary and victim, the Diamond Model provides a useful framework for visualizing the relationship between an intelligence point of view focused on the actions of an adversary and an operational point of view focused on the active cyber-defence cycle.

On the adversary's side, or top half of the Diamond Model, the focus is on the first phase of the Active Cyber Defense Cycle, consuming threat-intelligence information. Threat-intelligence information is consumed to determine whether an adversary uses infrastructure s/he owns, or infrastructure he has exploited and gained control over. Similarly, threat intelligence information can be consumed to determine whether an adversary statically uses the same infrastructure for extended periods of time, or dynamically changes the infrastructure used to avoid re-use of easy-to-identify indicators. On the capability side, an adversary's tactics, techniques, tools, and

procedures can be characterised to indicate how an adversary uses infrastructure and what s/he has or is seeking to accomplish on the defended network. Socio-political information regarding an adversary can also be incorporated to provide detail on an adversary's driving motivation—whether it be hacktivism, cybercrime, or state-sponsored espionage.

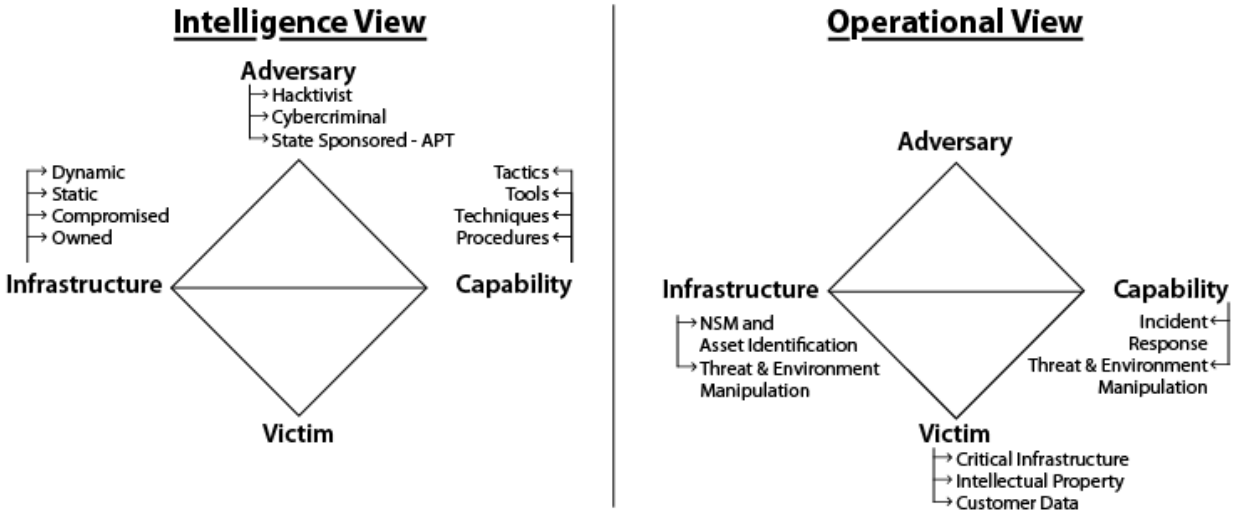


Figure 6: Comparison of intelligence and operational views and associated meta-features within the framework of 'The Diamond Model of Intrusion Analysis'

On the victim or network defender's side, the bottom half of the Diamond Model, the focus is on characterizing the remaining three phases of the Active Cyber Defense Cycle: asset identification and network security monitoring, incident response, and threat and environment manipulation. On the infrastructure side, asset identification and network security monitoring focus on ensuring all network devices are accounted for and monitored for indicators of adversary activity. Threat and environment manipulation includes but is not limited to actions such as changing firewall settings, updating rules within intrusion detection/prevention systems, or modifying network architecture. On the capability side, incident response concerns containment, mitigation, and remediation activities after an adversary has gained access to the defended network. Threat and environment manipulation involves things such as static and dynamic malware analysis. Socio-political information regarding a victim can also be included to identify key interests needing protection, such as critical infrastructure, intellectual property, or customer data.

These adversary-oriented intelligence and network defence-oriented operational views meet at the axis between infrastructure and capability, forming a cyber frontline of sorts. It is along this line of skirmish that adversary and defender engage each other to achieve their respective goals. Information collected along this line of skirmish through sensing and monitoring actions can feed data back into the intelligence and analysis cycles to produce refined and updated intelligence to inform network defence decisions. Synthesizing the two perspectives into a unified framework based on the Diamond Model provides a useful model for maintaining

situational awareness of both the activities of an adversary and the active cyber-defence activities and capabilities of network defenders, as **Figure 7**, below, illustrates.

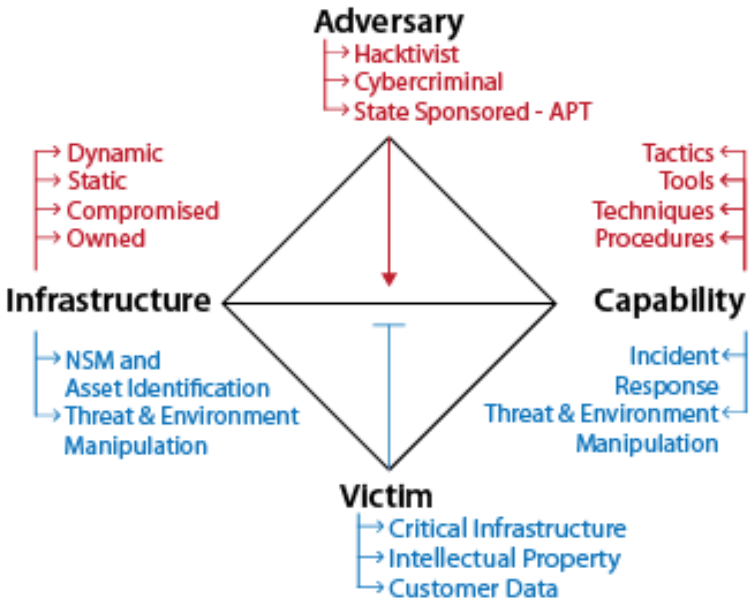


Figure 7: Synthesis of intelligence and operational views and associated meta-features within the framework of 'The Diamond Model of Intrusion Analysis'

The flexibility of both the Diamond Model of Intrusion Analysis and the Active Cyber Defense Cycle means they are easily adaptable to changes in the cyber environment. The Diamond Model can adapt to the emergence of new command and control infrastructures, new malware or changes in the tactics, techniques, and procedures employed by attackers; and the Active Cyber Defense Cycle easily incorporates new network defence tools and capabilities. This flexibility and adaptability is retained after the two models are integrated.

Integrating the Diamond Model and the Active Cyber Defense Cycle can also help network defenders better understand the relationship of their actions to core business goals and processes. The network defence-oriented operational view, particularly the Victim feature at the very bottom of the Diamond Model, makes explicit the core business asset or process to be defended. Visualizing the relationship between network defence actions and core business interests can facilitate active cyber-defence operations in two ways. First, it can help network defenders prioritize tactical defensive actions based on the relationship of a given threat to the asset most valued by the business. Second, it can assist network defenders in communicating with business executives at an operational or strategic level by making clear the relationship between a cyber threat and core business interests.

Future Work

This paper considered the integration of the Diamond Model of Intrusion Analysis (Caltagirone, Pendergast & Betz 2013) and Active Cyber Defense Cycle (Lee 2015a; 2015b) from a conceptual perspective only. The synthesis of operational and intelligence viewpoints offered in **Figure 7**, above, provides a mental framework designed to assist network defenders in visualizing the relationship between an adversary-focused consumption of threat intelligence and an operationally-focused active cyber-defence cycle. However, this framework was not tested within the context of a case study using data from a real-world intrusion. Significant benefit may be gained from an additional study testing the utility of the proposed integration through a cyber-defence exercise to validate or invalidate the proposed integration. Additionally, a case study approach may identify additional or more appropriate meta-features than those provided in **Figure 7**, above, and may help develop a course-of-actions matrix with a detailed menu of options a network defender may consider for responding to an adversary's malicious actions against a network.

Conclusion

The cyber-security community faces an environment of continual change. With each passing day, the scope and complexity of cyberspace increases. Networks to be defended are continually changed and updated to meet emerging business and operational requirements. The last decade has seen a rapid and massive expansion of mobile devices that has drastically increased the scope, complexity, and difficulty of cyber-security's task. The rate of change in cyberspace is expected to increase exponentially as the Internet of Things emerges. At the same time, malicious actors continue to find creative ways to exploit these same developments for illicit gain. In response to this malicious activity, the cyber-security community continually develops and refines security products. This has produced a situation in which problems resulting from technological developments are understood largely in technological terms and assumed to necessitate technological solutions. The history of cyber security since at least the mid-1980s has shown the limits of such an approach. The adoption of well-thought-out cyber-security strategies driven by a mature cyber-intelligence capability is the next logical step needed to better leverage the power of cyber-security technologies and to potentially change the balance of power between attacker and defender. The integration of structured cyber-intelligence analysis and active cyber-defence operations described in this paper is offered as a step towards this solution.

References

Caltagirone, S 2015, 'The cost of bad threat intelligence', viewed 6 February 2017, <http://www.activeresponse.org/the-cost-of-bad-threat-intelligence/>.

——, Pendergast, A & Betz, C 2013, 'The Diamond Model of Intrusion Analysis', viewed 6 February 2017, <http://www.activeresponse.org/wp-content/uploads/2013/07/diamond.pdf>.

Central Intelligence Agency 2009, 'Tradecraft primer: structured analytic techniques for improving intelligence analysis', viewed 6 February 2017, <https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/books-and-monographs/Tradecraft%20Primer-apr09.pdf>.

——n.d., 'Kid's zone', viewed 6 February 2017, <https://www.cia.gov/kids-page/6-12th-grade/who-we-are-what-we-do/the-intelligence-cycle.html>.

Cyber Intelligence Task Force 2011, 'Cyber intelligence: setting the landscape for an emerging discipline', Intelligence and National Security Alliance, viewed 5 April 2015, <http://www.oss-institute.org/storage/documents/Resources/studies/insa_cyber_ intelligence_2011.pdf>.

Cyber Intelligence Task Force 2013, 'Operational levels of cyber intelligence', Intelligence and National Security Alliance, viewed 15 May 2015, <http://csrc.nist.gov/cyberframework/framework_comments/20131213_charles_alsup_insa_ part3.pdf>.

Denning, DE & Neumann, PG 1985, 'Requirements and model for IDES—a real time intrusion detection expert system', viewed 18 May 2015, <http://faculty.nps.edu/dedennin/publications/ IDESReport SRI 1985.pdf>.

Director of National Intelligence n.d., Analysis 101: Participant Guide.

Eichin, MW & Rochlis, JA 1989, 'With microscope and tweezers: an analysis of the Internet virus of November 1988', viewed 18 May 2015, <http://www.utdallas.edu/~edsha/UGsecurity/ internet-worm-MIT.pdf>.

Farnham, G 2013, Tools and standards for cyber threat intelligence projects, SANS Institute, viewed 26 May 2015, <https://www.sans.org/reading-room/whitepapers/warfare/tools-standards-cyber-threat-intelligence-projects-34375>.

Federation of American Scientists n.d., 'The evolution of the U.S. intelligence community—an historical overview', viewed 6 February 2017, <http://fas.org/irp/offdocs/int022.html>.

Folker, RD & Bressette, KB 2012, 'Realizing the potential of analytics arming the human mind' *Air & Space Power Journal*, viewed 6 February 2017, <http://www.airpower.maxwell.af.mil/ digital/pdf/articles/2012-Nov-Dec/V-Robert-Folker-Bressette.pdf>.

Heuer, RJ 1999, 'Psychology of intelligence analysis', viewed 6 February 2017, <https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/books-and-monographs/psychology-of-intelligence-analysis/PsychofIntelNew.pdf>.

Krebs, B 2003, 'A short history of computer viruses and attacks', *The Washington Post*, viewed 6 February 2017, <http://www.washingtonpost.com/wp-dyn/articles/A50636-2002Jun26_2.html>.

——2015, 'Data breach at health insurer Anthem could impact millions', viewed 6 February 2017, <https://krebsonsecurity.com/2015/02/data-breach-at-health-insurer-anthem-could-impact-millions/>.

Lee, RM 2015a, 'Active cyber defense cycle', viewed 6 February 2017, <http://www.irongeek.com/i.php?page=videos/bsideshuntsville2015/active-cyber-defense-cycle-robert-m-lee>.

——2015b, *Threat intelligence in active cyber defense* (Part 1), viewed 1 March 2016, <https://www.recordedfuture.com/active-cyber-defense-part-1/>.

Lynn III, WJ 2010, 'Defending a new domain, the Pentagon's cyberstrategy', *Foreign Affairs,* viewed 6 February 2017, <http://www.foreignaffairs.com/articles/66552/william-j-lynn-iii/defending-a-new-domain>.

Marchio, J 2013, 'Analytic tradecraft and the intelligence community: enduring value, intermittent emphasis', *Intelligence and National Security,* DOI: 10.1080/02684527.2012.746415.

Marrin, S 2005, 'Intelligence analysis: turning a craft into a profession', viewed 8 March 2014, <https://www.e-education.psu.edu/drupal6/files/sgam/IA_Turning_Craft_into_Profession_Marrin.pdf>.

Mattern, T, Felker, J, Borum, R & Bamford, G 2014, 'Operational levels of cyber intelligence', *International Journal of Intelligence and Counterintelligence,* vol. 27, no. 4, pp. 702-719, DOI: 10.1080/08850607.2014.924811.

McGraw, G 2013, '"Active Defense" is irresponsible', viewed 6 February 2017, <https://www.cigital.com/justice-league-blog/2013/02/14/active-defense-is-irresponsible/>.

McFarlin, J 2015 'Hacking back: active defenses redux?', *Securityweek.com,* viewed 6 February 2017, <http://www.securityweek.com/hacking-back-active-defenses-redux>.

Poirier, WJ & Lotspeich, J 2013, 'Air Force cyber warfare now and the future', *Air & Space Power Journal,* viewed 5 May 2015, http://www.airpower.maxwell.af.mil/digital/pdf/articles/2013-Sep-Oct/F-Poirier_Lotspeich.pdf>.

Robb, CS, Silberman, LH, Levin, RC, McCain, J, Rowen, HS, Slocombe, WB, Studeman, WO, Wald, PM, Vest, CM & Cutler, L 2005, *The Commission on the Intelligence Capabilities of the United States Regarding Weapons of Mass Destruction: report to the president of the United States,* Commission on Intelligence Capabilities Regarding WMD, Washington, D.C., U.S.A.

Shackleford, D 2014, 'Analytics and intelligence survey 2014', SANS Institute, viewed 15 May 2015, <https://www.sans.org/reading-room/whitepapers/analyst/analytics-intelligence-survey-2014-35507>.

——2015, 'Who's using cyberthreat intelligence and how?', viewed 6 February 2017, <https://www.sans.org/reading-room/whitepapers/analyst/analytics-intelligence-survey-2014-35507> .

Target n.d., 'Data breach FAQ', viewed 6 February 2017, <https://corporate.target.com/about/shopping-experience/payment-card-issue-faq>.

Timberg, C, Nakashima, E, & Douglas-Gabriel, D 2014, 'Cyberattacks trigger talk of "hacking back"', *The Washington Post*, viewed 6 February 2017, <http://www.washingtonpost.com/ business/technology/cyberattacks-trigger-talk-of-hacking-back/2014/10/09/6f0b7a24-4f02-11e4-8c24-487e92bc997b_story.html>.

Townsend, T, Ludwick, M, McAllister, J, Mellinger, AO & Sereno, KA 2013, 'SEI emerging technology center: cyber intelligence tradecraft project summary of key findings, viewed 18 May 2015, <https://www.sei.cmu.edu/library/assets/ whitepapers/citp-summary-key-findings.pdf>.

U.S. Department of Defense 2013a, *Joint Publication 3-12(R) Cyberspace Operations*, viewed 5 January 2017, <http://www.dtic.mil/doctrine/new_pubs/jp3_12R.pdf>.

——2013b, *Joint Publication 2-0 Joint Intelligence*, viewed 22 May 2015, <http://www.dtic.mil/doctrine/new_pubs/jp2_0.pdf>.

U.S. Department of Homeland Security, Federal Bureau of Investigation 2016, *GRIZZLY STEPPE—Russian Malicious Cyber Activity* (JAR-16-20296), viewed 6 February 2017, <https://www.us-cert.gov/sites/default/files/publications/JAR_16-20296A_ GRIZZLY%20STEPPE-2016-1229.pdf>.

DDoS Attack Simulation to Validate the Effectiveness of Common and Emerging Threats

RJ Gordon

School of Engineering and Information Technology
Australian Centre for Cyber Security
University of New South Wales
Canberra, Australia
E-mail: r.gordon@unswalumni.com

Abstract: *Distributed Denial of Service (DDoS) attacks are a persistent cyber threat and a growing concern in computer security. This paper seeks to analyse DDoS attacks and the technologies that have been developed in an attempt to combat their effectiveness. This paper includes results from a DDoS simulation using commercial hardware appliances to both demonstrate and measure the effectiveness of DDoS attacks on a targeted victim. The simulation validates the relevance of these hardware appliances in identifying and reducing network vulnerabilities. This paper also provides insight on the current impact of DDoS attacks globally and the threat that these attacks pose in the future.*

Keywords: *Cybersecurity, Denial of Service Attack, DoS, DDoS, Cyberattack, Simulation, Computer Network Attack, ICMP Ping Flood, UDP Flood, HTTP Flood, Low-Rate DoS Attack, BreakingPoint*

Introduction

The rapid growth of collaborative environments, particularly cloud computing, expands the threat landscape for Denial of Service (DoS) attacks and leads to a large number of application developments for such environments (Zargar, James & Tipper 2013). Additionally, enhanced network capacity and the ever-expanding number of more interconnected devices have created conditions that are more vulnerable to exploitation by cyberattacks.

Distributed Denial of Service (DDoS) attacks are an advanced and evolved form of DoS as seen through their persistent use and increasing level of success on victims in recent years. Perhaps the most concerning factor is their ability to replicate legitimate traffic from a very wide distribution of networks, creating a problem that is difficult to solve using traditional IT security practices. DDoS attacks remain effective, providing a greater amount of network and/or resource congestion. This has resulted in a more widespread impact on interconnected users.

DDoS attacks provide a powerful tool for criminals and state-sponsored organisations to achieve political or economic objectives at the expense of their victims (Zargar, James & Tipper 2013). As nations continue to develop and test their cyber capabilities, individuals, businesses, and national infrastructure are in an increasingly vulnerable position. The development of prevention

techniques and overarching cyber strategies is as important now as it has ever been in the growing age of cyber warfare. However, there is uncertainty regarding how effective these techniques are in combating advanced DDoS attacks, which continue to improve in both efficiency and complexity.

Background

A DoS attack is considered an intentional and malicious attempt to disrupt, degrade, or deny services or resources to legitimate users of a computer network (Arora, Kumar & Sachdeva 2011). These attacks are best known for their ability to cause a website to crash from excess visitor traffic as generated by attackers, though they can be applied to achieve much broader outcomes. This definition of DoS matches that of Zargar, James & Tipper (2011, p. 2046) who further state that these types of attacks "have been known to the network research community since the early 1980s".

The impact of DoS attacks is becoming more significant because users are increasingly dependent upon the availability of the Internet globally. As noted by Leiner *et al.* (2003), the Internet has become far more important to current society as it changes the methods of conducting communication, business, and everyday life. The growth of Internet-based industries is staggering and shows no signs of slowing, which further exposes organisations to various forms of cyber attack, including those seen today and through the emergence of new technologies. As highlighted by Gupta, Joshi & Misra (2010, p. 268), "Internet usage is growing at an exponential rate as organizations, governments and citizens continue to increase their reliance on this technology".

A DDoS attack can disrupt, degrade, or deny either server resources or alternatively the bandwidth of a victim's network. This bottleneck congestion is typically observed at the link between an ISP's access router and a customer's domain router. Congestion can also be seen at the link between the customer's domain router and the victim. While DDoS attacks may not directly or permanently damage data, they do deliberately compromise the availability of resources and can cost the victim a substantial amount of time and money (Arora, Kumar & Sachdeva 2011).

The original DoS attacks could be mitigated with the use of suitable network equipment and appropriate security practices, which led to the evolution of DDoS attacks which were first seen in 1999. A DDoS attack differs from a standard DoS attack in the following manner: it uses a large number of machines to launch a coordinated attack against one or more targets (Gupta, Joshi & Misra 2010). This type of attack has proven to be far more difficult to defend as it uses a sophisticated approach which mimics legitimate users. DDoS attacks aim to overwhelm a target with a very large volume of useless traffic and are considered a major threat to the stability of the Internet (Nwaocha & Inyiama 2011).

The Cost of DoS Attacks on Society

As outlined by Thompson (2012, p. 58), "it is the expanding global reach and the low cost of entry and access that makes cyberspace a truly globalised or transnational concern for law enforcement and national security agencies alike". DDoS attacks can have financially devastating consequences on victim businesses, and research indicates they could cost a

company more than $100,000 per minute of downtime (Ponemon Institute 2012). The potential impact on an organisation presents itself as a very suitable justification to invest more heavily in IT security equipment, practices, and procedures which may prevent DDoS attacks. The financial consequences of having a company network that relies on Internet revenue being exploited with a DDoS attack is expected to be in far greater proportion than the cost of upgrading IT security to defend against it. As highlighted by Aamir & Arif (2013, p. 57), "attackers are using more sophisticated and automated tools to launch larger magnitudes of attacks at rapid speed, for which the defence has to be fast as well".

Fordyce (2013) suggests that DDoS attacks are used not only by hackers for 'political dissent', but also by others for economic or political purposes. Criminal organisations may seek financial benefits while governments might suppress particular networks for political gain. As an example, Wall Street is the leading financial centre of the world. It relies very heavily on electronic trading: and, noting its sensitivities, even the smallest degradation of services will affect global trading, costing millions of dollars to companies, individuals, or governments. Research indicates that DDoS attacks on financial sectors have increased and attackers seem to be more focused on this area with evolving strategies (Aamir & Arif 2013). Not surprisingly, organisations within this financial centre have invested a substantial amount on IT security specifically to defend against DDoS attacks. Such attacks can also inflict harm upon third parties indirectly, with no association to a target.

DoS Simulation, Emulation, and Testing

There are multiple options allowing an organisation to test its network environment without being exposed to the consequences of a real DDoS attack. These include hardware solutions that support the conduct of simulations and facilitate research to compare and contrast the various techniques seen with DDoS attacks, providing valuable network traffic data that can be measured and analysed. For example, Tomar & Tyagi (2014) conduct modelling of DoS flooding attacks using a traffic generator to create a DoS simulation. They compare the results and level of effectiveness of both TCP and UDP protocol attacks.

Modelling from Tomar & Tyagi (2014) demonstrates a router becoming congested when excess traffic is sent, from a combination of legitimate and attacker sources. The congestion is measured at the router by the number of dropped packets that fail to reach their destination, and the adverse impact these events have on legitimate users' traffic. This model shows similarity to the research published by Arora, Kumar & Sachdeva (2011). DDoS theory is successfully demonstrated through practical simulations such as these and leads to a greater understanding of their impacts in order to develop suitable prevention techniques. Further, by using more advanced simulation or emulation, an existing network can be tested for vulnerabilities. Large organisations, including those within the Wall Street financial centre, routinely conduct DDoS testing on their own networks to measure the performance and level of effectiveness of their cyber security defence mechanisms. Through the conduct of testing and analysis these hardware tools provide a means to improve upon any identified vulnerabilities and reduce their exposure to DDoS attacks.

Mirkovic *et al.* (2009) consider how to test DoS defences accurately and effectively. Part of the test criteria is to quantify the delay associated with detection and identification of attack packets,

in addition to the memory and CPU costs, that result in reduced processing capacity and latency across network equipment and servers.

The evaluation process used by Mirkovic *et al.* (2009) compares various test methods within their research to determine those which are most suitable. It includes the following elements:

- A testing approach such as a model, simulation, emulation, or deployment in an operational network;
- Test scenarios containing a combination of legitimate user and attacker traffic; and,
- A success metric to confirm the defence mechanism is reducing or eliminating a DoS threat.

During the evaluation, it was identified that emulation provides the greatest fidelity and is the recommended DoS test type (Mirkovic *et al.*), along with using a realistic network topology that scales up appropriately or is large enough not to require it. Whilst "inappropriate tools lead to incorrect results" (Mirkovic *et al.*, p. 3), using well developed tests, DDoS attacks can be accurately replicated to better understand their limitations.

Experimental Research

DDoS simulation and analysis using BreakingPoint

An Ixia BreakingPoint appliance was selected for use in conducting several DDoS attacks within a closed network cyber range. This testing was undertaken to validate the effectiveness of several common DDoS flooding attack techniques and to measure and compare the results observed with equipment against other DDoS attack simulations. The Ixia appliance is a commercially available tool built for conducting cyberattacks and for penetration testing. It is widely used by organisations globally to conduct testing of their own internal networks with the intent of identifying vulnerabilities and measuring the effectiveness of software and hardware solutions used to mitigate them.

DDoS attack network scenario

The scenario simulates an external network (the Internet) sending traffic through to an internal company network. The network design represented by **Figure 1**, below, was configured for testing in a BreakingPoint simulation. The intent was to create a realistic network configuration commonly used within an organisation, to demonstrate the effectiveness of a DDoS attack against it.

The test scenario was configured using Ixia virtual routers and a Palo Alto firewall appliance in-between, to filter traffic on route to the internal company network. Virtual routers were used to maintain consistency; only a single hardware device is introduced for testing, so as to accurately measure DDoS attack performance against it. The virtual routers were configured using the Ixia appliance. The aim of the simulation is to demonstrate and measure the effectiveness of a DDoS attack at a critical choke-point in the network.

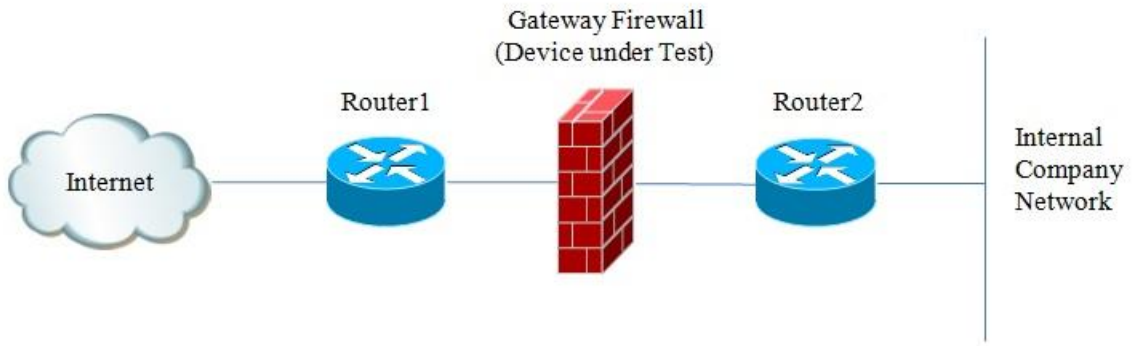


Figure 1: Network used to undertake a DDoS attack simulation

BreakingPoint software configuration

Figure 2, below, represents three large network ranges from Static Hosts 1/2/3 used to generate DDoS attacks from a total of 1000 IP addresses. This simulates a DDoS botnet attack leveraging 1000 legitimate, distributed users to generate a very large volume of useless traffic. This traffic is produced to significantly congest the company network, up to the point of achieving complete DoS. The device under test was a Palo Alto PA-2050 firewall appliance. The congestion and level of network performance degradation was measured by the packet loss experienced as traffic is sent out of Interface 1 (virtual Router 1), through the firewall, and into Interface 2 (virtual Router 2).

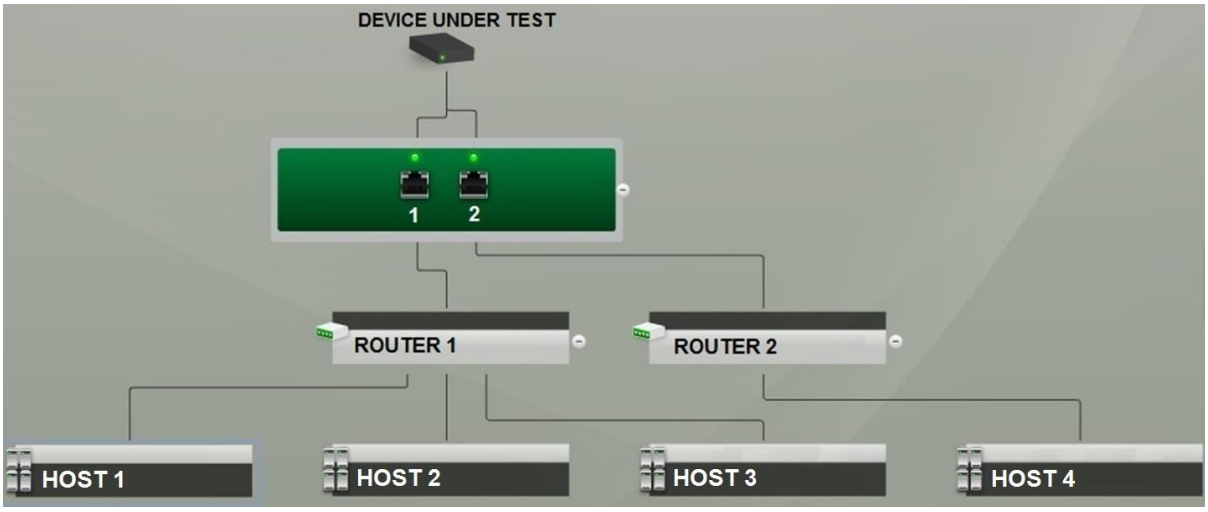


Figure 2: BreakingPoint network depicting attacker source IP address ranges (Hosts 1/2/3) that traffic through Router 1, into the device under test, and out Router 2, prior to reaching an internal company network (Hosts 4)

The relevant specifications of the PA-2050 firewall that have been tested to their limits include

- Maximum firewall throughout of 1.0 Gigabits/sec (Gbps);
- Maximum number of new sessions per second at 15,000;
- Maximum sessions at 250,000.

Exhausting the 1.0 Gigabits/sec throughput is expected to achieve network congestion, whilst exceeding the maximum number of sessions will exhaust the firewall's resources. In either instance, DoS should be demonstrated and measurable when very large numbers of packets fail to reach their destination.

All of the configurable variables for DDoS attack testing using BreakingPoint are located within the Session Sender. However, the specific variables used to scale the DDoS simulated attacks are the maximum flow creation rate and maximum concurrent flows. These values correlate to the above specifications used by the firewall appliance of maximum number of new sessions per second, and maximum concurrent flows. The Ixia appliance can deliver a maximum of 10 million concurrent flows, and a maximum flow creation rate of one million through its BreakingPoint software. This greatly exceeds the limits of the PA-2050 firewall and thus provides a suitable means to test its performance under stress.

Three DDoS flood attack types were tested: Internet Control Message Protocol (ICMP) ping flood attack (network layer), User Datagram Protocol (UDP) flood attack (transport layer), and a HTTP flood attack (application layer). The standard length of the test was two minutes to ensure sufficient data was captured, with an additional one second each for ramp up and ramp down at the beginning and end of the attack. The use of ramp up and ramp down ensures the test has some phase whereby bandwidth is exclusively dedicated to setting up sessions. Further testing was also conducted within intervals of 30 seconds, five minutes, and 10 minutes, although the results demonstrated no real difference in data.

Initial loopback test using an ICMP ping flood

A loopback test was initially conducted prior to the introduction of a device to be tested. The loopback was able to confirm the test environment was not subject to any traffic degradation, as measured through packet loss. An ICMP Ping Flood is the most suitable option for a loopback test because it also requests reply packets to be sent, and, thus, can accurately test in both directions of the network.

As illustrated by **Table 1**, below, the loopback test results demonstrate a 100% success rate of ICMP packets being sent and received. These results confirm no packet loss is experienced within the Ixia appliance in both directions, prior to a device being introduced into the network for testing. The loopback test verifies that the network produces results as intended without any anomalies.

Superflows		Frames / Sec				Data Rate (Mbps)				Packet Tx Success Rate
Creation Rate	Concurrent Flows	Tx (Int 1)	Rx (Int 2)	Tx (Int 2)	Rx (Int 1)	Tx (Int 1)	Rx (Int 2)	Tx (Int 2)	Rx (Int 1)	
1,000	100	630	630	630	630	4.98	4.98	4.98	4.98	100%
10,000	1,000	5,459	5,459	5,459	5,459	37.43	37.43	37.43	37.43	100%
100,000	10,000	100,000	100,000	100,000	100,000	257.3	257.3	257.3	257.3	100%
1,000,000	100,000	380,800	380,800	380,800	380,800	997	997	997	997	100%

Table 1: Sample of the ICMP ping flood loopback test results

ICMP ping flood attack simulation

This test scenario involved sending an overwhelming number of ICMP Ping packets from 1000 distributed users/attackers. The goals of the test were to measure the level of network congestion and seek to observe DDoS, whereby a very large number of packets fail to pass through the firewall and onto their destination.

An ICMP ping, also known as an ICMP echo request, is a standard transport layer protocol typically used as a diagnostic tool to test connectivity across a network (Stallings & Brown 2012). An ICMP ping generates a request from its source that seeks a response from its destination device. It is important to realise for the conduct of this test that reply packets should be sent from the Virtual Router 2, producing the exact same volume of return traffic back across the firewall. In the event of network congestion or complete DoS being observed, not only will packets fail to be received at their intended destination, but reply packets will also fail to reach Virtual Router 1 as well, in the other direction.

An ICMP ping flood is a widely used DDoS attack; and whilst it is becoming more common to block this protocol type within an organisation's external firewall, it remains a useful option for attackers and continues to be used widely on vulnerable networks (Gupta, Joshi & Misra 2010).

Table 2, below, represents a sample of the test data captured throughout the process. The superflow count within the sample is increased by a factor of 10 during each of the listed results below to demonstrate obvious changes in the data.

Superflows		Frames / Sec				Data Rate (Mbps)				Packet Tx Success Rate
Creation Rate	Concurrent Flows	Tx (Int 1)	Rx (Int 2)	Tx (Int 2)	Rx (Int 1)	Tx (Int 1)	Rx (Int 2)	Tx (Int 2)	Rx (Int 1)	
1,000	100	631	630	629	629	4.99	4.98	4.98	4.97	99.8%
10,000	1,000	5,459	4,893	4,893	4,885	37.43	34.26	34.26	34.2	89.6%
100,000	10,000	99,900	450	450	0	257.3	0.96	0.96	0	0.005%
1,000,000	100,000	380,800	251	224	0	997	0.58	0.58	0	0%

Table 2: Sample of the ICMP ping flood simulation test results

The above test results from an ICMP ping flood attack demonstrate that when the superflow rate is low, packet loss is negligible with only the occasional missing packet. However, as both the superflow rates are increased, packet loss is exponential. It is important to note that the bandwidth consumption is comparatively low, indicating that during the IMCP Ping Flood attack the maximum throughput is not the limiting factor of the device under test.

The ICMP Ping flood attack simulation demonstrates DoS being achieved against the firewall appliance through exhausting system resources as opposed to network resources. The sample test data within **Table 2**, above, indicates the firewall's threshold for superflow throughput is exceeded when the network data rate being transmitted is less than 100 Mbps. Of significance, this data rate is well below the 1.0 Gigabits/sec throughput capacity of the firewall. This simulation suggests that an appropriate selection of the type of DDoS attack provides flexibility for the attacker to reduce the number of required botnets to achieve DoS on a target network, as a single botnet is generally limited by available bandwidth.

During the simulation, the load on the firewall is progressively increased on a linear scale. However, once the firewall has reached saturation it fails completely. If the load on the firewall is subsequently reduced back to minimal levels, the firewall will continue to drop all packets. This behaviour demonstrates complete system failure, which has been achieved through DoS, and the firewall requires a manual reboot by the administrator before it begins to function correctly. Within this network scenario, once the firewall is considerably congested with an overwhelming amount of traffic, all packets being sent to it are dropped and the attacker can be considered successful in achieving DoS on the target network.

BreakingPoint is configured to operate within a large range of cyberattack test environments. Its test procedure, including many of the statistics obtained, function with use of the Transmission Control Protocol (TCP). Because ICMP is a non-TCP protocol, BreakingPoint produces very limited statistical information on ICMP ping flood attacks. Further, its built-in test criteria designed to set pass and fail levels are not configurable with an ICMP flood. Nevertheless, the results provided are effective in confirming DoS of the device under test.

UDP flood attack simulation

A standard UDP flood attack varies to an ICMP ping flood in that traffic is sent from a source without requesting a reply. UDP is a connectionless protocol and because of its relative simplicity, a UDP flood attack is a popular choice by attackers and can be generated using the available Low Orbit Ion Canon software (Asri & Pranggono 2015).

UDP packets can include a spoofed source IP address to protect the attacker from exposing the source network and location of the attacker, or series of attackers. Of note, a UDP flood attack using a botnet architecture and spoofed source IP addresses can protect the bot network, allowing the machines to remain anonymous and available for use in subsequent attacks.

During the process of a UDP flood attack, the traffic received at the destination is checked on the port that it is sent; and once it has been identified, no application is listening, the distant host will send a reply with an ICMP destination unreachable packet (Asri & Pranggono 2015). This process further congests the target device's resources as it processes the packets and distributes the reply.

Similar to the ICMP flood attack, the UDP flood attack simulation remains consistent by sending an overwhelming amount of UDP packets from 1000 source IP addresses to generate a very large volume of useless traffic, focused on the target network.

During the simulation, the load on the firewall is progressively increased on a linear scale. The sample results represented in **Table 3**, below, demonstrate DoS being achieved on the firewall appliance, causing it to drop all packets once it has been overloaded. At this point, the firewall requires a reboot before it will function correctly and begin undertaking packet inspection and filtering processes, which again demonstrates DoS has been achieved.

Superflows		**Frames / Sec**		**Data Rate (Mbps)**		**Packet Tx Success Rate**
Creation Rate	**Concurrent Flows**	**Tx (Int 1)**	**Rx (Int 2)**	**Tx (Int 1)**	**Rx (Int 2)**	
1,000	100	5,150	5,100	62.87	62.26	99%
10,000	1,000	48,860	48,390	596.5	590.8	99%
100,000	10,000	495,800	0	998	0	0%

Table 3: Sample of the UDP flood simulation test results

The UDP flood results differ from those of an ICMP flood since the bandwidth utilisation of the network is higher. The results suggest the firewall appliance can inspect UDP packets more easily with a lower processing burden on its resources. Given the simple nature of the UDP protocol, this matches its characteristics whereby processing overheads are minimal.

The results suggest a DDoS ICMP ping flood attack is more efficient than a DDoS UDP flood, whereby response traffic is generated, thus causing additional processing overheads. However, under the same conditions they undeniably produce much the same outcomes of DoS on the device under test.

HTTP flood attack simulation

A Hypertext Transfer Protocol (HTTP) flood attack is generally used to target web servers and applications; however, in order to reach a web server, in many instances a firewall exists in between them. For example The High Orbit Ion Canon, considered an upgraded version of the Low Orbit Ion Canon (LOIC) and also a popular choice by attackers, prevents firewall detection whilst undertaking a HTTP flood attack through the targeting of sub-pages (Asri & Pranggono 2015).

An HTTP flood attack typically sends HTTP GET and POST requests and is designed to consume the resources of a target server, as opposed to bandwidth. The advantage of a HTTP flood attack is its seeming legitimacy, making it difficult for security services such as a firewall to distinguish between legitimate or malicious HTTP traffic. It requires a connection to be established to the destination host, which then needs reply traffic to be sent, using processing power to do so. However, by design, an HTTP flood attack requires the attacking host machine to provide its source IP address to establish the connection. This compromises the source, which for a DDoS attack is likely to be a botnet running on a legitimate host (victim) machine.

This test scenario involved sending an overwhelming number of HTTP GET requests from 1000 distributed users/attackers. The BreakingPoint software produces more functional test results using an HTTP flood attack as Transmission Control Protocol (TCP) connection metrics are able to be collected, producing more thorough statistics within the simulation that could be used for further analysis. This includes the TCP connection rate, cumulative TCP connections, average TCP time between SYN and ACK packets, and a TCP state diagram that tracks packets as they are, for example, sent, received, held, and closed. However, it is worth noting that these metrics are less relevant within a DDoS attack but instead are of value in understanding and testing the other forms of cyberattacks BreakingPoint can generate.

The sample results represented in **Table 4**, below, are comparable to that of the other two attack types. During the simulation, the load on the firewall is again progressively increased on a linear scale, while the sample data within **Table 4** was selected to show more obvious results.

Superflows		**Frames / Sec**				**Data Rate (Mbps)**				**Packet Tx Success Rate**
Creation Rate	**Concurrent Flows**	**Tx (Int 1)**	**Rx (Int 2)**	**Tx (Int 2)**	**Rx (Int 1)**	**Tx (Int 1)**	**Rx (Int 2)**	**Tx (Int 2)**	**Rx (Int 1)**	
1,000	100	354	251	167	167	0.34	0.27	0.13	0.13	71%
10,000	1,000	5,406	2,972	1,981	1,981	4.94	3.13	1.51	1.51	55%
100,000	10,000	66,860	11,930	20,260	6,330	63.2	13.57	15.03	4.69	17.8%
1,000,000	100,000	499,700	0	0	0	372.6	0	0	0	0

Table 4: Sample of the HTTP flood simulation test results

During this simulation, DoS is achieved by exhausting the processing capacity of the firewall appliance, whilst only utilising a small portion of its maximum throughput. This DDoS attack type appears to be more efficient than a UDP flood as it requires less bandwidth overall to cause DoS on the firewall, potentially requiring a reduced number of botnet host machines to achieve the task.

Scenario Relevance to Real-world Networks

The intent of the simulation was to apply the theory of DoS attacks from a range of literature onto a practical scenario, whilst measuring the effectiveness of an attack. The BreakingPoint product successfully demonstrated the three varying types of DDoS attack used within a closed

network or sandbox environment. It is clear from the results produced that the hardware and software used within the simulation adequately provide a means to test the performance of an existing company network in defending such attacks, and to identify vulnerable points within its network.

The simplified network diagram represented in **Figure 1**, above, depicts a gateway firewall as a choke point to an internal company network. Within this scenario, once a DDoS attack has taken place, whilst the internal network is unaffected from communicating with local services behind Router 2, all Internet connectivity is lost. An alternative approach from an attacker is to target a specific server for a more focused attack. However, it is likely the traffic will still be required to traverse a company external firewall and possibly an internal firewall as well, prior to reaching its destination.

Of relevance to this typical network scenario, organisations increasingly rely on Internet access for daily operations. This may include offsite cloud server storage, server replication, a WAN link to a larger company headquarters or meshed with other branches, enterprise resource planning tools, databases, and much more. Thus, a DDoS attack on an organisation through its Internet connection can have a defining consequence, potentially reducing productivity, increasing costs, or reducing opportunities to earn revenue during the attack period and the subsequent outage that follows.

The three DDoS attack types were selected because they are not only common but also varied in their methods of achieving DoS. It is believed an ICMP ping flood attack in present day is likely to be blocked by standard firewall policies that prevent the protocol being used. The benefit of retaining access for this protocol is for flexibility with network testing amongst IT administrators. However, it could be considered an unnecessary risk, particularly for organisations that are likely to be targeted or are at risk of significant consequences as a result of a DDoS attack.

To demonstrate the ease at which an ICMP ping flood attack can be negated, the PA-2050 firewall was configured to block ICMP echo requests. Following the policy being applied, the DDoS attack was again produced, and the firewall successfully blocked all packets, including when the superflow rate was configured on BreakingPoint to its maximum.

A UDP flood is more complicated, as it can disguise packets used for legitimate purposes, including a variety of packet types, to avoid setting an obvious pattern. The risk of applying strict filtering processes for UDP packets is that it prevents legitimate traffic from being received into the network. Similarly, a HTTP flood attack is also difficult to filter, particularly if an organisation is hosting its own web server and requires various forms of HTTP requests to reach the server.

It is important to note that whilst this scenario depicts a standard company network with an Internet connection as the obvious point in which to conduct the attack, DoS attacks can be achieved on any network that has a point of connectivity for the attacker to reach. This includes private networks running over their own dedicated infrastructure, which may have a wireless point of entry, or access through a secondary network. As Bergin (2015, p. 386) notes, “a major

type of threat in wireless communications is denial of service (DoS)", which attempts to incapacitate network and computer resources. The susceptibility of emerging wireless technology used by the U.S. Department of Defense is not well understood, in particular that used with autonomous vehicles, including unmanned aerial vehicles and unmanned ground systems (Bergin 2015). Wireless networks have a finite amount of bandwidth available when compared to wired networks and can be more accessible than their counterparts, creating an easier target in some instances.

Use of Low-Rate DDoS Attacks

As identified in the results of the simulation above, the use of a low-rate DDoS attack is beneficial as it substantially reduces the likelihood of detection when an attack is taking place. Combining a sophisticated low-rate attack with the more specific targeting of a network device, such as a web server rather than the furthest network choke point, can have a different impact—in particular, one that is more discreet and lacks a pattern that can be detected by advanced security mechanisms.

This supports the research identified by Asri & Pranggono (2015) on the High Orbit Ion Cannon targeting web server system resources as opposed to bandwidth, whilst limiting its detection threshold by a firewall and using as few as 30-50 attacking machines. Additionally, Yang, Park & Chung (2013) categorises two DDoS attack types that are possible with only a small amount of traffic, including a Structured Query Language (SQL) search attack and a Mass Contents Request attack. The SQL search attack targets a database server hosted on a web server using an SQL query. It uses common search terms including the 'LIKE' keyword to produce mass search results that drain server resources. As the requests are seemingly legitimate, they have the potential to remain undetected. In contrast, a Mass Content Request attack requests the content files that have been uploaded to the website (Yang, Park & Chung 2013), with a focus on flash (.flv) and compressed (.zip) file extensions due to the time it takes to analyse these larger files.

These attacks employ a similar concept to that of a Domain Name Server (DNS) reflection attack, however, do not rely on DNS servers that are known to be closely monitored. The benefit of a low-rate attack could be a DDoS attack continuing for a sustained period, or the ability to recycle the technique on the same server in future without the exploit being mitigated by the victim.

Conclusion

DDoS attacks are an increasingly complex and persistent threat to society. Recent history has demonstrated the ability of these attacks to disrupt large-scale defended networks and services, which confirms their potential to cause much further damage in the future. Under state sponsorship, these attacks are also capable of achieving military objectives, perhaps even more successfully than through kinetic actions.

The evolution of DoS methods has seen an ever-increasing number of successful attacks globally, despite the greater uptake of cyber-security practices. DDoS techniques, particularly DDoS flood attacks, provide the ability for attackers to seamlessly access users anywhere in the world to generate network traffic on a mass scale, including from legitimate sources. The fact that these attacks impersonate regular network usage complicates defence mechanisms, which

continue to be reactionary in nature. Unless a DDoS attack produces a technique or signature that is known or detectable, through its use of distributed traffic, it becomes exceptionally challenging to filter.

The conduct of a DDoS simulation was beneficial in comparing the effectiveness of several common attack types. It was successful in corroborating the results from similar studies undertaken, whilst also demonstrating the value of utilising a commercial product to identify vulnerabilities on a typical network design. Importantly, the simulation determines the limits of a specific security device under test conditions to understand how it responds and the subsequent impact it has on that network. The results indicate where a typical choke point exists within a corporate network and emphasises the need for suitable cyber-security measures as close as possible to the Internet or other external network access point when defending against DoS attacks.

Application layer attacks appear to be some of the most successful DDoS attacks at present, due to the difficulty border routers, firewalls, and similar security perimeter devices have in detecting them. This is evidenced by the current popularity of the High Orbit Ion Cannon as an application layer attack tool, and the development from its predecessor, the Low Orbit Ion Canon (Asri & Pranggono 2015). Furthermore, a DDoS attack that targets the resources of a victim device as opposed to consuming available network bandwidth appears to be both more effective and less detectable. Low-rate DDoS attacks take this approach a step further by using more advanced and seemingly genuine methods of exhausting system resources, all whilst using minimal bandwidth.

Critical infrastructure and services are amongst the most sensitive networks that require the best protection. However, Supervisory Control and Data Acquisition (SCADA) networks have been shown to be amongst the most vulnerable as well, due in part to the enduring legacy design of many programmable logic controllers that remain widely in-use. If disrupted, degraded or denied altogether they could pose a very real threat to the lives of unsuspecting citizens. Additionally, businesses and other such organisations without the appropriate network security are vulnerable to DDoS attacks that can have widespread consequences. This underscores the need to revise these network architectures and to employ stronger defensive measures.

Analysis, thus, reveals vulnerabilities in both detection and prevention techniques against DDoS attacks. This highlights the need for a collaborative cyber-security mitigation strategy that incorporates individual users of networks, leadership within organisations, and network security experts. Additionally governments are required to provide support through legislation, research and development, and international engagement. It will be a challenge for ISPs to cooperate in order to become an integral part of this overall strategy. However, without their support and active involvement, DDoS attacks will continue to have a defining impact into the future.

Future Work

Further experimental work should include the testing of critical infrastructure networks with existing and emerging attack techniques. Analysing the impact of application layer attacks is particularly important, due to their low level of detection, the reluctance to block such traffic due to increasing dependence, and the overall level of effectiveness of this type of attack.

Experimentation into proposed solutions to mitigate the vulnerability of SCADA systems is also encouraged. This includes further testing on the implementation of a SCADA Security Device (SSD). The proposal by Rodrigues, Best & Pendse (2011) utilises Transport Layer Security (TLS) encryption and conducts a TCP handshake followed by a TLS handshake, prior to communicating with TCP or UDP securely over TLS. It is believed this may mitigate much of the threat that DDoS attacks present, amongst other such cyber-security attacks.

There is already substantial historical evidence demonstrating the ease with which these remote systems have been exploited in past decades. As SCADA networks take advantage of increasing connectivity through the Internet and their systems become more and more interconnected, the number of threat surfaces available in which a cyber-attack can reach and exploit these systems increases.

A final recommendation includes further research into a system capable of characterising packets to determine legitimacy. A proposed solution would see border routers holding incoming packets and creating copies of packet headers. These headers would then be passed on to a dedicated server to undertake inspection of traffic prior to a border router by either releasing the packets or dropping them altogether. Such a process can become considerably complex, and requires integration between routers and specialised servers, in addition to pre-defined classifications of packets. However, through a streamlined approach and a globally recognised protocol, owners of border routers such as Internet Service Providers (ISPs) will be in a stronger position to deny illegitimate traffic internally, and share, externally, information on cyber threats as they are identified.

References

Aamir, M & Arif, M 2013, 'Study and performance evaluation on recent DDoS trends of attack & defense', *Information Technology and Computer Science*, vol. 8, pp. 54-65.

Arora, K, Kumar, K & Sachdeva, M 2011, 'Impact analysis of recent DDoS attacks', *International Journal on Computer Science and Engineering*, vol. 3, no. 2, pp. 877-84.

Asri, S & Pranggono, B 2015, 'Impact of distributed Denial-of-Service attack on advanced metering infrastructure', *Wireless Personal Communications*, vol. 83, no. 3, pp. 2211-23.

Bergin, D 2015, 'Cyber-attack and defense simulation framework', *Journal of Defense Modelling and Simulation: Applications, Technology,* vol. 12, no. 4, pp. 383-92.

Fordyce, R 2013, 'DDoS Attacks as Political Assemblages', *PLATFORM: Journal of Media and Communication*, vol. 5, no. 1, pp. 6-20.

Gupta, B, Joshi, R & Misra, M 2010, 'Distributed Denial of Service prevention techniques', *International Journal of Computer and Electrical Engineering*, vol. 2, no. 2, pp. 268-76.

Leiner, B, Cerf, V, Clark, D, Kahn, R, Kleinrock, L, Lynch, D, Postel, J, Roberts, L & Wolff, S 2009, 'A brief history of the Internet', *ACM SIGCOMM Computer Communication Review*, vol. 39, no. 5, pp. 22-31.

Mirkovic, J, Fahmy, S, Reiher, P & Thomas, R 2009, 'How to test DoS defenses', *2009 Cybersecurity Applications & Technology Conference for Homeland Security, CATCH'09. Cybersecurity Applications & Technology, IEEE 2009*, pp. 103-17.

Nwaocha, V & Inyiama, H 2011, 'Establishing an effective combat strategy for prevalent cyber-attacks', *International Journal of Computer Science and Information Security*, vol. 9, no. 5, pp. 142-8.

Ponemon Institute 2012, *Cyber security on the offense: a study of IT security experts*, pp. 1-29, viewed 18 May 2015, <http://security.radware.com/uploadedFiles/Resources_and_Content/Attack_Tools/CyberSecurityontheOffense.pdf>.

Rodrigues, A, Best, T & Pendse, R 2011, 'SCADA security device: design and implementation', *Cyber Security and Information Intelligence Research: Proceedings of the Seventh Annual Workshop*, pp. 1-4.

Stallings, W & Brown, L 2012, *Computer security principles and practice*, Pearson, Upper Saddle River, New Jersey, USA.

Thompson, M 2012, 'The cyber threat to Australia', *Australian Army Journal*, no. 188, pp. 57-66.

Tomar, K & Tyagi, S 2014, 'Quantifying the impact of flood attack on transport layer protocol', *International Journal on Computational Sciences & Applications*, vol. 4, no. 6, pp. 79-87.

Yang, J, Park, M & Chung, T 2013, 'A study on low-rate DDoS attacks in real networks, *2013 International Conference on Information Science and Applications*, pp.1-4.

Zargar, S, James, J & Tipper, D 2013, 'A survey of defense mechanisms against Distributed Denial of Service (DDoS) flooding attacks', *IEEE Communications Surveys & Tutorials*, vol. 15, no. 4, pp. 2046-69.

Critical Infrastructure: Where Do We Stand Today? A Comprehensive and Comparative Study of the Definitions of a Critical Infrastructure

C Gallais[1], E Filiol[2]

[1]*TEVALIS*
Rennes, France
E-mail: cecilia.gallais@tevalis.fr

[2]*E.S.I.E.A*
Laval, France
E-mail: filiol@esiea.fr

Abstract: *The concept of 'critical infrastructure' has become a key issue as far as the cyber dimension is concerned. All industrialized nation-states that depend on information and communication technologies have defined this concept or established a list of critical sectors to identify their critical infrastructures. Despite the high number of definitions, none of them considers a realistic view of a critical infrastructure as it tends to be reduced to its simple computerized dimension. The survey of definitions of critical infrastructure presented in this paper highlights the omissions in these definitions. This paper suggests the need for a new definition of critical infrastructure—a definition which includes the missing elements identified herein.*

Keywords: *Critical Infrastructure, Definition, Human Component, External Component, Protection, Security*

Introduction

National plans for the protection of 'critical infrastructures' thrive pretty much everywhere—Australia, Canada, Japan, Germany, the United States of America, the European Union, and even in the African nation-states of Mauritius and Kenya—which proves the importance of the concept in contemporary security thought. Most of these plans include a definition of critical infrastructure, since defining the term is the first logical step before implementing plans or programs to defend it.

Historically, the first definition of 'critical infrastructure' appeared in Presidential Decision Directive (PDD) 63 dating back to 1998 in the U.S. At that time, a critical infrastructure consisted of those physical and cyber-based systems that were essential to the minimum operations of the economy and the government. Since the appearance of this initial definition, several others have followed.

Despite the variety and the great number of definitions, no single definition provides a complete and accurate description of what constitutes a critical infrastructure; important components are

left out. To highlight these omissions, the authors compiled a survey of the definitions of critical infrastructure. The survey is principally based on the *International CIIP handbook of 2008/2009* (Brunner & Suter 2009) and its previous versions (Wenger, Metzger & Dunn 2002; Dunn & Wigert 2004; Abele-Wigert & Dunn 2006). Even if the subject of these documents is the critical information infrastructure—which can be seen in broad outline as a part of critical infrastructure and "refers exclusively to the security and protection of the IT connections and IT solutions within and between the individual infrastructure sectors" (German Federal Office for the Security of Information Technologies 2004)—critical infrastructures are mentioned and several definitions are provided. However, there is no official distinction between critical infrastructure and critical information infrastructure, so the terms become interchangeable in some countries.

The first part of this paper surveys the definitions of critical infrastructure and relies on a document from the Organization for Economic Co-operation and Development (Gordon & Dion 2008). In the second part of the paper, different conclusions regarding the various definitions of critical infrastructure are presented, and, finally, a new definition of critical infrastructure is proposed.

What Is a Critical Infrastructure?

Twenty definitions of critical infrastructure are presented in this section, including eleven from European nation-states and organisations, three from American nation-states and organisations, five from Asian and Pacific nation-states and organisations, and one from an African nation-state.

Each of the definitions presented is the most recent that can be found. Of course, with the great number of documents on critical infrastructure published globally (especially on the protection of such), it is quite difficult to be sure if the chosen definitions are really the most recent ones.

European nation-states and organisations

For the European Union, the Council Directive 2008/114/CE defines critical infrastructure as

> an asset, system or part thereof located in Member States which is essential for the maintenance of vital societal functions, health, safety, security, economic or societal well-being of people, and the disruption or destruction of which would have a significant impact in a Member State as a result of the failure to maintain those functions. (Council of the European Union 2008)

The Austrian Federal Chancellor defines critical infrastructures as

> natural resources; services; information technology facilities; networks; and other assets which, if disrupted or destroyed would have serious impact on the health, safety, or economic well-being of the citizens or the effective functioning of the Government. (Austrian Federal Chancellor 2006)

In Belgium, as stipulated by law on 1 July 2011, a critical infrastructure is

> an installation, system or part thereof, of federal interest, which is essential for the maintenance of vital societal functions, health, safety, security, economic or societal well-being of people, and which, if disrupted or destroyed, would have a significant impact. (Service Public Fédéral Intérieur 2011)

The National Strategy for Critical Infrastructure Protection of Germany defines critical infrastructure as

> the organizational and physical structures and facilities of such vital importance to a nation's society and economy that their failure or degradation would result in sustained supply shortages, significant disruption of public safety and security, or other dramatic consequences. (Federal Ministry of the Interior 2009)

The Hungarian definition of a critical infrastructure is based on the definition of the European Union: critical infrastructures are

> the interconnected, interactive, and interdependent infrastructure elements, establishments, services, and systems that are vital for the operation of the national economy and public utilities to maintain an acceptable level of security for the nation, individual lives, and private property, as well as concerning the maintenance of the economy, the public health services, and the environment. (Brunner & Suter 2009)

For the Netherlands, critical infrastructure includes

> the business enterprises and public bodies that provide the goods and services essential for the day-to-day lives of most people in the Netherlands. (Government of the Netherlands)

The report NOU 2006:6, about the protection of critical infrastructures and critical societal functions in Norway, defines critical infrastructures as

> the facilities and systems that are necessary to maintain the functions that are critical for society. These functions cover basic needs in the society and contribute to a sense of safety in the population. (Royal Norwegian Ministry of Government 2007)

In Poland, the National Critical Infrastructure Protection Program defines critical infrastructure according to the Act on Crisis Management: critical infrastructure shall be understood as

> the systems and functional sites forming their part which are mutually related, such as building sites, facilities, installations, key services for the safety of the state and its citizens and serving to ensure efficient functioning of the public administration authorities, as well as institutions and entrepreneurs. (Rządowe Centrum Bezpieczeństwa 2015)

In Spain, the Law 8/2011 defines critical infrastructure as

> those installations, networks, systems, physical equipment, and information technologies, whose interruption or destruction would have a grave impact on the health, security, social or economic well-being of citizens or on the efficient functioning of the state institutions and of the public administration. (Jefatura del Estado 2011)

The Federal Council's Basic Strategy for Critical Infrastructure Protection defines critical infrastructure in Switzerland as

> infrastructures whose disruption, failure, or destruction would have a serious impact on public health, public and political affairs, the environment, security, and social and economic well-being. (Federal Council of Switzerland 2012)

The same definition can be found in a more recent article by Brem (2011).

The United Kingdom's critical national infrastructure is defined by the government as

> those facilities, systems, sites, information, people, networks and processes, necessary for a country to function and upon which daily life depends. It also includes some functions, sites and organizations which are not critical to the maintenance of essential services, but which need protection due to the potential danger to the public (civil nuclear and chemical sites for example). (Centre for the Protection of National Infrastructure)

American nation-states and organisations

In Canada, the National Strategy for Critical Infrastructure defines critical infrastructure as

> processes, systems facilities, technologies, networks, assets and services essential to the health, safety, security or economic well-being of Canadians and the effective functioning of government. Critical infrastructure can be stand-alone or interconnected and interdependent within and across provinces, territories and national borders. Disruptions of critical infrastructure could result in catastrophic loss of life, adverse economic effects, and significant harm to public confidence. (Canadian Government 2010).

This definition can also be found on the Public Safety Canada website (Government of Canada).

The U.S. Patriot Act of 2001 defines critical infrastructure as

> systems and assets, whether physical or virtual, so vital to the United States that the incapacity or destruction of such systems and assets would have a debilitating impact on security, national economic security, national public health or safety, or any combination of those matters. (United States Congress 2001)

According to the Department of Homeland Security's website, this definition is still currently used (U.S. Department of Homeland Security).

During a session about the protection of critical infrastructures in 2007, the NATO Parliamentary Assembly admitted that there is no universally agreed upon definition of a critical infrastructure. But this term "is generally understood as those facilities and services that are vital to the basic operations of a given society, or those without which the functioning of a given society would be greatly impaired" (NATO Parliamentary Assembly 2007).

Asian and Pacific nation-states and organisations

The Asia-Pacific Telecommunity, in the report of the South Asian Telecommunication Regulator's Council (SATRC) titled *Critical information infrastructure protection and cyber security* and adopted in April 2012, defines critical infrastructure as

> the computers, computer systems and/or networks, whether physical or virtual, and/or the computer programs, computer data, content data and/or traffic data so vital to a country that the incapacity or destruction of or interference with such systems and assets would have a debilitating impact on security, national or economic security, national public health and safety, or any combination of those matters. (SATRC Working Group on Policy and Regulations 2012)

In Australia, the Attorney-General's Department website says that critical infrastructure

> delivers services essential to our daily lives, such as power, water, health services, communications systems and banking. (Attorney-General's Department of the Australian Government)

The Trusted Information Sharing Network's website provides more information as it defines Australian critical infrastructure as

> those physical facilities, supply chains, information technologies and communication networks, which if destroyed, degraded or rendered unavailable for an extended period, would significantly impact on the social or economic wellbeing of the nation, or affect Australia's ability to conduct national defense and ensure national security. (Trusted Information Sharing Network)

The same definition can be found in the *Critical infrastructure resilience strategy* (Australian Government 2010).

Critical infrastructure in Japan is defined by *The second action plan on information security measures for critical infrastructures*. According to this plan, critical infrastructure is

> the basis of people's social lives and economic activities formed by business that provide services which are extremely difficult to be substituted by others. If its function is suspended, deteriorated or become[s] unavailable, it could have significant impacts on people's social lives and economic activities. (Information Security Policy Council 2009)

For New Zealand, a presentation at the International Disaster and Risk Conference in Davos, defined critical infrastructure as

> infrastructure necessary to provide critical services, whose interruption would have a serious adverse effect on New Zealand as a whole or on a large proportion of the population, and which would require immediate reinstatement. (Helm 2008)

Malaysia calls a critical infrastructure a Critical National Information Infrastructure (CNII) and defines it as

> those assets (real and virtual), systems and functions that are vital to the nations that their incapacity or destruction would have a devastating impact on:
> - National economic strength; Confidence that the nation's key growth area can successfully compete in global market while maintaining favourable standards of living.
> - National image; Projection of national image towards enhancing stature and sphere of influence.
> - National defence and security; guarantee sovereignty and independence whilst maintaining internal security.
> - Government capability to functions; maintain order to perform and deliver minimum essential public services.
> - Public health and safety; delivering and managing optimal health care to the citizen. (Critical National Information Infrastructure [CNII] 2016)

African nation-states

The Cabinet Secretary Interior and Co-ordination of National Government defines a critical infrastructure as "the totality of Critical Infrastructure Assets"; the critical infrastructure assets are the

> designated physical and virtual assets or facilities, whether owned by private or public entities which are designated as such under this Act as essential to the provision of vital services to Kenyans for their social and economic wellbeing, and which if destroyed, degraded or rendered unavailable, would impact on the social or economic wellbeing of the nation or affect Kenya's ability to conduct national defense and security. (Nkaissery 2015)

List of critical sectors

As stated in national plans for the protection of critical infrastructures, the role of the critical sectors is to "facilitate identification, prioritization, assessment and protection of critical information infrastructure through information sharing and reporting" (Republic of Mauritius 2016). Therefore, it is not a surprise when a definition of critical infrastructure is almost always followed by a list of critical sectors.

The first mention of critical infrastructure sectors is found in the Executive Order 13010 of July 15, 1996 (Clinton 1996). This list identified the sectors which were necessary to the effective functioning of the society.

The list of critical sectors tends to be specific to each nation-state or organisation. A sector may be included for historical, geographic, socio-political, or traditional reasons, which can explain the differences between the lists of critical sectors. Lists of critical sectors are presented in the Appendix of this paper (see **Tables 1-8**) from nearly forty-six nation-states (such as the United States, Germany, Switzerland, India, and Kenya) and organisations (such as the European Union and the Asia-Pacific Telecommunity).

Some nation-states have developed a list of critical sectors without ever having defined the term critical infrastructure. In the case of the members of the European Union, they may not have a definition of critical infrastructure of their own because they are probably content with the European Union's definition. For other nation-states, it is not that easy to explain the absence of a definition.

However, some nation-states such as Austria do not have an official list of critical sectors—the one given in this paper was developed by some experts and is not an official definition from the Austrian government—but they do have a definition of critical infrastructure.

Despite the great variety of lists of critical sectors, most of the nation-states and organisations seem to agree on the importance of specific critical sectors. As can be seen in the histogram below (**Figure 1**), the transport sector is mentioned in more than 95% of the lists of critical sectors that were gathered for this study; the energy sector is mentioned in more than 86% of the lists of critical sectors that were gathered for this study; and, the communication technology sector is mentioned in more than 84% of the lists of critical sectors that were gathered for this study. Two other critical sectors——finance and water—follow closely behind in regular appearance on these lists.

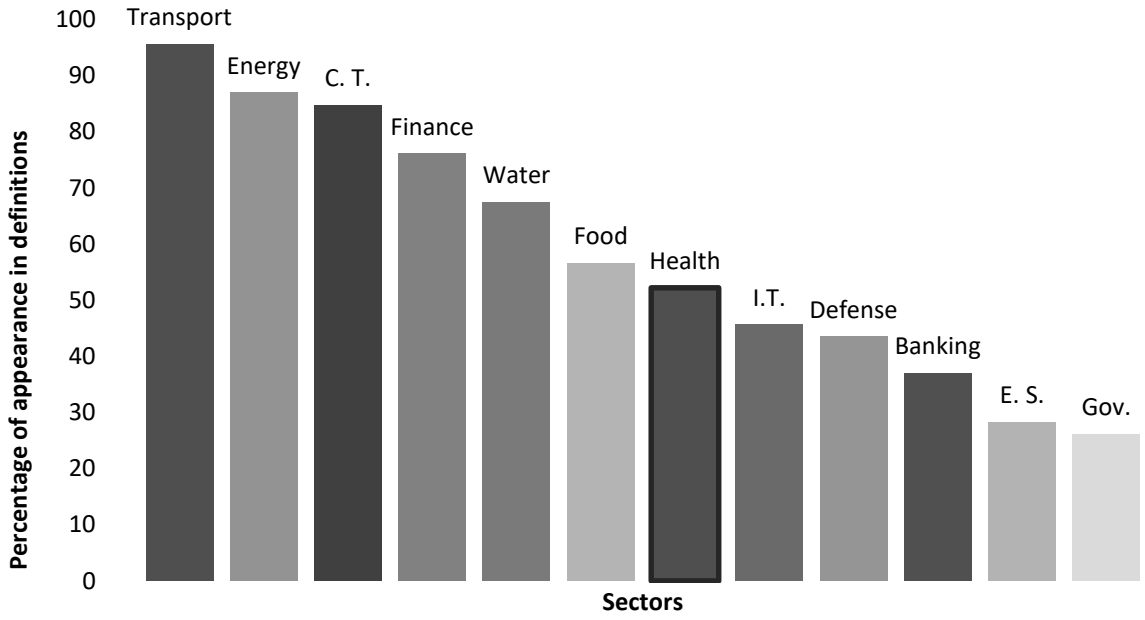


Figure 1: Histogram of the most cited sectors in lists of critical sectors

(In **Figure 1**, above, C.T. stands for Communication Technology, I.T. for Information Technology, E.S. for Emergency Services, and Gov. for Government.)

What can be said about these definitions?

There are more definitions of critical infrastructure and lists of critical sectors. Furthermore, most of the definitions have undergone many modifications, and they will certainly undergo others as more and more actors understand and grasp the importance of this concept. For example, Belgium-based Elia has its own definition of critical infrastructure, which can be found on its website (Elia).

Also, in all of the documents gathered for this survey, the definition of a critical infrastructure is usually divided into two parts. First there is the list of its components; and, second, there are the consequences of its disruption, damage, or destruction.

The Components of a Critical Infrastructure

Many components of a critical infrastructure are identified in the different definitions presented above, including assets, systems, or networks, but the list of these components tends to become shortened in length over time. For example, the 2005 *Green paper on a European programme for critical infrastructure protection* (Commission of the European Communities) gives a more complete definition than the one from the 2008 Council Directive 2008/114/CE (Council of the European Union). In fact, the list of the components of a critical infrastructure is the part of the definition which differs the most from nation-state to nation-state.

In the following section, the components missing from these definitions are identified, and the consequences of their absence is discussed.

Missing components

The following omissions have been noted in at least one other source: Filiol (2011). Among all of the components cited in the various definitions, the absence of human factors is perhaps most immediately noticeable. Only one of the definitions presented in this paper mentions humans as part of a critical infrastructure—despite the fact that humans are essential for the functioning of every existing infrastructure, critical or not. The UK is the only nation-state that clearly includes the human factor as a component of a critical infrastructure. Some may say that a system, which is a component whose presence is acknowledged in many definitions, could be defined as being comprised of people, processes, and technology. That does not change the fact that humans are not clearly stated as a component of a critical infrastructure, and thus the definitions can mislead those in charge of critical infrastructure security with respect to the importance of people.

Also of note is the lack of some ‘intelligence’ perspective, thus allowing a broader and more operational view as far as the cyber dimension is concerned. As an example, no existing definition takes interdependencies with external components into account, thus providing only a very narrow-minded view of an infrastructure, which is considered then only as a completely isolated structure. Indeed, even though some of the definitions mention the concept of interdependency—such as those from Canada, Hungary, and Poland—the interdependencies taken into account are only the ones within the critical infrastructure itself or with other critical

infrastructures, but never with basic infrastructures—including subcontractors, suppliers, data-centres, or others.

In addition, a critical infrastructure's environment—both cultural and political—should be considered. Attackers could use these environments to trigger a strike, for example, which could disturb the transport of needed resources or finished products.

The consequences of these absences

As stated previously, the human component is missing from almost all of the definitions which are presented above, despite the fact that people are essential for the functioning of critical infrastructures. Mitnick and Simon (2003) consider humans to be the weakest link of security. Mitnick and Simon's work demonstrates that, despite the use of the best possible security protection items, it is possible for an attacker to obtain access to critical information or critical objects simply by using social engineering techniques.

As an example, Mitnick and Simon (2003) show how an attacker, or a manipulator in this case, can obtain a username and the corresponding password simply by asking its owner for the information while pretending to be a staff member of the organisation's information security office. And, with this username and password, the manipulator has everything he or she needs to gain access to the company's network and locate the elements he or she is looking for.

While these stories may not be persuasive enough due to their fictional nature, the real-life cases involving Edward Snowden and Wikileaks demonstrate that humans can be a major flaw in any security scenario. In these cases, however, not much can be done to prevent employees working of their own free will from sharing confidential information with others.

It is interesting to note that a definition of an infrastructure, dating back to 1996, briefly mentions this human component. Indeed the Executive Order 13010 defined infrastructure as

> "the framework of interdependent networks and systems comprising identifiable industries, institutions (including *people* and procedures) [emphasis added], and distribution capabilities that provide a reliable flow of products and services essential to the defense and economic security of the United States, the smooth functioning of government at all levels, and society as a whole". (qtd. in Moteff & Parfomak 2004)

Since then, the human component has been cast aside in definitions.

In addition to the absence of the human component in definitions of critical infrastructure, these definitions also lack references to external components and environments. (Of note, however, is the work completed by Filiol and Raynal (2009), which includes both the human component and external components/environments in the planning of an attack to delay the departure of a military ship.)

The absence of these components in an official definition effectively creates a weak link that can be used by attackers to target security policies and achieve their goals. Without mention (and

defence) of these two components, definitions of critical infrastructures are more vulnerable to outside actors.

Distinction between a critical and a basic infrastructure

The distinction between a critical and a basic infrastructure is an element which does appear clearly in all of these definitions—it is in the criticality of the consequences of their disruption, damage, or destruction. The definitions differ when it comes to the domains on which the disruption or destruction may have serious consequences, although some of the domains appear frequently, such as public safety, public security, or the social and economic well-being of citizens. However, the notion of criticality is always there.

The criticality of an infrastructure depends as much on the infrastructure itself as on its relations with other infrastructures. So two kinds of criticality are identified: the inherent criticality that occurs when an infrastructure is critical in and of itself, and the external criticality that occurs when an infrastructure is critical for other infrastructures.

Identifying inherent critical infrastructure can be more straightforward than identifying external critical infrastructures for a few reasons. First, some infrastructures, such as energy suppliers, come to mind instantaneously. Also, the lists of critical sectors are useful in identifying inherent critical infrastructure. In the end, however, nothing ensures that all inherent critical infrastructures will be easily identified.

What is certain is that identifying external critical infrastructures is more challenging than identifying inherent ones. This is the case because infrastructures are classified as external infrastructures only after a first disaster has occurred. Therefore, it can be difficult, even impossible, to predict the consequences of a disaster (be it disruption, damage, or destruction). As a result, this identification and classification is mostly done *a posteriori*.

The importance taken by the critical sectors

Identifying critical sectors is a necessary and important precursor to identifying and defining critical infrastructure. Many nation-states, like Estonia, Finland, France, Italy, or Sweden, have a list of critical sectors but no official definition of ‘critical infrastructure’. And several others have proposed a definition many years after they presented a list of critical sectors. In fact, establishing the list of critical sectors seems to have taken priority over establishing a definition of critical infrastructure. Some nation-states use their list of critical sectors to identify at least their inherent critical infrastructures, even if they still ignore identifying their external critical infrastructures.

In addition, problems and challenges still arise once a list of critical sectors is established. These lists, like definitions of critical infrastructure, are dynamic, rather than static, things. In real time, lists and definitions change and have impacts on one another. The list of critical sectors tends to undergo more modifications than the definition does. As sectors are added or removed, their names are changed; subsectors may be present or not; sectors may become subsectors; or subsectors may become sectors (for example, emergency services). Often, complex lists of critical sectors attend simple definitions of critical infrastructures.

While important, lists of critical sectors are not enough to ensure protection of critical infrastructures. A complete definition of critical infrastructure (a definition that specifies what must be protected) is necessary to guarantee security. The following section suggests the necessary components of that definition.

Discussion Toward an Enlarged and More Suitable Definition

The actual definitions of critical infrastructure appear restrictive, static, and local, as they are mainly dictated by the vision of the defender. So, to go against this trend, a definition is proposed that is based on the vision of the attacker, a definition which takes into account the greatest number of elements possible.

Critical infrastructure can be a company, an institution, an organisation, facilities (U.S. Department of Transportation 2002), services, and equipment (Filiol & Gallais 2016), whether regional, national, or international, which, if disrupted, damaged, or destroyed, would have a serious impact on the health, safety, security, or economic well-being of citizens or the effective functioning of governments and others infrastructures depending on it. It includes humans, which, if co-opted, diverted, or eliminated, could lead to the disruption, damage, or destruction of the critical infrastructure.

It also includes:

- installations (such as access, buildings, sites);
- equipment (for example, computers, printers, hard drives);
- resources, whether physical or natural;
- networks, whether physical (like electricity or water) or virtual (such as the Intranet, or the Internet);
- data, whether physical or virtual (confidential data, like passwords or access codes, procedures, organisation charts);
- information and communication technology facilities;
- services;
- processes;
- assets, including the corporate image;
- systems or parts thereof;
- other infrastructures with which strong dependencies exist (suppliers of services or products, for example) (Filiol & Raynal 2009), which if disrupted, damaged, or destroyed would have a serious impact on the health, safety, security, or well-being of a population (including employees) or the effective functioning of the critical infrastructure.

In fact, it includes any element which could lead to the disruption, damage, or destruction of the critical infrastructure. These elements can also be found in the political and cultural environment of the infrastructure.

Surely all of an infrastructure's elements which can lead to its disruption, damage, or destruction cannot be identified and enumerated, as this task is impossible, since the security aspect lies in the ability of the attackers to be innovative, creative, and, in essence, unpredictable. Attackers

can turn an element that is thought to be inoffensive into a weapon. This may explain why some definitions, such as those used by the Swiss and the Dutch, do not go into detail regarding the components of a critical infrastructure.

Conclusion

In this paper, the authors have presented a variety of definitions of 'critical infrastructure'. They have also identified the key weaknesses of those definitions (the absence of the human component and the lack of an 'intelligence' perspective). The authors have also suggested the security consequences of these weaknesses. In essence, these omissions can lead to the failure of cyber defence policies and techniques, as they create clear breaches in the security of critical infrastructure—breaches which are easy for attackers to use to advantage (Filiol & Gallais 2016; Gallais 2017).

This is why a new definition of critical infrastructure is proposed, based on the attacker's point of view rather than on the defender's point of view. The former includes the human component and the external components and some other elements such as the political and cultural environments.

Because it is quite difficult to have a clear vision of what a critical infrastructure actually is the next step is to design a methodology which is capable of finding visible and hidden weaknesses to make this task less impossible.

References

Abele-Wigert, I & Dunn, M 2006, *International CIIP handbook 2006*, vol. 1, Center for Security Studies, ETH Zurich, Switzerland.

Associazione Italiana esperti in Infrastrutture Critiche 2011, *AIIC - Italian Association of Critical Infrastructures' Experts*, viewed 5 December 2016, <http://www.infrastrutturecritiche.it/aiic-en/>.

Attorney-General's Department of the Australian Government, *Critical infrastructure resilience*, viewed 5 December 2016, <http://www.ag.gov.au/NationalSecurity/InfrastructureResilience/Pages/default.aspx>.

Australian Government 2010, *Critical infrastructure resilience strategy*, viewed 5 December 2016, <http://ccpic.mai.gov.ro/docs/Australian+Government+s+Critical+Infrastructure+Resilience+Strategy.pdf>.

Austrian Federal Chancellor 2006, *Anfrage*, viewed 5 December 2016, <http://www.parlament.gv.at/PAKT/VHG/XXII/J/J_04641/imfname_067709.pdf>.

Brem, S 2011, 'The Swiss programme on critical infrastructure protection', *The CIP Report*, The Center for Infrastructure Protection and Homeland Security, vol. 9, no. 12, pp. 23+, viewed 2 January 2017, < http://cip.gmu.edu/wp-content/uploads/2013/06/107_The-CIP-Report-June-2011_International.pdf>.

Brunner, EM & Suter, M 2009, *International CIIP handbook 2008/2009,* Center for Security Studies, ETH Zurich, Switzerland.

Canadian Government 2010, *National strategy for critical infrastructure,* viewed 25 December 2016, <http://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/srtg-crtcl-nfrstrctr/srtg-crtcl-nfrstrctr-eng.pdf>.

Centre for the Protection of National Infrastructure, *Critical national infrastructure,* viewed 5 December 2016, <http://www.cpni.gov.uk/about/cni/>.

Clinton, WJ 1996, Executive Order EO 13010 on Critical Infrastructure Protection, 15 July, viewed 19 March 2017, <https://fas.org/irp/offdocs/eo13010.htm>.

Commission of the European Communities 2005, *Green paper on a European programme for critical infrastructure protection,* viewed 5 December 2016, <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52005DC0576&from=EN>.

Council of the European Union 2008, *Council Directive 2008/114/CE of 8 December 2008 on the identification and designation of European critical infrastructure and the assessment of the need to improve their protection,* viewed 5 December 2016,<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:345:0075:0082:EN:PDF>.

Critical National Information Infrastructure (CNII) 2016, *CNII Portal,* viewed 2 December 2016, <http://cnii.cybersecurity.my/main/about.html>.

Dunn, M & Wigert, I 2004, *International CIIP Handbook 2004: an inventory and analysis of protection policies in fourteen countries,* Center for Security Studies, ETH Zurich, Switzerland.

Elia, *Critical infrastruture,* viewed 4 December 2016, <http://www.elia.be/en/safety-and-environment/Security/critical-infrastructure>.

Federal Council of Switzerland 2012, *The Federal Council's basic strategy for critical infrastructure protection,* viewed 5 December 2016, <http://www.babs.admin.ch/content/babs-internet/en/aufgabenbabs/ski/_jcr_content/contentPar/tabs/items/downloads/tabPar/downloadlist/downloadItems/67_1460980334945.download/natstratski2012de.pdf>.

Federal Ministry of the Interior 2009, *National strategy for critical infrastructure protection,* viewed 5 December 2016, <http://www.bmi.bund.de/cae/servlet/contentblob/598732/publicationFile/34423/kritis_englisch.pdf>.

Filiol, E 2011 'Operational aspects of a cyberattack: intelligence, planning and conduct', *Cyberwar and information warfare,* ed. D Ventre, Wiley & Sons, Hoboken, NJ, U.S.A., pp. 245-284.

——& Raynal, F 2009, Cyberguerre : de l'attaque du bunker à l'attaque dans la profondeur. *Revue Défense Nationale,* no.3, mars, p. 74.

——& Gallais, C 2016, *Combinatorial optimization of operational (cyber) attacks against large-scale critical infrastructures - the vertex cover approach*, 11th International Conference on Cyber Warfare and Security (ICCWS) 2016, Boston, March 17-18th, 2016.

Gallais, C 2017, *Formalization, algebraic and combinatorial analysis of generalized cyber attack scenario*, Ph D thesis, ENSAM, Paris.

German Federal Office for the Security of Information Technologies 2004, *Critical infrastructure protection : survey of world-wide activities*, viewed 16 July 2013, <https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/Kritis/paper_studie_en_pdf.pdf?__blob=publicationFile>.

Gordon, K & Dion, M 2008, *Protection of 'critical infrastructure' and the role of investment policies relating to national security*, viewed 14 August 2013, <http://www.oecd.org/daf/inv/investment-policy/40700392.pdf>.

Government Centre for Security, *National critical infrastructure protection programme adopted by the Council of Ministers*, viewed 25 July 2013, <rcb.gov.pl/eng/?p=791>.

Government of Canada, *Public safety Canada / critical infrastructure*, viewed 4 December 2016, <http://www.publicsafety.gc.ca/cnt/ntnl-scrt/crtcl-nfrstrctr/index-eng.aspx>.

Government of the Netherlands, *Crisis, national security and terrorism : protecting critical infrastructure*, viewed 16 July 2013, <http://www.government.nl/issues/crisis-national-security-and-terrorism/protecting-critical-infrastructure>.

Government 2000, *Emergency preparedness act (Estonia)*, viewed 5 December 2016, <http://www.ifrc.org/docs/idrl/233EN.pdf>.

Government 2006, *Arrêté du 2 juin 2006 fixant la liste des secteurs d'activités d'importance vitale et désignant les ministres coordonnateurs desdits secteurs*, viewed 5 December 2016, <http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000423259&dateTexte=&categorieLien=id>.

Härkönen, T 2007, *The Finnish critical infrastructure protection; state crisis management model and situation awareness*, viewed 8 October 2013, <http://www.helsinki.fi/aleksanteri/civpro/thdata/cip.htm >.

Helm, P 2008, *Critical infrastructure resilience : perspective from New Zealand.* viewed 5 December 2016, <http://idrc.info/fileadmin/user_upload/idrc/former_conferences/idrc2008/presentations2008/Helm_Patrick_Owen_Critical_Infrastructure_Protection_A_Perspective_from_New_Zealand.pdf>.

Information Security Policy Council 2009, *The second action plan on information security measures for critical infrastructures*, viewed 5 December 2016,<http://www.nisc.go.jp/eng/pdf/actionplan_ci_eng_v2.pdf >.

Jefatura del Estado 2011, *Ley 8/2011, de 28 de Abril, por la que se establecen medidas para la protectión de las infraestructuras críticas*, viewed 5 December 2016,< http://www.boe.es/boe/dias/2011/04/29/pdfs/BOE-A-2011-7630.pdf>.

Mitnick, KD & Simon, WL 2003, *The art of deception*, Wiley & Sons, Hoboken, NJ, U.S.A.

Moteff, JD 2011, *Critical infrastructures: background, policy, and implementation*, Congressional Research Service, Washington, D.C., U.S.A.

——& Parfomak, P 2004, *Critical infrastructure and key assets: definition and identification*, Congress Research Services, Washington, D.C., U.S.A.

NATO Parliamentary Assembly 2007, *162 CDS 07 E rev 1—The protection of critical infrastructures*, viewed 4 December 2016, <http://www.nato-pa.int/default.asp?CAT2=1159&CAT1=16&CAT0=2&COM=1165&MOD=0&SMD=0&SSMD=&STA=0&ID=0&PAR=0&LNG=0>.

Nkaissery, J 2015, *The critical infrastructure protection bill 2015*, viewed 2 December 2016, <http://www.icta.go.ke/downloads/critical-bill.pdf>.

Republic of Mauritius 2016, *National cyber security strategy 2014-2019*, viewed 2 December 2016, <https://www.itu.int/en/ITU-D/Cybersecurity/Documents/National_Strategies_Repository/Mauritius_2014_National%20Cyber%20Security%20Strategy%20-%202014%20-%20EN.pdf>.

Royal Norwegian Ministry of Government 2007, *National guidelines on information security 2007-2010*, viewed 5 December 2016, <https://www.oecd.org/norway/41671072.pdf>.

Rządowe Centrum Bezpieczeństwa 2015, *The national critical infrastructure protection programme*, viewed 5 December 2016, <http://rcb.gov.pl/wp-content/uploads/NPOIK-2015_eng-1.pdf>.

SATRC Working Group On Policy And Regulations 2012, *SATRC report on critical information infrastructure protection and cyber security*, Asia-Pacific Telecommunity (APT), viewed 2 January 2017, <http://www.apt.int/sites/default/files/Uploadfiles/SATRC/SAPIII%20Outcomes/SATRC-SAPIII-04_WG_PR_Cybersecurity_Report.doc.

Service Public Fédéral Intérieur 2011, 1er Juillet 2011 - Loi relative à la sécurité et la protection des infrastructures critiques, *Moniteur Belge N.205, Vendredi 15 Juillet 2011, Deuxième édition*, 15 July, p. 42320.

Trusted Information Sharing Network (TISN), *Critical infrastructure*, viewed 5 December 2016, <http://www.tisn.gov.au/Pages/Critical_infrastructure.aspx>.

United States Congress 2001. Patriot Act of 2001, 107th Congress, 1st Session (H.R. 3162), 24 October, Section 1016, viewed on 16 March 2017, <http://www.usapatriotact.com/USA_Patriot_Act.doc>.

U.S. Department of Homeland Security, *What is critical infrastructure?*, viewed 4 December 2016, <http://www.dhs.gov/what-critical-infrastructure>.

U.S. Department of Transportation 2002, *Effects of catastrophic events on transportation system management and operations*, viewed 2 January 2017, <http://ntl.bts.gov/lib/jpodocs/repts_te/13754_files/13754.pdf>

Wenger, A, Metzger, J & Dunn, M 2002, *International CIIP handbook 2002, an inventory of protection policies in eight countries*, Center for Security Studies, ETH Zurich, Switzerland.

Appendix

The critical sectors of American, African, Asian, European, and Pacific nation-states are presented in the following tables. To most effectively present the data, some critical sectors were merged or renamed because they were very similar or one was a sub-sector of another. For example, Railways and Aviation were merged into Transport; likewise, Electrical Power in Energy, and Manufacturing were merged into Industry.

The results are divided into eight tables (**Tables 1-8**), as they cannot be easily presented in a single table. **Tables 1-3** show the critical sectors of nation-states with an established definition of critical infrastructure. They are presented in alphabetical order. **Tables 4-8** show the critical sectors of nation-states without an established definition of critical infrastructure. They are also presented in alphabetical order.

When the source of data is not indicated, the list of critical sectors comes from the document where the definition of critical infrastructure is found.

	APT	Australia	Austria	Belgium	Canada	E.U.
Administration						X
Banking	X	X				
Broadcasting						X
Chemical Industry						X
Communication Technology	X	X	X	X	X	X
Defense						X
Emergency Services	X					X
Energy	X	X	X	X	X	X
Finance	X	X		X	X	X
Food		X			X	X
Government	X				X	
Health		X			X	X
Industry					X	
Information Technology	X		X		X	X
Legal Order			X			X
Nuclear Industry						X
Oil and Gas	X					X
Public Order	X					X
Research						X
Safety/Security					X	X
SCADA Systems						X
Space						X
The Internet						X
Transport	X	X	X	X	X	X
Vital Goods			X			
Water	X	X			X	X

Table 1: Critical sectors of nation-states with definition, Asia-Pacific Telecommunity to European Union

Table Notes: "APT" stands for the Asia-Pacific Telecommunity. "E.U." stands for the European Union. The list of critical sectors of Canada comes from the *National strategy for critical infrastructure* (Canadian Government 2010).

	Germany	Hungary	Japan	NATO	Netherlands	New Zealand
Administration	X		X			
Banking		X				X
Communication Technology	X	X	X	X	X	X
Culture	X					
Cyber Infrastructure				X		
Defense		X		X	X	
Disaster Control & Management	X					
Emergency Services	X			X	X	X
Energy	X	X	X	X	X	X
Finance	X	X	X	X	X	X
Food	X	X		X		X
Government			X	X		X
Health	X	X	X	X	X	X
Industry		X				
Information Technology	X	X	X	X	X	
Insurance			X		X	
Legal Order	X			X		X
Logistics						X
Media	X					
Networks						X
Oil and Gas			X			X
Postal Services		X				
Public Order						X
Rescue Services	X				X	
Safety/Security		X			X	X
Space					X	
The Internet						X
Transport	X	X	X	X		X
Vital Goods					X	
Waste						X
Water	X	X	X	X	X	X

Table 2: Critical sectors of nation-states with definition, Germany to New Zealand

Table Notes: The list of critical sectors of NATO comes from *162 CDS 07 E rev 1—The protection of critical infrastructures* (NATO Parliamentary Assembly 2007).

	Norway	Poland	Spain	UK	USA	Switzerland
Administration		X	X			X
Agriculture					X	
Banking	X				X	
Chemical Industry		X	X		X	
Communication Technology	X	X	X	X	X	X
Defense	X				X	
Disaster Control & Management	X					
Emergency Services	X			X	X	
Energy	X	X	X	X	X	X
Environment	X					
Finance	X	X	X	X	X	X
Food	X	X	X	X	X	X
Government	X			X	X	
Health	X	X	X	X	X	X
Industry					X	X
Information Technology		X	X		X	X
Legal Order	X					
Materials					X	
National Monuments & Icons					X	
Nuclear Industry		X	X		X	
Oil and Gas	X					
Postal Services					X	
Public Order	X					
Rescue Services	X	X				
Research			X			
Safety/Security						X
Satellite-Based Infrastructures	X					
Social Welfare/Social Services	X					
Space			X			
Transport	X	X	X	X	X	X
Vital Goods						
Waste	X				X	X
Water	X	X	X	X	X	X

Table 3: Critical sectors of nation-states with definition, Norway to Switzerland

Table Notes: "UK" stands for the United Kingdom. "USA" stands for the United States of America. The list of critical sectors of the USA comes from *Critical infrastructures: background, policy, and implementation* (Moteff 2011). The list of critical sectors of Switzerland comes from "The Swiss programme on critical infrastructure protection" (Brem 2011).

	Argentina	Brazil	Chile	Czech Republic	Denmark	Estonia
Agriculture			X			
Banking		X				
Chemical Industry			X			
Communication Technology		X			X	X
Defense	X	X			X	
Emergency Services						X
Energy	X	X	X		X	X
Finance		X				X
Food			X			X
Health		X				X
Insurance						X
Media	X		X		X	
Postal Services					X	X
Public Order						X
Safety/Security		X				
Social Welfare/Social Services						X
Transport	X	X	X	X	X	X
Transectoral			X			
Water	X	X				

Table 4: Critical sectors of nation-states without definition, Argentina to Estonia

Table Notes: The list of critical sectors of Argentina, Chile, Czech Republic, and Denmark come from *Protection of 'critical infrastructure' and the role of investment policies relating to national security* (Gordon & Dion 2008). The list of critical sectors of Brazil comes from the *International CIIP handbook 2008/2009* (Brunner & Suter 2009). The list of critical sectors of Estonia comes from *Emergency preparedness act (Estonia)* (Government 2000).

	Finland	France	Greece	Iceland	India	Italy
Administration		X				
Agriculture			X	X		
Banking			X	X	X	X
Chemical Industry	X					
Communication Technology	X	X		X	X	X
Construction	X					
Defense		X			X	X
Emergency Services						X
Energy	X	X	X	X	X	X
Finance	X	X	X	X	X	X
Food	X	X	X	X		X
Government						X
Health	X	X				X
Industry	X	X				
Information Technology	X	X				X
Insurance	X				X	
Legal Order		X			X	
Media		X	X			
Nuclear Industry					X	
Oil and Gas					X	
Postal Services				X		X
Public Order					X	
Research		X				
Space		X			X	
The Internet						X
Transport	X	X	X	X	X	X
Transectoral				X		
Waste	X					
Water	X	X				X

Table 5: Critical sectors of nation-states without definition, Finland to Italy

Table Notes: The list of critical sectors of Finland comes from *The Finnish critical infrastructure protection; state crisis management model and situation awareness* (Härkönen 2007). The list of critical sectors of France comes from *Arrêté du 2 juin 2006 fixant la liste des secteurs d'activités d'importance vitale et désignant les ministres coordonnateurs desdits secteurs* (Government 2006). The lists of critical sectors of Greece and Iceland come from *Protection of 'critical infrastructure' and the role of investment policies relating to national security* (Gordon & Dion 2008). The list of critical sectors of India comes from *International CIIP handbook 2008/2009* (Brunner & Suter 2009). The list of critical sectors of Italy comes from *AIIC - Italian Association of Critical Infrastructures' Experts* (Associazione Italiana esperti in Infrastrutture Critiche 2011).

	Ireland	Kenya	Korea	Latvia	Lithuania	Luxembourg
Administration			X			
Agriculture	X			X	X	
Banking	X			X		
Broadcasting			X			
Communication Technology	X	X	X	X		X
Defense			X	X		
Disaster Control & Management			X			
Emergency Services			X			
Energy	X	X	X	X		X
Finance	X		X	X		
Food	X			X	X	
Government			X			
Health					X	
Information Technology		X				
Media			X		X	X
Nuclear Industry			X			
Oil and Gas			X			
Postal Services	X			X		X
Safety/Security		X	X			
Transport	X	X	X	X	X	X
Water	X					X

Table 6: Critical sectors of nation-states without definition, Ireland to Luxembourg

Table Notes: The lists of critical sectors of Ireland, Latvia, Lithuania, and Luxembourg come from *Protection of 'critical infrastructure' and the role of investment policies relating to national security* (Gordon & Dion 2008). The list of critical sectors of Kenya comes from *The critical infrastructure protection bill, 2015* (Nkaissery 2015). The list of critical sectors of Korea comes from *International CIIP handbook 2008/2009* (Brunner & Suter 2009).

	Mauritius	Malaysia	Mexico	Portugal	Romania	Russia
Agriculture		X	X			
Banking		X	X	X		
Broadcasting	X					
Chemical Industry			X	X	X	
Communication Technology	X	X	X	X		X
Defense		X	X	X	X	X
Disaster Control & Management						X
Domestic & Foreign Policy						X
Emergency Services		X				
Energy	X	X	X			
Finance	X	X	X	X		X
Food		X	X			
Government	X	X				
Health	X	X				
Industry	X					
Information Technology	X	X				X
Legal Order						X
Logistics	X					
Media			X			
Postal Services			X	X		
Safety/Security		X				
Science & Technology						X
Transectoral			X			
Transport	X	X	X	X	X	
Waste			X			
Water	X	X	X	X	X	

Table 7: Critical sectors of nation-states without definition, Mauritius to Russia

Table Notes: The list of critical sectors of Mauritius comes from *National cyber security strategy 2014-2019* (Republic of Mauritius 2016). The list of critical sectors of Malaysia comes from the *CNII Portal* website (Critical National Information Infrastructure [CNNI] 2016). The lists of critical sectors of Mexico, Portugal, and Romania come from *Protection of 'critical infrastructure' and the role of investment policies relating to national security* (Gordon & Dion 2008). The list of critical sectors of Russia comes from *International CIIP handbook 2008/2009* (Brunner & Suter 2009).

	Singapore	Slovak Republic	Slovenia	Sweden
Banking	X			
Communication Technology	X		X	X
Defense			X	
Energy	X		X	X
Finance	X			X
Food	X			
Health	X			
Industry				X
Information Technology	X			
National Command Systems				X
Postal Services			X	
Safety/Security	X			
SCADA Systems				X
The Internet				X
Transport	X	X	X	X
Water	X		X	X

Table 8: Critical sectors of nation-states without definition, Singapore to Sweden

Table Notes: The lists of critical sectors of Singapore and Sweden come from *International CIIP handbook 2008/2009* (Brunner & Suter 2009). The lists of critical sectors of Slovak Republic and Slovenia come from *Protection of 'critical infrastructure' and the role of investment policies relating to national security* (Gordon & Dion 2008).

An Analysis of the Social-Media Technology, Tactics, and Narratives Used to Control Perception in the Propaganda War Over Ukraine

KJ Boyte

Department of Business, English, and Language Arts
Cabrillo College
San Francisco Bay Area, U.S.A.
E-mail: kenboyte@gmail.com

Abstract: *Framed by the historic 2013-2015 conflict in Ukraine—widely described by Western media as a 'Cold-War'-style clash between the Russian Federation and the United States/NATO alliance—this qualitative case study examines how social media was used as a platform for propaganda warfare waged by clandestine bloggers and special operations commandos (known as trolls) deployed worldwide by state and non-state actors, and digitally linked and informed by data-mining, to manipulate public perceptions of the events by controlling an element of rhetoric known as narratives.*

Keywords: *Social Media, Propaganda, Narratives, Information War, Information Operations, Ukraine, NATO, Russian Federation, United States, Trolls, Internet, Cyber War*

Introduction

This qualitative case study focuses on the uses of social media (Kaplan & Haenlein 2010) and an element of propaganda (Pedro 2011; Troianovski 2014; U.S. Department of Defense 2010) known as *narratives* (Defense Advanced Research Projects Agency [DARPA] 2011; Freeman 2012; Lim 2011; Mattis 2009) during the 2013-2015 conflict in Ukraine involving the Russian Federation, NATO, and the United States (Geers 2015; Gertz 2015; Woehrel 2015). The conflict was historic in terms of the extent cyber-warfare tactics supported information operations (Giles 2015; Jaitner 2015) to control the interpretation of political events (Pocheptsov 2014). The study also exemplifies the evolution of military strategies and doctrines of information warfare as a strategic dimension of what Russian General Valery Gerasimov has described as *hybrid warfare* (Bachmann & Gunnneriusson 2015; McDermott 2016; U.S. Army Special Operations Command 2015).

Defined by Kaplan and Haenlein (2010, p. 60) as a "group of Internet-based applications that build on the ideological and technological foundation of Web 2.0, which allows the creation and exchange of user-generated content," social media has become increasingly militarised (Gertz 2015; NATO 2016) in pro-democracy movements throughout Eastern Europe (Hollis 2011), the Middle East (Ketchley 2014), North Africa (Esseghaier 2013), and China (Tkacheva *et al.* 2013). It has been specifically used as a tool for political organisations (Pappic & Noonan 2011) and as a weapon of propaganda warfare (Lange-Jonatmishvili & Suetoka 2015; Murphy & White 2007; Shirky 2011).

Of particular interest to this case study of the 2013-2015 Ukraine conflict—predominantly described by Western media as an end-of-the-world confrontation between the Russian Federation and the United States/NATO alliance (Allam 2014; Limnell 2014; Parry 2014; Rothrock 2014; Snyder 2014; Stern 2014; Walker 2014)—is a form of propaganda known as *narratives*, also called 'semantic frames' (Butler 2010; Fillmore & Baker 2012) and 'cognitive frames' (Lakoff 2009), which are important because the way in which reality is portrayed "actively participates in a strategy of containment, selectively producing and enforcing what will count as reality" (Butler 2010, p. xiii). Explaining the importance of narratives to national security, the current U.S. Secretary of Defense and former U.S. Marine Corps four-star general James Mattis (2009), while serving as Supreme Allied Commander of NATO, stated, "Our military must operate within a Structure Concept that highlights the battle of the narrative… NATO cannot at this point in history surrender any part of the warfighting spectrum". Reiterating this point, DARPA, the research-and-development wing of the U.S. Department of Defense that created the Internet and stealth technology, reported that stories "change the course of insurgencies, frame negotiations, play a role in political radicalization, [and] influence the methods and goals of violent social movements" (Lim 2011). Russian television anchor Dmitry Kiselyov, who also heads a government information agency (Ennis 2014), noted that in the military context, social media is regularly employed to intentionally distort reality: "Previously, there was artillery preparation before an attack. Now, it's informational preparation" (qtd. in Dougherty 2014, p. 2).

In light of these concerns about the uses of social media for propaganda warfare, this study attempts to answer the following three research questions:

(1) How did the Russian Federation, NATO, and the United States use social media to control the perception of political events during the 2013-2015 conflict in Ukraine?
(2) How has the use of social media, as exemplified in the Ukraine conflict, changed the traditional model of propaganda warfare?
(3) What were the dominant narratives used by the main players to frame the Ukraine conflict?

After describing the methodology used to collect and analyse data for this case study, the paper presents a brief history of information warfare and the militarisation of social media, followed by sections focusing on how the Russian Federation used social media as a platform for propaganda during the Ukraine conflict, the subsequent response by NATO and the United States, the dominant narratives and counter-narratives used to shape public perception of the international incident, and the findings of this case study.

Methodology

This non-experimental case study utilised a qualitative method to collect and analyse data, beginning with Internet searches and textbook research to identify and locate mainstream media reports, articles from military journals and academic institutions, as well as reports from NATO and NATO-member nations, which included various U.S. government agencies, such as the Congressional Research Service, the Defense Advanced Research Projects Agency (DARPA), the U.S. Department of Defense, and the U.S. Department of State. To further identify specific counter-narratives—disseminated via social media—which were supportive of the NATO/U.S.

position, the author collected and analysed 3,046 tweets posted by the English-language Twitter account @WeAreUkraine during the 2013-2015 Ukraine conflict, although only a small sampling of those postings are considered herein. The next section of the paper presents a historical overview of information warfare and social media.

Information Warfare and Social Media

Information has been used as a weapon in warfare from time immemorial and today continues to be wielded by governments and non-state actors as an instrument of power (Hardy 2005; Murphy & White 2007; Shirky 2011). Referring to one aspect of information warfare that seems to have been important to both the United States and the Russian Federation since World War II, Simpson (1996, p. 15) reported that

> Psychological warfare is not new of course. It is a modern coalescence and development of very old methods. Some of the earliest human civilizations used symbols, masks, and totems as instruments of power, and the ancient military philosopher Sun Tzu documented the use of relatively sophisticated 'psychological' tactics in both warfare and civil administration as early as the fifth century B.C.E.

The term *propaganda*, conceptually linked to psychological operations, originated in the 17th century, initially meaning "'a society of cardinals' whose task was to oversee and facilitate foreign missions of the Catholic Church" (Pynnoniemi & Racz 2016, p. 27). Although not included in the U.S. Department of Defense's (2016) updated *Dictionary of military and associated terms*, 'propaganda' was earlier defined by the DoD (2010) as "any form of adversary communication, especially of a biased or misleading nature, designed to influence the opinions, emotions, attitudes, or behaviors of any group in order to benefit the sponsor, either directly or indirectly". Alternatively, Roger Vandomme (2010, p. 12) of the Canadian military described 'propaganda' as

> an array of psychological actions, executed by an institution or organization, that determine public perception of events, people, or issues, so as to indoctrinate or recruit a population and to make it act in a certain way.

By comparison, Dmitry Tulchinskly, the Berlin bureau chief of the Russian news agency *Rossiya Segodnya*, defined the term as "the tendentious presentation of facts… It does not mean lying" (Troianovski 2014).

Other similar terms popularly used throughout the literature include *information operations*, *information warfare*, *strategic communications*, *perception management*, and *influence operations* (Garfield 2007, emphasis added). Overlapping to some degree, the semantic boundaries between these terms are imprecise. For example, describing "the military's role in the broadest possible construct" (Martemucci 2007, p. 8), the U.S. Department of Defense (2016, p. 113) defined 'information operations' as "the integrated employment, during military operations, of information-related capabilities in concert with other lines of operations to disrupt, corrupt or usurp the decision-making of adversaries and potential adversaries while

protecting our own". By comparison, the DoD (2016, pp. 223-224) defined 'strategic communications' as:

> Focused United States Government efforts to understand and engage key audiences to create, strengthen, or preserve conditions favorable for the advancement of USG interests, policies, and objectives through the use of coordinated programs, plans, themes, messages, and products synchronized with the actions of all instruments of national power.

Semantically similar to strategic communications but not included in the DoD's (2016) updated *Dictionary of military and associated terms*, 'perception management' was defined as

> Actions to convey and/or deny selected information and indicators to foreign audiences to influence their emotions, motives, and objective reasoning as well as to intelligence systems and leaders at all levels to influence official estimates, ultimately resulting in foreign behaviors and official efforts favorable to the originator's objectives. In various ways, perception management combines truth projection, operations security, cover and deception, and psychological operations. (Martemucci 2007, p. 6)

Perhaps reflecting an evolution of conceptual thinking, or a change in the organisation or classification of designated activities, the terms 'influence operations' and 'psychological operations' also are not included in the DoD's (2016) updated *Dictionary of military and associated terms*.

Attempting to distinguish between the military doctrines of the United States and the Russian Federation, which both consider information to be important in modern warfare, Heickero (2010) reported that Moscow's view is more similar to Chinese thinking, although both the Russian Federation and the United States consider the concept of deception to be a core component of IO/IW. Explaining that some of the perceived differences between U.S. and Russian psychological operations may result from dissimilarities in organisational structures and historical contexts, Heickero (2010) additionally reported that both countries also recognise the need to achieve information superiority to win future wars. In this regard, according to Unwala and Ghori (2015), on 5 February 2010, the Russian Federation updated its military doctrine—which since 2000 had been written in defensive language—with references to the modernisation and development of "forces and resources for information warfare".

Regarding the United States' experience with information operations, propaganda was a central cause of World War I and a driving force behind U.S. entry into that international conflict (Larson 1966). The U.S. government continued to deploy information operations in World War II (Cali & Romanych 2005), when U.S. Army General William 'Wild Bill' Donovan led the Office of Strategic Services, the predecessor of the Central Intelligence Agency, and embraced the German military strategy *Weltanschauungskrieg* ('war of the worldviews') (Pocheptsov 2014). Simpson (1996, p. 24) reported that the "use of the new term [translated into English in 1941 as 'psychological warfare'] quickly became widespread throughout the U.S. intelligence community". By 1948, George Kennan of the U.S. State Department advocated using words as

weapons of political warfare even in times of peace to create confusion and uncertainty about world events similar to what the Prussian military scholar Carl von Clausewitz described in 1873 as 'the fog or war' (Tucker 2015).

However, many of the United States' information operations since World War II—beginning with *Radio Free Europe* (1950) and *Radio Liberty* (1953) targeting audiences in Eastern Europe and the Soviet Union—have been associated with intelligence failures and generally suffered from a lack of cultural knowledge about target populations (Bayles 2014; Cox 2006; Davis 2012; Dizard 2004; Garfield 2007; Munoz 2012; Trent & Doty 2005; Vanden Brook & Locker 2012; Wright 2009). In fact, according to U.S. Ambassador George V. Allen, director of the U.S. Information Agency from 1957 to 1960, "Americans are the worst propagandists" (qtd. in Bayles 2014, p. 136).

In the historical context of such information operations, this case study focuses on the uses of social media, the latest platform of propaganda warfare, to control international public perception of the 2013-2015 conflict in Ukraine, which was part of the Soviet Union until the break-up of that super-power nation in the early 1990s. Keenly aware of the power of propaganda during that global clash of ideologies, the mostly young demonstrators, who gathered in the central area of Kyiv (or Kiev) known as *Maidan Nezalezhnosti* ('Freedom Square'), widely used social media for organising events and waging information warfare throughout the 2013-2015 conflict in Ukraine. This practice began with the political protests that erupted in November 2013 following the government's decision not to pursue closer ties with the European Union, and continued through the overthrow of Russian-leaning President Viktor Yanukovych in February 2014 and the subsequent Russian annexation of Crimea the following month (Geers 2015; Gertz 2015; Woehrel 2015).

Some of the other reported uses of social media for propaganda warfare during the Ukraine conflict included spreading traditional disinformation, rumours, and half-truths via *trolls*—paid and organised bloggers (both military and civilian) who disseminate propaganda and otherwise behave disruptively online to derail conversations in opposition to their own political agendas (Fillingham 2015; Freedom House 2013; Gregory 2014a, 2014b; Khazan 2013; NATO 2016; Szwed 2016). The widespread use of trolls in the information operations reported on in the next sections of this study is similar to other programs in China, Iran, North Korea, and elsewhere (Sindelar 2014; Valentino-Devries & Yadron 2015), described as "rampant" in 22 of the 60 nations examined by the non-governmental organisation Freedom House (2013).

Because the 140-character limit for messages presents an incomplete and distorted view of the world, some analysts consider the social media platform Twitter, founded in March 2006 and headquartered in San Francisco, to be best suited for propaganda warfare (Gross 2011; Naughton 2013; Stern 2014). Based on an analysis of 3.6 million tweets collected with the main hashtag #Euromaidan between 25 November 2013 and 28 February 2014, a preliminary report by the Social Media and Political Participation Lab at New York University found that the number of tweets increased from 120,000 in the first weeks of the protest in Ukraine during late 2013 to about 250,000 in the 24-hour period following the violence of 18 February 2014. During this time, the proportion of tweets in English increased from 44 percent to 53 percent, compared to tweets in Russian and Ukrainian. Although the number of English-language tweets

was "augmented by a much larger population of users from the international community," 69 percent of the tweets originated in Ukraine (Tucker, Metzger & Barbera 2014). In an article for the English-language newspaper *The Kyiv Post*, Kapliuk (2013) reported the frequency of the tweets on 26 November 2013 hashtagged #Euromaindan as occurring "every one or two seconds". The identity of those responsible for these postings remains unclear. The next section of this article focuses on the Russian Federation's use of social media as a platform for propaganda during the 2013-2015 conflict in Ukraine.

Advanced Russian Propaganda Tactics in Information War

As the result of recent doctrinal changes and experiences launching cyber attacks against Estonia in 2007 (*BBC News* 2007) and Georgia in 2008 (Hollis 2011; Pomerleau 2016), the Russian Federation has increasingly used social media to support military offenses in hybrid warfare, "which consists of deliberate disinformation campaigns supported by actions of the intelligence organizations designed to confuse the enemy and achieve strategic advantage at minimal cost" (Snegovaya 2015, p. 9). Describing the 2013-2015 conflict as the "first cognitive war in the world," Pocheptsov (2014) reported that "the seizure of the territory in the physical space was made both at the expense of generated uncertainty and ambiguity, and at the expense of reinterpretation, which were aimed at blocking the counteraction".

As part of its development of technologies to launch and defend against social media attacks, increasingly funded and prioritized since 2011 (Brown 2015), the Russian Federation also has been paying English-speaking Russians to post pro-Russian comments on news reports made by *Fox News*, the *Huffington Post*, and other U.S. news sources (Khazan 2013; Sindelar 2014). Other tactics of Russian trolls have included spamming photo-shopped images of staged events via fake social media accounts "to destabilize entire communities and districts", according to Gregory (2014b), who also explained, "Putin's invisible social media campaign includes fiction writers posting on fake Facebook (or the Russian version *Vkontakte*) accounts, pretending to have witnessed some horrendous crime committed by Ukrainian extremists". Describing Russia's audience-influencing techniques as advanced and involving organised and coordinated disinformation operations on social media and news portals, NATO reported, "A scheme for troll activity can be described in three phases: luring, taking the bait, and hauling in" (Szwed 2016, p. 7).

The Russian Federation's worldwide expanding information operations has also included the 2005 launch of the foreign-language *Russia Today* (*RT*) television network (broadcast in English, Spanish, and Arabic), as well as foreign-language web sites and social media (Troianovski 2014). According to Bidder (2013), *RT* is the most popular foreign broadcaster in Chicago, New York, and San Francisco. In 2014, Miskimmon and O'Loughlin reported that "Russia's response to the West is actually more sophisticated. By being so blatant about their propaganda efforts, they attempt to lead audiences to distrust all politics and media, thereby neutralizing any Western propaganda".

The U.S. and NATO Response

Scrambling for information control, the U.S. State Department now operates more than 350 Twitter accounts (Kastrenakes 2015; Suebsaeng 2014), including the Russian-language @ProgressForUkraine (Rothrock 2014). The U.S. Department of Defense also operates Twitter

accounts for all of its combat commands (Barnes & Yadron 2015). These online initiatives additionally support traditional information operations, including intensified broadcasting efforts in Eastern Europe via *Radio Free Europe/Radio Liberty* and *Voice of America* (Weed 2014; Woehrel 2015).

Based on a 2011 memo from the U.S. Secretary of Defense stressing the importance of the "new war or words" (Cigalese 2013, p. 109), the Defense Advanced Research Projects Agency (DARPA) has also been developing the technology "to take propaganda to the next level, … hoping to do so by controlling the very way their targets perceive and interpret the flow of incoming information" (Dvorsky 2011). In 2011, the DoD agency unveiled the Social Media in Strategic Communications Program to address threats to national security posed by social media and enemy narratives (Sanchez n.d.; U.S. Army Special Operations Command 2014). Explaining some of DARPA's strategies for using social media tactically in information warfare, an agency researcher explained,

> Since everyone is potentially an influencer on social media and is capable of spreading information, our work aims to identify and engage the right people at the right time on social media to help propagate information when needed. (Quinn & Ball 2014)

The U.S. State Department created a similar agency in 2011 to coordinate and strengthen government efforts to use social media (U.S. Department of State n.d.; White House 2011).

Related to the mission of controlling enemy narratives, the U.S. military's Central Command (CENTCOM) has been developing "an online persona management service" that will allow military personnel to each operate worldwide up to ten online personas "to counter violent extremist and enemy propaganda outside the U.S." (Fielding & Cobain 2011). According to testimony before the U.S. Senate's Armed Services Committee in 2010, the purpose of the program is to "disrupt recruitment and training of suicide bombers; deny safe haven for our adversaries; and counter extremist ideology and propaganda" (Fielding & Cobain 2011). In July 2014, NATO also established the Strategic Communications Center of Excellence in Riga, Latvia, as a partnership between Estonia, Germany, Italy, Latvia, Lithuania, Poland, the UK, and the U.S. to "focus on providing an alternative to the official Russian narrative" (NATO 2014).

Because social media is considered an intelligence asset, special operations forces within the U.S. Department of Defense have also been using it for data mining to determine the identities of individuals in the 'sniper scope', as well as their social connections, to better inform soldiers' decisions about whether to pull the trigger (Tucker 2015). Already, technologies are available that "connect every Twitter or Facebook post to a specific location," and the U.S. Department of Defense is developing a technology that can provide a much more detailed, real-time, and geo-spatial analysis of social media postings and the identities of the people associated with related social networks (Gertz 2015). Rawnsley (2011) reported that NATO has been using Twitter "as an intel source to aid in bomb-targeting decisions". According to a 2014 article in *Defense World*, DARPA has also been developing a computer-powered "smart rifle" for snipers that uses lasers to 'tag' targets "much like tagging photographs on Facebook … [t]his data can then be relayed or streamed for immediate intelligence sharing". In view of these technological

advances, the next section of the paper considers the political narratives, manifest by militarised social-media technology and evolving doctrines of warfare, used to shape public perception of the 2013-2015 conflict in Ukraine.

Dominant Narratives of the Ukraine Conflict

In a report for the Finnish Institute of International Affairs that analysed the official statements from the Russian Federation's Ministry of Foreign Affairs, Pynnoniemi and Racz (2016) identified several narratives used by the Russian government to frame the public perception of political events in Ukraine, including the narrative of Nazi atrocities and ethnic cleansing, the narrative of a *coup d'état*, the narrative of Western geopolitical expansion to Russia's sphere of influence, the narrative of the tale of one nation, and the narrative of the Crimea operation as a legitimate action.

All of these narratives are historically situated in conflicts. For example, to trigger memories of Nazi atrocities in Crimea during World War II, some statements from the Russian government used the term *karatel* ('punisher'). On 4 May 2014, the Russian Ministry of Foreign Affairs stated, '"While Ukrainian punishers (karatel) are conducting their operations in Eastern Ukraine carrying out cleansing on some populated areas…, there is an actual information blockade regarding the tragic events in this country in the West'" (qtd. in Pynnoniemi & Racz 2016, pp. 22-23).

Blaming the deaths of about 100 civilians and police who died as a result of the 'anti-terrorist operations' against protesters on 18 and 20 February 2014, the Russian Ministry of Foreign Affairs stated, "We again see that the United States…are in fact attempting to impose a 'Western vector' on their development dictating to the authorities of a sovereign country what they should do" (qtd. in Pynnoniemi & Racz 2016, pp. 79-80).

Similar to this narrative of a *coup d'état*, the narrative of Western geopolitical expansion framed the conflict as a binary between a 'passive' Russia and an 'active' West. For example, citing NATO's expansion into Eastern Europe as causing the conflict in Ukraine, the Russian Ministry of Foreign Affairs stated on 12 March 2014, "This is also evidence of a desire to shift the blame onto Russia and to present it as a geopolitical adversary in order to justify the bloc's existence" (qtd. in Pynnoniemi & Racz 2016, pp. 84-85).

Finally, regarding the narrative of the tale of one nation, on 3 March 2014, the Russian Federation's Ministry of Foreign Affairs described Russia's relationship with Ukraine as historical, calling Ukraine "a fraternal country with which we have many ages of shared history" (qtd. in Pynnoniemi & Racz 2016, pp. 90-91).

Countering the Russian narratives, Western media generally characterised the conflict in Ukraine as a 'Cold-War'-style propaganda war between the Russian Federation and the United States/NATO alliance (Allam 2014; Limnell 2014; Parry 2014; Rothrock 2014; Snyder 2014; Stern 2014; Walker 2014). For example, reporting on the conflict in Ukraine in March 2014, Snyder (2014) wrote, "From Moscow to London to New York, the Ukrainian revolution has been seen through a haze of propaganda". In a report for the U.S. Army publication *Stars and Stripes* that implies the need for strong corrective action, Allam (2014) also stated

> Russia's actions…are outpacing the U.S. responses with bolder and more provocative statements from the Kremlin each day, according to analysts who watch Moscow's messaging. One battle in this information war is waged 140 characters at a time, with the State Department and Russia's Foreign Ministry squaring off on Twitter.

To corroborate these political narratives with those widely circulated via social media, as well as to better understand how political activists used social media during the Ukraine conflict, this author collected and analysed tweets from the English-language Twitter account @WeAreUkraine, which was part of an unpublished study conducted for a graduate-level research course in applied linguistics. Among the findings of the study was that the social media postings reflected international media coverage of the Ukraine conflict, which was generally characterised as a 'Cold-War'-style confrontation between the Russian Federation and the United States/NATO alliance (Allam 2014; Limnell 2014; Parry 2014; Rothrock 2014; Snyder 2014; Stern 2014; Walker 2014). More specifically, these postings can be classified according to five political themes: (1) narrative of a Russian military invasion, (2) narrative of despotism (criticism of President Putin), (3) narrative of propaganda war, (4) narrative of Slavic ancestry and ties to Europe, and (5) narrative of NATO intervention. While perusing the data below, readers should be aware of the non-standard spelling and usage of the English language popular among Twitter users, resulting from the 140-character restrictions of the Twitter format. In some cases (for example, references to President Putin as 'Pu'), the abbreviated language seems to be used for rhetorical effect.

Regarding the first political theme identified, illustrating the narrative of a military invasion, some tweets implied that Russian provocateurs and agents were behind the separatist movement in Ukraine (for instance, 8 April 2014: "Last night I listened to radio traffic of Donetsk separatists on Zello. Some coordinators wt stark RU accent and do not know town geography"). Other postings used satire in describing purported actions of the Russian military (for instance, 17 April 2014: "TV tower 'seized' by separatists in Slovyarisk. Ukrainian TV is off, Russian TV is back on"). In other posts, Russian soldiers were called 'green men' (for instance, 10 March 2014: "Green men in Crimea are being rotated, some fresh aliens are brought in"); 'tourists' (for instance, 7 April 2014: "RU tourists mistook opera house for regional govt building&tried to storm it"); 'bandits' (for instance, 14 March 2014: "Lavrov lied again! RU bandits beat & killed pl in Donetsk); and, 'thugs' (for instance, 6 April 2014: "While RU thugs stormed regional govt in Donetsk today, ppl had a wedding with fireworks in nearby downtown hotel. Life goes on despite Pu").

Another dominant theme of the tweets analysed concerned criticism of Russian President Vladimir Putin, who is generally blamed for the current crisis in Ukraine (for instance, 17 March 2014: "Having started trouble in Crimea, Pu unequivocally proved to all world powers that Russia must be neutralized"). Other tweets also seemed to attribute diabolical motives to the Russian leader (for instance, 9 March 2014: "He doesn't care about 'cold.' The worse the better. If world turns against him, he'll have perfect excuse for N. Korean misery in RU"). Additionally, President Putin was accused of stirring up tensions of the Cold War to divert attention away from the Russian economy (for instance, 23 March 2014: "[Putin] needs external enemy to keep

public attention off internal problems"). Other postings portray him as dishonest (for instance, 14 March 2014: "Putin, Lavrov, Churkin and other RU politicians and media redefined my notions of cynicism, lies and chutzpah with fresh depth of meaning") and trying to reunite the Soviet Union (for instance, 22 March 2014: "PU is hell-bent on reviving USSR").

The theme of propaganda warfare also is dominant in many of the tweets, where terms such as 'brainwashed' and 'zombified' were used to describe Russians and pro-Russian supporters. For example, on 9 March 2014, @WeAreUkraine tweeted, "Feel sorry 4 Russians who r brainwashed by propaganda not to see obvious mismanagement&robbery of their country by Putin&Co". Another posting on 10 March 2014, stated: "Poor Russians. They have been zombified for years by one of the world's most powerful propaganda machines." Similar postings were made on 10 March 2014 (for instance, "Direct psy warfare—new paradigm in military science") and 11 March 2014 (for instance, "Only now did I appreciate what a dark and terrible thing propaganda is. These lies are venom").

Slavic identity is another theme found in the tweets (for example, 22 March 2014: "I prefer 2 be great Ukrainian kozak, rather than Malorussian cossak. My hero is Mazepa, rather than Khmeinitsky"). Although the curators of @WeAreUkraine recognised the common heritage shared by the peoples of Ukraine and the Russian Federation, they also acknowledged that the citizenry of both countries are different, which is emphasized in many tweets (for instance, 19 March 2014: "True, Kievan Rus is our common root. But modern RU doing everything to ruin friendship with other Slavs"). In differentiating Ukraine from the Russian Federation, some posts also emphasized Ukraine's historical heritage with Europe (for instance, 8 March 2014: "…whereas the first constitution of a republic in Europe was written in 1710 by Pylyp Orlyk, a kozak Hetman. Freedom is in our blood"). Other tweets focused on Ukraine's legacy of liberty (for example, 9 March 2014: "We are descendants of Kosaks, self-organization and freedom is our mantra. And God is on our side"). Some postings also expressed aspirations to join the NATO alliance (for instance, 8 March 2014: "Ukraine has never been that unified. NATO here we come").

Related to this theme, as early as 9 March 2014, @WeAreUkraine pleaded for Western intervention (for instance, "There are reports tht RU missile cruiser Moskva is moving toward UA mainland to secure amphibious assault in Herson Reg. US/EU, please act"). Numerous other tweets also called for NATO intervention. For example, on 11 March 2014, @WeAreUkraine posted, "…international community must act"; which was followed on 24 March 2014 by: "NATO must take a tougher position on UA crisis. If east and south UA annexed, RU will get a powerful boost to its military capacity". Similarly, on 17 April 2014, @WeAreUkraine tweeted: "Thrashing of RU at UN Security Council is not enough. The world should recognize RU as dangerous international aggressor and act accordingly".

Discussion

Regarding the 2013-2015 Ukraine conflict involving the Russian Federation and the United States/NATO alliance, defined by the extensive use of social media, this study attempted to answer the following three research questions:

(1) How did the Russian Federation, NATO, and the United States use social media to control the perception of political events during the 2013-2015 conflict in Ukraine?
(2) How has the use of social media exemplified by the Ukraine conflict changed the traditional model of propaganda warfare?
(3) What were the dominant narratives used by the main players to frame the Ukraine conflict?

To answer the first research question, the author examined mainstream media reports, articles from Western military journals and academic institutions, as well as reports from NATO and NATO member nations, which this author acknowledges could also be part of broader disinformation operations intended to skew public perception and academic inquiry. In addition to the reported uses by protesters for organising political events and waging information warfare (Geers 2015; Gertz 2015; Woehrel 2015), the study found that state and non-state actors paid civilian and military bloggers known as trolls to spread propaganda via social media during the 2013-2015 conflict in Ukraine (Fillingham 2015; Freedom House 2013; Gregory 2014a, 2014b; Khazan 2013; NATO 2016; Szwed 2016), although the identities and motives of those responsible for the social media postings examined in this study remain highly questionable.

Reportedly, the main players in the Ukraine conflict also used social media for surveillance purposes (Gertz 2015; Lutz 2014; Pappalardo 2013; Rawnsley 2011) and data-mining (Tucker 2015). Treverton and Miles (2014, p. 20), of the Swedish National Defense College, explained, "If social media and a smart phone 'can turn any human into a geo-located collector', they can also turn any human into an intelligence collection target".

Regarding the second research question (How has the use of social media exemplified by the Ukraine conflict changed the traditional model of propaganda warfare?), the involvement of suspected military and civilian bloggers known as trolls (Fillingham 2015; Freedom House 2013; Gregory 2014a, 2014b; Khazan 2013; NATO 2016; Szwed 2016) seems to have blurred the distinction between civilians and military combatants (Treverton & Miles 2014), thus expanding the battlespace to "include the civil, commercial, and private infrastructure of a nation by targeting mass beliefs and perceptions" (Molinari 2005, p. 20).

This study also found that Web 2.0 technology seems to be contributing to the development of a new, two-way model of propaganda (Pedro 2011), where "elite propagandizing interacts with target audiences who play an active role in the production of meaning" (Dauber & Winkler 2014, p. 7). Similarly, explaining that the purpose of propaganda is to incite anger and fear, Treverton and Miles (2014) reported that the use of social media is changing the way conflicts are mediated—allowing enemy combatants to broadcast duelling narratives, attempting to legitimate their causes, at lightning speed, although much of their messaging may be relegated as unbelievable because of the suspected involvement of trolls (Chen 2015).

In terms of the third research question (What were the dominant narratives used by the main players to frame the Ukraine conflict?), this study additionally found that Western sources generally portrayed the international confrontation over Ukraine as a 'Cold-War'-style propaganda war between the Russian Federation and the NATO/U.S. alliance. In this regard, despite the concerns of U.S. and NATO officials about losing the information war to the

Russian Federation (Allam 2014; Powell 2014; Rothrock 2014), it seems that the West's technologically-driven strategies to counter the narratives of the Russian Federation were quite effective, at least in terms of the widespread dissemination throughout Western media and academia of the narrative of a propaganda war. Accordingly, in the same way Miskimmon and O'Loughlin (2014) reported that Russia's advanced propaganda efforts "attempt to lead audiences to distrust all politics and media, thereby neutralizing any Western propaganda," the most damning allegations against the West made by the Russian Federation seem to have been marginalised as 'Cold-War' rhetoric, at least among English-speaking audiences seeking information about the ordeal from official channels.

This study only considered English-language sources and social-media postings. The characterisation of the Ukraine conflict by foreign-language sources could be different, a likely possibility that should be investigated by those linguistically capable. However, just because the Western narrative of a propaganda war was dominant does not mean that the consumers of mainstream media reports and academic sources believed the narrative; the dominance may have more to do with the monopolisation of media ownership (Bagdikian 2004) than any inherent rhetorical qualities of the media message. Moreover, considering the difficulty of proving causality, the actual impact of the Western messaging on the 'hearts and minds' of target audiences is unclear beyond simply crowding out alternative perspectives. In this regard, more problematic is that the dominance of the Western narrative of a propaganda war may have misdirected attention away from important actors who benefited politically and/or economically from the Ukraine conflict, such as the 'corrupt' Russian oligarchs who took control of Ukraine and the country's assets following the break-up of the Soviet Union (Woehrel 2015).

Among related questions that need further investigation are how paid civilian and military bloggers, and other foreign governments and non-state actors more concealed, continue to dominate social media internationally, create false consensuses, and agitate citizens into taking political action. For example, was the @WeAreUkraine Twitter account examined in this study really operated by citizen activists in Ukraine, as the site claims, or was it merely a front for military bloggers? Certainly, in a preliminary report for the Social Media and Political Participation Lab at New York University, Tucker, Metzger & Barbera (2014) found that 31 percent of the tweets hashtagged #Euromaidan came from outside of Ukraine during the 24-hour period following the political violence in Kyiv on 18 February 2014. Who was sending these tweets? What was their motivation?

Additionally, regarding the electronic bombardment of potential disinformation during the height of the 'Freedom Square' protests, how did the ringing, beeping, and vibrating of computers, cell phones, and other mobile media devices in the pockets, purses, and backpacks of the young activists impact their physical actions and reactions to the seemingly never-ending stream of revolutionary messaging? Can any linkages be established between specific messaging and the behaviour of protesters? In an article for *Stratfor Intelligence*, Pappic and Noonan (2011) reported that "At the end of the day, for a social media-driven protest movement to be successful, it has to translate social media membership into street action". Several studies, including Gallagher (2014) and Zhang *et al.* (2015), appear to suggest that research has been underway seeking to make causal links between social media messaging and human behavioural responses possible. Already, Internet platforms have been used to stifle political expression and

support conformity (Newman 2015). Similarly alarming is the fact that other sources, such as Makinen (2016), report that adversaries have used social media to manipulate U.S. elections.

There is a legitimate need for national security and reasonable methods for providing national defence, as government leaders attempt to understand, counter, and control enemy narratives in the future, especially considering the numerous failures of U.S. information operations since World War II (Bayles 2014; Cox 2006; Davis 2012; Dizard 2004; Garfield 2007; Munoz 2012; Trent & Doty 2005; Vanden Brook & Locker 2012; Wright 2009). However, national defences against information warfare should not depend exclusively on high-tech gadgetry and algorithms (DARPA 2011), particularly because of the limited capacity of human beings to process information (Bracken & Shubik 2001) and because of the vulnerabilities that technologies create. That is to say, the World Wide Web, which the United States reportedly began developing in the 1960s to respond to a Soviet nuclear strike (Gervaise 2012), has certainly made the citizenry of the world more vulnerable, in ways that the government scientists who created the Internet allegedly did not anticipate.

Developing technologies for information warfare is important, but it is not all that should be done. In addition to the ongoing need for experienced human intelligence agents on the ground in this advanced age of machinery (U.S. Department of Defense 2014), governments also must educate citizenry to think critically and refrain from overreacting to rumours of wars and catastrophes. In terms of connecting the dots and connecting with the target audiences of foreign nations, the United States also must overcome a history of failed information operations that have suffered from a lack of cultural knowledge about target populations (Bayles 2014; Cox 2006; Davis 2012; Dizard 2004; Garfield 2007; Munoz 2012; Trent & Doty 2005; Vanden Brook & Locker 2012; Wright 2009). In this regard, this author has supported a team at the U.S. Army's Defense Language Institute in developing products to provide cultural training about the peoples and nations of the world for service members deployed internationally.

Conclusion

With the development of technology—including radio, television, and the Internet—the impact of propaganda warfare is significantly greater today than in the past (Larson 1966). According to Mark Laity, the Chief of Strategic Communications at NATO's Supreme Headquarters Allied Powers Europe,

> The threat of conventional warfare has changed and we have to recognize that information can be a weapon. Whether used for disinformation, deception, or plain fabrication to create false narratives, we have to be aware and be able to respond to this challenge. (NATO 2015)

This qualitative case study has provided an open-source overview of how social media and an element of propaganda known as narratives were used during the 2013-2015 information war over Ukraine, when advanced technologies made it possible for governments and non-state actors to leverage the Internet and social media for manipulating public perception of political events. The study also exemplifies the evolution of military tactics and doctrines of information warfare. Regardless of the true identities of Internet trolls or the extent of their involvement during the Ukraine crisis, it is evident that Web 2.0 technology has allowed users of social

media to participate in the production, distribution, and perpetuation of a new generation of propaganda warfare characterised by a localised, bottom-up, and interactive model (Dauber & Winkler 2014; Pedro 2011).

In view of concerns expressed during the Ukraine conflict by U.S. and NATO policymakers worried about losing the information war (Allam 2014; Powell 2014; Rothrock 2014), this case study of the uses of social media and propaganda is important because words often make the difference between peace and war (Larson 1966). The use of such politically-charged language in the information war over Ukraine, according to the dominant narratives of Western media, hearkened back to the height of the Cold War and included calls for using nuclear weapons against targets in the United States and the Russian Federation (Blank 2015; Diamond & Botelho 2015; Keck 2015; Tan 2015). Future research might consider whether the people of the world are really made safer by governments who perpetually wage information warfare, even in times of peace, to create fear, anger and confusion. In addition, researchers might investigate to what extent social media can be used to bring about peaceful rather than doomsday ends?

References

Allam, H 2014, 'Showdown over Ukraine sparks cold war-style propaganda', *Stars and Stripes*, 3 May, viewed 16 June 2016, <http://www.stripes.com/news/europe/showdown-over-ukraine-sparks-cold-war-style-propaganda-battle-1.281181>.

Bachmann, SD & Gunneriusson, H 2015, 'Russia's hybrid warfare in the east: the integral nature of the information sphere', *Georgetown Journal of International Affairs*, pp. 198-211, viewed 22 January. 2017, <https://www.researchgate.net/publication/277953401_RUSSIA'S_HYBRID_WARFARE_IN_THE_EAST_USING_THE_INFORMATION_SPHERE_AS_INTEGRAL_TO_HYBRID_WARFARE>.

Bagdikian, B 2004, *The new media monopoly*, 20th ed., Beacon Press, Boston, MA, U.S.A.

Barnes, JE & Yadron, D 2015, 'U.S. probes hacking of military Twitter accounts by pro-Islamic State group', *The Wall Street Journal*, 12 January, viewed 17 June 2016, <http://www.wsj.com/articles/u-s-investigating-apparent-hack-of-military-twitter-account-by-islamic-militants-supporters-1421086712>.

Bayles, M 2014, *Through a screen darkly: popular culture, public diplomacy, and America's image abroad*, Yale University Press, New Haven, CT, U.S.A.

BBC News 2007, 'Estonia hit by Moscow cyber war', 17 May, viewed 18 February 2017, <http://news.bbc.co.uk/2/hi/europe/6665145.stm>.

Bidder, B 2013, 'Russia today: Putin's weapon in the war of images', *Spiegel International*, trans. C Sultan, 13 August, viewed 4 July 2016, <http://www.spiegel.de/international/business/putin-fights-war-of-images-and-propaganda-with-russia-today-channel-a-916162.html>.

Blank, S 2015, 'Putin's thinly veiled threat of nuclear attack is working', *Newsweek*, 5 June, viewed 16 June 2016, <http://www.newsweek.com/putins-thinly-veiled-threat-nuclear-war-working-339817>.

Bracken, P & Shubik, M 2001, 'Gaming in the information age: theory and purpose', *Naval War College Review*, vol. LIV, no. 2, pp. 47-60, viewed 3 July 2016, <https://www.usnwc.edu/getattachment/db8161f6-f600-4f49-ac05-efc56f8d0876/War-Gaming-in-the-Information-Age--Theory-and-Purp>.

Brown, J 2015, 'Why are Russian trolls spreading hoaxes in U.S?', Public Broadcasting Service *Newshour*, 8 June, viewed 16 June 2016, <http://www.pbs.org/newshour/bb/russian-trolls-spreading-online-hoaxes-u-s/>.

Butler, J 2010, *Frames of war: when is life grievable?* Verso, London, UK, viewed 17 June 2016, <http://humanities.wisc.edu/assets/misc/Butler.pdf>.

Cali, R & Romanych, M 2005, 'Counterpropaganda: an important capability for joint forces', *Sphere*, pp. 11-3, U.S. Strategic Command Joint Information Operations Center, Lackland Air Force Base, TX, U.S.A., viewed 22 January 2017, <http://www.au.af.mil/info-ops/iosphere/iosphere_fall05_cali.pdf>.

Chen, A 2015, 'The agency', *New York Times Magazine*, 2 June, viewed 16 June 2016, <http://www.nytimes.com/2015/06/07/magazine/the-agency.html?_r=0>.

Cigalese, C 2013, 'Redefining information operations,' *Joint Forces Quarterly*, vol. 69, no. 2, pp. 109-12, National Defense University, Fort McNair, Washington, D.C., U.S.A., viewed 8 July 2016, <http://ndupress.ndu.edu/Portals/68/Documents/jfq/jfq-69/JFQ-69_109-112_Cicalese.pdf>.

Cox, J L 2006, 'Information operations in Operation Enduring Freedom and Iraqi Freedom: what went wrong?', U.S. Army Command and General Staff College, School of Advanced Military Studies, Fort Leavenworth, KS, U.S.A., viewed 22 January 2017, <https://fas.org/irp/eprint/cox.pdf>.

Dauber, CE & Winkler, CK 2014, 'Radical visual propaganda in the online environment: an introduction', *Visual propaganda and extremism in the online environment*, eds. C E Dauber & C K Winkler, Strategic Studies Institute, U.S. Army War College, Carlisle Barracks, PA, U.S.A., pp. 1-30, viewed 20 Feb. 2017, <http://docplayer.net/3752312-Visual-propaganda-and-extremism-in-the-online-environment-editors-carol-k-winkler-cori-e-dauber-press-u-s-army-war-college.html>.

Davis, D 2012, 'Dereliction of duty II: senior military leaders' loss of integrity around Afghan war effort', personal blog, 6 February, viewed 22 Jan. 2017, <http://s3.documentcloud.org/documents/291793/dereliction-of-duty-ii-january-15-2012.pdf>.

Defense Advanced Research Projects Agency (DARPA) 2011, 'Broad agency announcement: narrative networks', 7 October, viewed 3 July 2016, <http://www.robo-hunter.com/uploads/files/560c1c5fe87df.pdf>.

Defense World 2014, 'Future assault rifle to tag targets much like social media photo tagging,'16 January, viewed 17 June 2016, <http://www.defenseworld.net/news/9843/Future_ Assault_ Rifles_To_Tag_Targets_Much_Like_Social_Media_Photo_Tagging#.V2QcnelCgb0>.

Diamond, J & Botelho, G 2015, 'U.S.-Russia military tit for tat raises fears of greater conflict', *CNN*, 19 June, viewed 16 June 2016, <http://www.cnn.com/2015/06/17/politics/russia-us-military-threats-rise-ukraine/>.

Dizard, WP 2004, *Inventing public diplomacy: the story of the U.S. information agency*, Lynne Rienner Publishers, Boulder, CO, U.S.A.

Dougherty, J 2014, 'Everyone lies: the Ukraine conflict and Russia's media transformation', Harvard Kennedy School, Shorenstein Center of Media, Politics, and Public Policy, U.S.A., July, viewed 16 June 2016, <http://shorensteincenter.org/everyone-lies-ukraine-conflict-russias-media-transformation/>.

Dvorsky, G 2011, 'Propaganda 2.0 and the rise of narrative networks', *Institute for Ethics and Emerging Technologies*, U.S.A., 19 October, viewed 17 June 2016, <http://ieet.org/index.php/IEET/print/4934>.

Ennis, S 2014, 'Dmitry Kiselyov: Russia's chief spin doctor', *BBC News*, 2 April, viewed 19 June 2016, <http://www.bbc.com/news/world-europe-26839216>.

Esseghaier, M 2013, 'Tweeting out a tyrant: social media and the Tunisia revolution', *Journal of Mobile Media*, vol. 1, no. 1, March, viewed 2 July 2016, <http://wi.mobilities.ca/tweeting-out-a-tyrant-social-media-and-the-tunisian-revolution/>.

Fielding, N & Cobain, I 2011, 'Revealed: U.S. spy operation that manipulates social media', *The Guardian*, 17 March, viewed 16 June 2016, <https://www.theguardian.com/technology/2011/mar/17/us-spy-operation-social-networks>.

Fillingham, Z 2015, 'Internet trolls: propaganda's final frontier', *Geopolitical Monitor*, 16 March, viewed 16 June 2016, <https://www.geopoliticalmonitor.com/internet-trolls-the-final-frontier-of-propaganda/>.

Fillmore, CF & Baker, C 2012, 'A frames approach to semantic analysis', *The Oxford handbook of linguistic analysis*, eds. B Heine & H Narrog, Oxford University Press, Oxford, UK, viewed 16 June 2016, <http://lingo.stanford.edu/sag/papers/Fillmore-Baker-2011.pdf>.

Freedom House 2013, 'Freedom on the Net 2013: a global assessment of Internet and digital media', Washington, D.C., U.S.A., 3 October, viewed 2 July 2016, <https://freedomhouse.org/sites/default/files/resources/FOTN%202013_Full%20Report_0.pdf>.

Freeman, CW, Jr. 2012, 'Hashbara and the control of narrative as an element of strategy', remarks to the Jubilee Conference on the Council of Foreign and Defense Policy, Moscow, Russian Federation, 1 December, viewed 16 June 2016, <http://chasfreeman.net/hasbara-and-the-control-of-narrative-as-an-element-of-strategy/>.

Gallagher, S 2014, 'Air Force research: how to use social media to control people like drones', *ARS Technica*, Cambridge, MA, U.S.A., 17 July, viewed 2 July 2016, <http://arstechnica.com/information-technology/2014/07/air-force-research-how-to-use-social-media-to-control-people-like-drones/>.

Garfield, A 2007, 'The U.S. counter-propaganda failure in Iraq', *Middle East Quarterly*, vol. 14, no. 4, pp. 23-32, viewed 29 June 2016, <http://www.meforum.org/1753/the-us-counter-propaganda-failure-in-iraq>.

Geers, K, ed., 2015, *Cyber war in perspective: Russian aggression against Ukraine*, NATO Cooperative Cyber Defense Center of Excellence, Tallinn, Estonia, viewed 19 January 2017, <https://ccdcoe.org/sites/default/files/multimedia/pdf/CyberWarinPerspective_full_book.pdf>.

Gertz, B 2015, 'Special ops targets social media', *The Washington Times*, 18 March, viewed 17 June 2016, <http://www.washingtontimes.com/news/2015/mar/18/inside-the-ring-special-ops-targets-social-media/>.

Gervaise, M 2012, 'Cyber attacks and the law of war', *Journal of Law and Cyber Warfare*, vol. 1, no. 1, pp. 525-79, viewed 24 June 2016, <http://www.jlcw.org/wp-content/uploads/2013/04/2012-JLCW-Winter-Vol_1_1.pdf>.

Giles, K 2015, 'Russia and its neighbors: old attitudes, new capabilities', *Cyber war in perspective: Russian aggression against Ukraine*, ed. K Geers, NATO Cooperative Cyber Defense Center of Excellence, Tallinn, Estonia, pp. 19-28, viewed 19 January 2017, <https://ccdcoe.org/sites/default/files/multimedia/pdf/CyberWarinPerspective_full_book.pdf>.

Gregory, PR 2014a, 'Inside Putin's campaign of social media trolling and faked Ukrainian crimes', *Forbes*, 10 May, viewed 16 June 2016, <http://www.forbes.com/sites/paulroderickgregory/2014/05/11/inside-putins-campaign-of-social-media-trolling-and-faked-ukrainian-crimes/#260ad513629d>.

——2014b, 'Putin's new weapon in the Ukraine propaganda war: Internet trolls', *Forbes*, 9 December, viewed 16 June 2016, <http://www.forbes.com/sites/paulroderickgregory/2014/12/09/putins-new-weapon-in-the-ukraine-propaganda-war-internet-trolls/#5a57b21559e5>.

Gross, D 2011, '5 ways Twitter changed the way we communicate', *CNN News*, 12 March, viewed 22 January 2017, <http://www.cnn.com/2011/TECH/social.media/03/21/twitter.birthday.communication/>.

Hardy, CK 2005, 'Information operations as an element of national power: a practitioner's perspective on why the United States can't get it right', master's thesis, U.S. Army War College, Strategic Studies Institute, Carlisle Barracks, PA, U.S.A., viewed 22 Jan. 2017, <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1000.1570&rep=rep1&type=pdf>.

Heickero, R 2010, 'Emerging cyber threats and Russian views on information warfare and information operations', FOI, Swedish Defense Research Agency, Stockholm, Sweden, viewed 14 January 2017, <http://www.highseclabs.com/data/foir2970.pdf>.

Hollis, DC 2011, 'Cyberwar case study: Georgia 2008', *Small War Journal*, 6 January, pp. 1-9, viewed 17 June 2016, <http://smallwarsjournal.com/jrnl/art/cyberwar-case-study-georgia-2008>.

Jaitner, ML 2015, 'Russian information warfare: lessons from Ukraine', *Cyber war in perspective: Russian aggression against Ukraine*, ed. K Geers, NATO Cooperative Cyber Defense Center of Excellence, Tallinn, Estonia, pp. 87-94, viewed 19 January 2017, <https://ccdcoe.org/sites/default/files/multimedia/pdf/CyberWarinPerspective_full_book.pdf>.

Kaplan, AM & Haenlein, M 2010, 'Users of the world unite! The challenges and opportunities of social media,*' Business Insider*, vol. 53, pp. 59-68, viewed 18 February 2017, <http://www.sciencedirect.com/science/article/pii/S0007681309001232>.

Kapliuk, K 2013, 'Role of social media in EuroMaidan movement essential', *Kyiv Post*, 1 December, viewed 16 June 2016, <http://www.kyivpost.com/article/content/ukraine/role-of-social-media-in-euromaidan-movement-essential-332749.html>.

Kastrenakes, J 2015, 'U.S. government wants to get better at social media to fight ISIS propaganda', *The Verge*, 17 February, viewed 20 June 2016, <http://www.theverge.com/2015/2/17/8051081/isis-propaganda-us-government-social-media-tactics>.

Keck, Z 2015, 'Showdown: U.S. slams Russia over nuclear war threats', *The National Interest*, 27 April, viewed 19 June 2016, <http://nationalinterest.org/blog/the-buzz/showdown-us-slams-russia-over-nuclear-war-threats-12737>.

Kennan, GF 1948, 'On organizing political warfare', Wilson Center Digital Archive, History and Public Policy Program [U.S.A.], 30 April, viewed 22 January 2017, <http://digitalarchive.wilsoncenter.org/document/114320>.

Ketchley, N 2014, 'How social media spreads protest tactics from Ukraine to Egypt', *The Washington Post*, 14 February, viewed 20 June 2016, <https://www.washingtonpost.com/news/monkey-cage/wp/2014/02/14/how-social-media-spreads-protest-tactics-from-ukraine-to-egypt/>.

Khazan, O 2013, 'Russia's online-comment propaganda army', *The Atlantic*, 9 October, viewed 16 June 2016, <http://www.theatlantic.com/international/archive/2013/10/russias-online-comment-propaganda-army/280432/>.

Lakoff, G 2009, *The political mind: a cognitive scientist's guide to your brain and its politics*, Viking, New York, NY, U.S.A.

Lange-Jonatmishvili, L & Suetoka, S 2015, 'Strategic communications and social media in the Russia Ukraine conflict,' *Cyber war in perspective: Russian aggression against Ukraine*, ed. K Geers, NATO Cooperative Cyber Defense Center of Excellence, Tallinn, Estonia, pp. 103-11, viewed 19 June 2016, <https://ccdcoe.org/sites/default/files/multimedia/pdf/CyberWarinPerspective_Lange_Svetoka_12.pdf>

Larson, A 1966, 'The present status of propaganda in international law', *Law and Contemporary Problems*, pp. 439-51, Duke Law School, Durham, NC, U.S.A., viewed 16 June 2016, <http://scholarship.law.duke.edu/lcp/vol31/iss3/2/>.

Lim, D 2011, 'DARPA wants to master the science of propaganda', *Wired*, 18 October, viewed 16 June 2016, <https://www.wired.com/2011/10/darpa-science-propaganda/>.

Limnell, J 2014, 'Ukraine crisis proves cyber conflict is reality of modern warfare', *The Telegraph*, 19 April, viewed 22 January 2017, <http://www.telegraph.co.uk/technology/internet-security/10770275/Ukraine-crisis-proves-cyber-conflict-is-a-reality-of-modern-warfare.html>.

Lutz, C 2014, 'Is social media a dangerous force against democracy?', The Aspen Institute, Santa Barbara, CA, U.S.A., 6 August, viewed 24 June 2016, <http://www.economist.com/node/17848401>.

Makinen, J 2016, 'Chinese social media platform plays a role in U.S. rallies for NYPD officer', *The Los Angeles Times*, 24 February, viewed 1 July 2016, <http://www.latimes.com/world/asia/la-fg-china-liang-protests-20160224-story.html>.

Martemucci, G 2007, 'Regaining the high ground: the challenges of perception management in national strategy and military operations', master's thesis, U.S. Joint Forces Staff College, Joint Advanced Warfighting School, Norfolk, VA, U.S.A., 17 June, viewed 19 January 2017, <http://www.dtic.mil/dtic/tr/fulltext/u2/a468873.pdf>.

Mattis, J 2009, 'Launching NATO new strategic concept,' NATO, 7 July, viewed 16 June 2016, <http://www.nato.int/cps/en/natolive/opinions_56392.htm>.

McDermott, R 2016, 'Learning from today's wars: does Russia have a Gerasimov doctrine?', *Parameters*, vol. 46, no. 1, pp. 97-105, Strategic Studies Institute, U.S. Army War College, Carlisle Barracks, PA, U.S.A., viewed 22 January 2017, <http://strategicstudiesinstitute.army.mil/pubs/parameters/issues/Spring_2016/12_McDermott-pdf>.

Miskimmon, A & O'Loughlin, B 2014, 'Weaponizing information: Putin, the West and competing strategic narratives of Ukraine', *European Geostrategy* blog, vol. 6, no. 99, viewed 7 July 2016, <http://www.europeangeostrategy.org/2014/12/weaponising-information-putin-west-competing-strategic-narratives-ukraine/>.

Molinari, R J 2005, 'Winning the minds in hearts and minds: a systematic approach to information operations as part of counterinsurgency warfare', School of Advanced Military Studies, U.S. Army Command and General Staff College, Fort Leavenworth, KS, U.S.A., 26 May, viewed 16 June 2016, <http://www.au.af.mil/au/awc/awcgate/sam/winning_minds_molinari.pdf>.

Munoz, A 2012, 'U.S. military information operations in Afghanistan: effectiveness of psychological operations 2001-2010', National Defense Research Institute (RAND), Santa Monica, CA, U.S.A., viewed 17 June 2016, <http://www.rand.org/content/dam/rand/pubs/monographs/2012/RAND_MG1060.pdf>.

Murphy, DM & White, JE 2007, 'Propaganda: can a word decide a war?', *Parameters*, fall, pp. 15-27, Strategic Studies Institute, U.S. Army War College, Carlisle Barracks, PA, U.S.A., viewed 17 June 2016, <http://strategicstudiesinstitute.army.mil/pubs/parameters/articles/07autumn/murphy.pdf>.

NATO 2014, 'Seven allies establish NATO's Strategic Communications Center of Excellence in Latvia', 7 July, viewed 17 June 2016, <http://www.atlanticcouncil.org/blogs/natosource/seven-allies-establish-nato-s-strategic-communications-center-of-excellence-in-latvia>.

——2015, 'Press release: NATO adopts new technology to understand counter Russian propaganda', 11 May, viewed 17 June 2016, <https://toinformistoinfluence.com/2015/05/11/press-release-nato-adopts-new-technology-to-understand-and-counter-rtvussian-propaganda/>.

——2016, 'Internet trolling as a tool of hybrid warfare: the case of Latvia', Strategic Communications Center of Excellence: Riga, Latvia, 25 January, viewed 19 June 2016, <http://www.stratcomcoe.org/internet-trolling-hybrid-warfare-tool-case-latvia-0>.

Naughton, J 2013, 'Twitter and the transformation of democracy', *The Guardian*, 14 September, viewed 21 February 2017, <https://www.theguardian.com/commentisfree/2013/sep/14/twitter-flotation-facebook-politics-social-network>.

Newman, A 2015, 'Government uses psy-ops, trolls, propaganda to push conformity", *The New American*, 8 July, viewed 30 June 2016, <http://www.thenewamerican.com/world-news/europe/item/21215-government-uses-psy-ops-trolls-propaganda-to-push-conformity>.

Pappalardo, J 2013, 'NSA data mining: how it works', *Popular Mechanics*, 11 September, viewed 7 July 2016, <http://www.popularmechanics.com/military/a9465/nsa-data-mining-how-it-works-15910146/>.

Pappic, M & Noonan, S 2011, 'Social media as a tool for protest', Stratfor Intelligence, Austin, TX, U.S.A., 3 February, viewed 1 July 2016, <https://www.stratfor.com/weekly/20110202-social-media-tool-protest>.

Parry, R 2014, 'Ukraine, through the U.S. looking glass—anti-Russian propaganda in the mainstream media', Global Research, Montreal, Canada, 18 April, viewed 22 January 2017, <http://www.globalresearch.ca/ukraine-through-the-us-looking-glass-anti-russian-propaganda-in-the-mainstream-media/5378303>.

Pedro, J 2011, 'The propaganda model in the early 21st century: part II,' *International Journal of Communication*, vol. 5, pp. 1906-1926, viewed 16 June 2016, <http://eprints.sim.ucm.es/24140/1/Pedro%20(2011b).pdf>.

Pocheptsov, G 2014, 'The first cognitive war in the world (Ukraine, Crimea, Russia)', viewed 4 July 2016, <http://www.academia.edu/10057232/FIRST_COGNITIVE_WAR>.

Pomerleau, M 2016, 'Cyber operations come out of the shadows,*' Defense Systems*, viewed 12 July 2016, <https://defensesystems.com/articles/2016/05/05/us-cyber-war-isis.aspx>.

Powell, A 2014, 'Is the U.S. losing the propaganda war with Russia?', University of Southern California Annenberg Center on Communications, Leadership, & Policy, Los Angeles, CA, U.S.A., 2 June, viewed 22 January 2017, <https://communicationleadership.usc.edu/news/is-the-us-losing-the-propaganda-war-with-russia/>.

Pynnoniemi, K & Racz, A, eds. 2016, 'Fog of falsehood: Russian strategy of deception and the conflict in Ukraine', Finnish Institute of International Affairs, Helsinki, Finland, viewed 17 January 2017, <www.fiia.fi/assets/publications/FIIAReport45_FogOfFalsehood.pdf>.

Quinn, B & Ball, J 2014. 'U.S. military studied how to influence Twitter users in DARPA-funded research', *The Guardian*, 8 July, viewed 17 June 2016, <https://www.theguardian.com/world/2014/jul/08/darpa-social-networks-research-twitter-influence-studies>.

Rawnsley, A 2011, 'Pentagon wants a social media propaganda machine', *Wired*, 15 July, viewed 17 June 2016, <https://www.wired.com/2011/07/darpa-wants-social-media-sensor-for-propaganda-ops/>.

Rothrock, K 2014, 'The U.S. is being routed by Russia in the Information War over Ukraine', *The New Republic*, 16 April, viewed 16 June 2016, <https://newrepublic.com/article/117394/us-losing-russia-information-war-over-ukraine>.

Sanchez, S n.d., 'Narrative networks,' Defense Advanced Research Projects Agency (DARPA), Arlington, VA, U.S.A., viewed 3 July 2016, <http://www.darpa.mil/program/narrative-networks>.

Shirky, C 2011, 'The political power of social media: technology, the public sphere, and political change', *Foreign Affairs*, January/February, viewed 21 June 2016, <https://www.foreignaffairs.com/articles/2010-12-20/political-power-social-media>.

Simpson, C 1996, *Science of coercion: communication research and psychological warfare 1945-1960*, Oxford University Press, Oxford, U.K., viewed 22 January 2017, <https://historicalunderbelly.files.wordpress.com/2012/12/0195102924.pdf>.

Sindelar, D 2014, 'The Kremlin's troll army', *The Atlantic*, 12 August, viewed 16 June 2016, <http://www.theatlantic.com/international/archive/2014/08/the-kremlins-troll-army/375932/>.

Snegovaya, M 2015, 'Russia report: Putin information warfare in Ukraine: Soviet origins of Russian hybrid warfare', Institute for the Study of War, Washington, D.C., U.S.A., viewed 7 July 2016, <http://understandingwar.org/sites/default/files/Russian%20Report%201%20Putin's%20Information%20Warfare%20in%20Ukraine-%20Soviet%20Origins%20of%20Russias%20Hybrid%20Warfare.pdf>.

Snyder, T 2014, 'Ukraine: the haze of propaganda', *The New York Review of Books*, 1 March, viewed 19 June 2016, <http://www.nybooks.com/daily/2014/03/01/ukraine-haze-propaganda/>.

Stern, D 2014, 'The Twitter war: social media's role in Ukraine unrest: social media networks are powerful propaganda tools in the Ukraine crisis', *National Geographic*, 10 May, viewed 16 June 2016, <http://news.nationalgeographic.com/news/2014/05/140510-ukraine-odessa-russia-kiev-twitter-world/>.

Suebsaeng, A 2014, 'The State Department is actively trolling terrorists on Twitter', *Mother Jones*, 5 March, viewed 20 June 2016, <http://www.motherjones.com/politics/2014/02/state-department-cscc-troll-terrorists-twitter-think-again-turn-away>.

Szwed, R 2016, 'Framing of the Ukraine-Russia conflict in online and social media', NATO Strategic Communications Center of Excellence, Riga, Latvia, viewed 3 July 2016, <www.stratcomcoe.org/download/file/fid/6141>.

Tan, SL 2015, 'Russian analyst urges nuclear attack on Yellowstone National Park and San Andreas fault line', *Sydney Morning Herald*, 31 March, viewed 19 June 2016, <http://www.smh.com.au/world/russian-analyst-urges-nuclear-attack-on-yellowstone-national-park-and-san-andreas-fault-line-20150330-1mbl14.html>.

Tkacheva, O, Schwartz, LH, Libicki, MC, Taylor, JE, Martini, J & Baxter, C 2013, *Internet freedom and political space*, National Defense Research Institute (RAND), Santa Monica, CA, U.S.A., viewed 22 June 2016, <http://www.rand.org/content/dam/rand/pubs/research_reports/RR200/RR295/RAND_RR295.pdf>.

Trent, S & Doty, JL 2005, 'Marketing: an overlooked aspect of information operations', *Military Review*, vol. 85, no. 4, pp. 70-74, U.S. Army Combined Arms Center, Fort Leavenworth, KS, U.S.A., viewed 18 February 2017, <http://www.au.af.mil/au/awc/awcgate/milreview/trent.pdf>.

Treverton, G T & Miles, R 2014, *Social media and intelligence*, Swedish National Defense College, Stockholm, Sweden, viewed 16 June 2016, <https://www.fhs.se/documents/Externwebben/forskning/centrumbildningar/CATS/publikationer/Social%20Media%20and%20Intelligence.pdf>.

Troianovski, A 2014, 'Russia ramps up information war in Europe', *The Wall Street Journal*, 21 August, viewed 16 June 2016, <http://www.wsj.com/articles/russia-ramps-up-information-war-in-europe-1408675046>.

Tucker, JA, Metzger, M & Barbera, P 2014, 'SMaPP lab data report: Ukraine protests, 2013-2014', Social Media and Political Participation Lab, New York University, New York, NY, U.S.A., viewed 16 June 2016, <https://wp.nyu.edu/smapp/wp-content/uploads/sites/1693/2016/04/Ukraine_Data_Report.pdf>.

Tucker, P 2015, 'What your Facebook posts mean to U.S. special operations forces', *Defense One*, 29 January, viewed 17 June 2016, <http://www.defenseone.com/technology/2015/01/what-your-facebook-posts-mean-us-special-forces/104031/>.

Unwala, A & Ghori, S 2015, 'Brandishing the cybered bear: information warfare and the Russia-Ukraine conflict', *Military Cyber Affairs*, vol. 1, no. 1, pp. 1-11, viewed 30 June 2016, <http://scholarcommons.usf.edu/cgi/viewcontent.cgi?article=1001&context=mca>.

U.S. Army Special Operations Command 2015, *Little green men: a primer on modern Russian unconventional warfare, Ukraine 2013-2014*, Fort Bragg, NC, U.S.A., viewed 22 January 2017, <http://www.jhuapl.edu/ourwork/nsa/papers/ARIS_LittleGreenMen.pdf>.

——2014, *Counter-unconventional warfare*, white paper, Fort Bragg, NC, U.S.A., 26 September, viewed 16 June 2016, <https://info.publicintelligence.net/USASOC-CounterUnconventionalWarfare.pdf>.

U.S. Department of Defense 2010, *Joint publication 3-13.2: psychological operations*, U.S. Government Printing Office for the Joint Chiefs of Staff, Washington, D.C., U.S.A., 7 January, viewed 16 June 2016, <https://fas.org/irp/doddir/dod/jp3-13-2.pdf>.

——2014, *Joint publication 3-13: information operations*, U.S. Government Printing Office for the Joint Chiefs of Staff, Washington, D.C., U.S.A., 20 November, viewed 3 July 2016, <http://www.dtic.mil/doctrine/new_pubs/jp3_13.pdf>.

——2016, *Dictionary of military and associated terms*, 15 October, viewed 19 January 2017, <http://www.dtic.mil/doctrine/new_pubs/dictionary.pdf>.

U.S. Department of State n.d., 'Global Engagement Center', viewed 17 June 2016, <http://www.state.gov/r/gec/>.

Valentino-Devries, J & Yadron, D 2015, 'Cataloging the world's cyber forces', *The Wall Street Journal*, 11 October, viewed 22 June 2016, <http://www.wsj.com/articles/cataloging-the-worlds-cyberforces-1444610710>.

Vanden Broak, T & Locker, R 2012, 'U.S. info ops programs dubious, costly', *USA Today*, 24 February, viewed 2 July 2016, <http://usatoday30.usatoday.com/news/military/story/2012-02-29/afghanistan-iraq-military-information-operations-usa-today-investigation/53295472/1>.

Vandomme, R 2010, 'From intelligence to influence: the role of information operations', Canadian Force College, Center for National Security Studies, Canada, viewed 22 January 2017, <http://www.cfc.forces.gc.ca/237/251/vandomme-eng.pdf>.

Walker, S 2014, 'Russia and U.S. take their petty war of words over Ukraine on Twitter', *The Guardian*, 9 April, viewed 22 January 2017, <https://www.theguardian.com/world/2014/apr/09/russia-us-social-media-feelings-ukraine-clear>.

Weed, MC 2014, 'U.S. international broadcasting: background and issues for reform', Congressional Research Service, Washington, D.C., U.S.A., 2 May, viewed 3 July 2016, <https://www.fas.org/sgp/crs/row/R43521.pdf>.

The White House Office of the Press Secretary 2011, 'Executive Order 13584—developing an integrated strategic counterterrorism communications initiative', Washington, D.C., U.S.A., 9 September, viewed 17 June 2016, <https://www.whitehouse.gov/the-press-office/2011/09/09/executive-order-13584-developing-integrated-strategic-counterterrorism-c>.

Woehrel, S 2015, 'Ukraine: current issues and U.S. policy', Congressional Research Service, Washington, D.C., U.S.A., 12 February, viewed 16 June 2016, <http://fpc.state.gov/documents/organization/224484.pdf>.

Wright, BD 2009, 'Selling America, ignoring Vietnam: the United States Information Agency in South Vietnam, 1954-1960', master's thesis, University of British Columbia, Vancouver, BC, Canada, viewed 22 January 2017, <https://open.library.ubc.ca/cIRcle/collections/ubctheses/24/items/1.0070856>

Zhang, J, Brackbill, D, Yang, S & Centola, D 2015, 'Efficacy and causal mechanisms of an online social media intervention to increase physical activity: results of a randomized controlled trial', *Preventive Medicine Reports*, vol. 2, pp. 651-657, viewed 21 June 2016, <http://ac.els-cdn.com/S2211335515001072/1-s2.0-S2211335515001072-main.pdf?_tid=d5752082-37ca-11e6-bd9f-00000aab0f01&acdnat=1466525678_f4c69b9f485807770fdaefa61725c971>.

Managing Cybercrimes Through the Implementation of Security Measures

OK Enigbokan, N Ajayi

Discipline of Information Systems and Technology
College of Law and Management Studies
University of KwaZulu-Natal, South Africa
E-mail: jaidodabom@yahoo.com; ajayi@ukzn.ac.za

Abstract: *Today's global environment has seen tremendous growth in the number of online transactions and Internet subscribers. This increase is creating a situation in which businesses are now largely dependent on information systems and their inherent technologies. The increase, however, is also causing a rise in the volume and extent of cybercrimes and security lapses. Hence, countries, organisations, and individuals are regularly being faced with the challenge of protecting their privacy and integrity over the Internet. This study was conducted to identify effective security measures for managing cybercrimes and protecting organisations' information and information systems from cyber-related crimes.*

Keywords: *Information Systems, Cyber Security, Cybercrimes*

Introduction

The world has experienced momentous growth in areas such as telecommunications, online transactions, and social media because of the Internet and its related technologies (Dlamini & Modise 2012). The benefits of the Internet have been enormous, and its usage has grown rapidly. Hence, organisations are increasing their bandwidth capacity in order to meet users' demands (Grobler, van Vuuren & Zaaiman 2011). Organisations are, however, faced with an uphill task of providing Internet-related services to the increasing number of Internet users while at the same time ensuring that the users' information is well protected from cybercriminals (Grobler, van Vuuren & Jannie 2011; Grobler, van Vuuren & Leenan 2013; Dlamini & Modise 2012). The protection of data and information systems has remained a major concern, due to the rapid developments in computing technology (Grauman 2012). Organisations are increasingly becoming victims of cybercrime, and are, hence, continuously looking for effective means of preventing and managing cyber threats and cybercrimes (Grobler, van Vuuren & Zaaiman 2011; Grobler, van Vuuren & Leenan 2013; Dlamini & Modise 2012). Investigations by Norton (2013) reveal that a total of $113 billion has been lost by cyber victims to cybercrimes globally, with an estimated average of $298 per victim. A research study by Wolfpack (2012) revealed that in June 2012, one in every 171 emails was identified as a threat and one in every two emails was designated as spam. In order to conduct this study and to provide the reader with a good list of what organisations should be doing, security experts from three information security consulting firms in South Africa were interviewed. The findings from these interviews are summarised in this paper.

Literature Review

The Internet has been described as an environment that allows people to express themselves without any form of control or restriction (Lessig 1998). According to the High Representative of the European Union for Foreign Affairs and Security Policy (2013), the Internet facilitates the involvement of everyone in the sharing of information and networking. However, it is very difficult for nations of the world to control the behaviour of Internet users (Lessig 1998). Dlamini and Modise (2012) stated that although there are huge benefits from the Internet, it still presents a number of challenges—such as anonymity, pervasiveness, and encryption. The Internet has become an environment where cybercriminals lure people into divulging vital information. The U.S. Department of Homeland Security (2011) noted that the Internet is presently filled with evolving malware authors who design malicious programs to exploit flaws in information systems, with the aim of causing substantial damage.

The vulnerabilities exploited in information systems enable cybercriminals to easily steal, among other things, information, individuals' identities, and money (U.S. Department of Homeland Security 2011). Wolfpack (2012) stated that financial institutions have lost much financially after failing to effectively secure information systems from cyber attacks. The South African Post Bank's Johannesburg robbery attack, for example, which came in the form of unauthorised access to the bank's information systems by criminals, resulted in a loss of more than $6.7 million (Wolfpack 2012). Sony Corporation also was a victim of cybercrime due to a breach in data caused by hackers, resulting in the disclosure of personal information (such as emails, names, and login details) from more than 77 million users. Sony lost an estimated $171million as a result of this attack (Jooste 2012).

Cybercrimes are becoming very hard to identify; and, even when identified, their impacts are usually significant (Fortinet 2013; PricewaterhouseCoopers [PwC] 2011; Internet Policy Task Force 2011). These crimes vary from distributing malware, to phishing, and keylogging (PwC 2011). Several authors have shown that cyber-security measures are becoming necessary in managing cybercrimes (Vande Putte & Verhelst 2013; Pilling 2013; von Solms & van Niekerk 2013). Wolfpack (2012) stated that effectively managing cybercrimes would involve the creation of cyber-security institutes, a comprehensive national cyber-security structure, and cyber-security response teams. Fortinet (2013) also suggested that cooperation among nations in preventing attackers from regularly registering domains, the integration of response teams like CSIRT (Computer Security Incident Response Team) into security firms, and the proper implementation of security measures by organisations would go a long way in managing cybercrimes.

Methodology

A qualitative research approach featuring a non-probability sampling method was used for this study. A purposive sample was used to select respondents who were cyber-security professionals who have accumulated cyber-security knowledge over the years on various information or information systems' security procedures, standards, and tools.

Data collection was completed by conducting in-depth interviews with cyber-security experts from three different information security consulting firms in South Africa. The experience of the cyber-security experts provided rich insights on managing cybercrimes through the

implementation of security measures. In-depth interviews were adopted as the most appropriate means of data collection because they provide the investigator with a broader perspective, more detailed explanations, and richer insights on the phenomenon under study (Brown & Suter 2012). The interview data was analysed using the thematic data analysis technique—which allowed the researchers to classify and to extract a list of common themes or patterns from the interview data in order to generate an expression of the common voices of the respondents who were interviewed (Anderson 2007).

Findings and Discussion

Analysis of the interviews with cyber-security experts yielded responses that identified some security measures that can be adequately tailored to any organisation when managing cybercrimes. These are discussed below.

Vulnerability assessments and penetration tests

Security firms employ vulnerability assessments and penetration tests in order to properly examine the breadth and depth of a client's network or security posture, and also to determine loopholes that can be easily exploited by a cybercriminal. One interviewee stated that penetration tests are often carried out to compromise an organisation—just like a real hacker—in order to gain access to valuable details, such as emails of executives and board members. Organisations that conduct regular penetration tests by simulating attacks on their web applications, for example, are able to prevent financial losses and attacks—such as malware infections, Denial of Service attacks, and sabotage—that are often aimed at compromising the services and hosts. According to Dimkov *et al*. (2010), a successfully simulated attack during penetration testing against an organisation's network serves as an indication to the organisation that its security mechanisms are not well-aligned. The misalignment between an organisation's security mechanisms can easily be exploited using various forms of attack. Regular penetration tests also enable an organisation to comply with audit regulatory standards, for example the Payment Card Industry Data Security Standard (PCI- DSS) (Shivayogimath 2014).

In conducting a vulnerability assessment on an organisation's network, it is paramount to ensure that the tools that are adopted for the assessment are similar to those used by the potential attackers (Boyce 2001). This ensures that the methodologies and attack techniques employed by the attackers targeting an organisation's information system infrastructure can easily be duplicated. Organisations should adopt vulnerability assessment tools (such as Nmap, Nessus, and Whisker) that enable them, for example, to scan their networks, identify what service a host computer is running, and identify vulnerabilities on their web servers.

The common themes identified in this study regarding the types of security assessments that can be conducted by organisations, are discussed below.

- **Web application vulnerability assessments.** This form of assessment allows organisations to look for flaws and vulnerabilities in their web applications, before the vulnerabilities are exploited by attackers. These assessments enable organisations to be more thorough when evaluating every aspect of the security of a web application. They also allow more rigorous examination of security pertaining to session state, authentication, and application connections. Most web application vulnerability

assessments carried out via automated tools protect an organisation's network from attacks, such as cross-site scripting attacks, SQL injection, and Denial of Service (DoS). Tools such as Portswigger's Burp Suite are specifically run on an organisation's web application layers. After conducting a full scan through several components of the web application, the tool generates a report containing certain vulnerabilities (technical and logical) in the web application that can be mitigated, and best practices for the various programming languages used in developing the application. It then categorises certain identified vulnerabilities and their severity levels before providing a solution.

- **Mobile application assessments.** This form of assessment is carried out to determine if applications specific to mobile devices were developed using security best practices and measures, and at the same time, to test mobile devices' susceptibility to likely attacks. Mobile application assessments are often employed to ensure that data, applications, and networks are well protected. These assessments help minimise the risk to an organisation's brand, proprietary data, and market standing, and also make it difficult for other organisations to develop competitive intelligence. Behavioural assessments (a form of mobile application assessment) can be conducted with the use of a sandbox environment to, for example, thoroughly watch network traffic, file access, text messages, and phone calls. Static code analysis (mobile application code review), which is another form of mobile application assessment, can be used to detect triggered events and long-running timers on mobile applications. Mobile assessments are often done by running automated tools such as Burp Suite on customer-facing mobile applications in financial organisations. The software tools identify and examine aspects of the application (for example, dynamic contents and form fields) via objective measures that might leave the application vulnerable to attacks. The analytical results obtained using the various tools (whether manual or automated) can be used to determine potential business risk to an organisation.
- **Infrastructure assessment.** These are divided into internal and external assessments. Internal infrastructure assessment involves penetration testers going to clients' sites to conduct vulnerability scanning in an authenticated mode. This enables them to 'sniff' the traffic on a client's network to identify various forms of cybercrimes. On the other hand, external infrastructure assessment is usually carried out from outside an organisation's network using penetration testing tools (such as Metasploit and N-map) to analyse the organisation's corporate network infrastructure. The corporate network infrastructure can be accessed via the Internet or from third-party networks. Such assessments can help determine the level of exposure of an organisation's corporate network to external attacks.
- **Wi-Fi assessment.** This assessment helps organisations discover vulnerabilities specific to their Wireless Local Area Networks (WLAN) before attackers discover and exploit those weaknesses. By conducting wireless assessments on their networks, organisations can, for example, check compliance, identify unauthorised access points, and calculate the maximum distance that wireless traffic can be received. Organisations can choose to perform Wi-Fi assessments manually or via automated methods. Depending on the type of assessment conducted, tools such as Airsnort, Airmagnet, and standalone solutions with intrusion detection can be used to alert organisations in real-time to any discrepancies, changes, and/or suspicious activities on their network.

Incident response

According to respondents, the response ability of an organisation to a cyber incident and the availability of necessary strategic steps are factors that can help reduce the magnitude of a cyber attack. The response speed and strategy to address a cyber incident are dependent on the quality of an organisation's incident response process. One interview respondent stated that organisations that detect ongoing communications with information system assets and command and control (C&C) servers in their networks can employ incident response to enable them to effectively analyse and manage the identified malware or botnet.

Risk analysis was identified by respondents as a type of incident response strategy. According to respondents, risk analysis creates a platform for organisations to assess their risk levels by conducting simulated attacks that are often targeted at their employees. This finding is consistent with the work of Cichonski *et al*. (2012). One interview respondent stated that simulating attacks on employees creates a certain perspective on how much risk an organisation or its employees are exposed to. By simulating attacks, organisations can determine the kind of user education required for their employees. It also helps organisations determine whether there is a need to improve their attack response capabilities, and to identify other steps (from a more technical perspective) for mitigating attacks.

Perimeter defences and host-hardening procedures

Perimeter defences

Perimeter defence requires the creation of a tightly secure outer boundary around information systems in order to effectively control network traffic at every outgoing and incoming information channel (Ahmad, Maynard & Park 2014). Respondents advised that organisations should employ two-phased firewalls to protect the perimeter and the nodes of their network. This ensures that, if an attacker tries to access the subnet of the organisation's network, the information systems on that subnet are easily fire-walled and protected. Possible perimeter defence mechanisms identified by some of the respondents are discussed below.

- **Intrusion Prevention Systems (IPS).** This is an advanced and sophisticated mechanism for preventing network attacks (Shafi *et al*. 2010). The ability of an IPS to create logs of all events, to send automated emails, and to make phone calls, for example, makes them very vital when developing an organisation's defence strategy (Shafi *et al*. 2010). Respondents stated that an IPS can also be designed with capabilities to discover sophisticated attacks, to prevent various types of intrusion attempts, and to learn on its own about unknown future attacks. Network-based IPSs are usually located at the network gateway. Also, network-based IPSs are designed to intercept network traffic and to identify malicious or suspicious content within the traffic before taking the necessary steps to block potential attacks. Network-based IPSs protect network resources by employing attack signatures to discover known attacks and the malicious behaviour of data contents that pass through the network.
- **Intrusion Detection Systems (IDS).** IDSs are usually designed to report on a series of attacks or on more specific forms of attacks and on the extent of attack propagation, to name a few (Ahmad, Maynard & Park 2014). Some of the respondents stated that IDSs usually utilise signature- and anomaly-based detection paradigms. Signature-based

detection mechanisms employ patterns of common attacks to match and discover known network intrusions. They completely filter attack traffic based on the attack signature and separate good traffic based on a good signature. On the other hand, anomaly-based detection mechanisms also raise an alarm over observed activities on a network that differ significantly from the accustomed normal usage report. The combination of an IDS and localised firewall has been identified in the findings of this study as an effective means of protecting the information, information system, and internal subnets of organisations.

- **Web Application Firewall (WAF).** A Web Application Firewall represents a shielding mechanism created to protect web applications that can be accessed through Hypertext Transfer Protocol (HTTP) (Pubal 2015). Respondents stated that web application firewalls sit in front of most web applications and are capable of preventing application-level attacks (such as an SQL injection or a cross-site scripting attack) by monitoring network activity, or by identifying, alerting, and blocking malicious traffic that does not follow certain defined network rules or guidelines. Web Application Firewalls are especially effective during vulnerability exercises as they decrease remediation time and render unnecessary alteration of the source code in a web application. Some respondents stated that if WAFs are used as a part of a security monitoring infrastructure, they can significantly boost visibility into application traffic far beyond what is expected from an IDS or firewall.

Network segregation was identified by respondents as a perimeter defence type. Network segregation or segmentation is a technique that is often employed to limit the method and level of access to critical organisational information by certain information systems or certain individuals without the required permission (Australian Signals Directorate 2012). In most cases, these are often implemented on an organisation's network gateway. Network segmentation is the process of partitioning a network into smaller sub networks. It also involves designing and enforcing certain network rules and guidelines that control or monitor which information system is allowed to communicate with other kinds of information systems (Australian Signals Directorate 2012).

One cyber-security analyst stated that organisations should be very careful when using their own web servers for network assessments since that might expose them to attacks. He advised that an organisation should rent a server that is not connected to its network in order to enable it to easily carry out various kinds of network assessments or to make use of a part of its network that is completely segmented from critical data.

Some respondents advised that, in order to prevent the exploitation of an organisation's virtual box software (for example, Oracle VirtualBox software), organisations should properly segregate or isolate certain network services. The segregation or isolation of certain network services, for example, would ensure that, once the VirtualBox software of an organisation (with a branch in South Africa) is exploited, attackers would not easily have access to another VirtualBox software of that same organisation (with a branch in Germany). In most cases, attackers exploit the VirtualBox software located on the same network in an organisation. Hence, organisations should focus more on the proper isolation of their VirtualBox software that might need to

communicate/access the Internet. This can be achieved by not grouping VirtualBox software that requires Internet connectivity with those that do not require Internet connectivity.

According to Sobh and Aly (2011), data encryption techniques and other security mechanisms (such as tunnelling) are constantly needed to secure organisational assets and services from unauthorised users, to expose any back door or Secure Shell (SSH) ports to attackers, and to prevent critical data from being modified without detection as it passes through the Internet. Respondents advised that there should be adequate measures on the ground to control the use of SSH by an organisation. These SSH restrictions ensure that employees without necessary authorisation cannot easily access an administrator's machine/admin user when accessing a remote system. Also, the implementation of SSH restrictions or controls enables organisations to block the IP addresses of users who try to gain access to their network after three failed login attempts. One Information Technology (IT) manager explained that, in order to ensure that organisations' systems are adequately secured over the Internet, organisations should employ security techniques and tools such as shared key Open Authentication (OATH) management measures and run Fail2ban software (which scans log files for malicious IP addresses) on their machines.

Host-hardening procedures

Respondents stated that, before the initialisation of a project and the deployment of security architectures, organisations should properly address constraints on database tables and users' read/write access to tables in the database. One cyber-security analyst stated that since users of web servers are often writing to and pulling information out of a database, Database Administrators (DBAs) should enforce tight security controls. Respondents recommended that DBAs should ensure that users of web servers have the required permission to write to or pull information from a database. Read-only access should be given to users who are mainly pulling information from a database while write access should be given by DBAs to users that regularly write to a database. The write access given to users should only be applicable to specific tables in order to effectively harden the databases and to prevent users from changing data in special tables that only the administrator should have access to. Another host-hardening procedure involves DBAs ensuring that, between the web and the database server, an organisation's virtual box has access only to a database port.

Possibilities for employing host-hardening procedures for securing the Wi-Fi devices of organisations as recommended by respondents include

- using Media Access Control (MAC) address filtering;
- ensuring that guests on the guest Wi-Fi network cannot access an organisation's internal network; and,
- ensuring that demilitarised zones (DMZs) are correctly set up.

Interviewees further identified hard-disk encryption as a type of host-hardening procedure for securing the information and information systems of organisations. It was also suggested that hard-disk encryption be used to secure the laptops of employees who work remotely on a regular basis. One cyber-security analyst stated that a very effective way of securing the files on an employee's Operating System (OS) is by ensuring the employee's laptop requests user

authentication by password at the BIOS level before booting the laptop's OS. The password is used to decrypt the laptop's hard drive so that the employee can access the information on the laptop. Without the correct authentication, there would be no other means to recover an encrypted laptop's OS. In line with this response, Casey and Stellatos (2008) argued that the huge risks related to the exposure of Personally Identifiable Information (PII), coupled with the increased scrutiny from customers and media, has driven most organisations to adopt security solutions that encrypt data at rest. They further explained that encryption is one of the most effective security measures for preventing unauthorised access to data.

Network and log analysis

Respondents indicated that log analysis is an effective security measure that helps organisations monitor activities carried out by cybercriminals via Internet Protocols (IPs) on their networks. During log analysis, fraudulent IP addresses that have been used for cybercriminal activities can easily be traced. Respondents in this study indicated that, in order for organisations to be able to keep track of what is going on in their networks, they should employ Splunk software for strengthening, collecting, visualising, and indexing log and machine data. However, despite the huge benefits apparent from checking logs, most organisations hardly check their logs. This makes them prone to stealthier forms of cybercrimes. Network analysis, on the other hand, enables organisations to easily examine their network bases and identify malicious protocols. It also helps to evaluate and identify whom an attacker is communicating with and what they are stealing from a network. Respondents further advised that organisations employ network analysis for identifying and hacking into C&C communications in order to locate the section of organisations' networks that is transmitting to a C&C server. This would prevent the creation and propagation of botnets on an organisation's network in the future.

Security Information and Event Managers (SIEMs) were identified by respondents as log and network analysis mechanisms for securing the information and information systems. Organisations that employ SIEMs are able to effectively monitor, detect, and mitigate attacks on their networks with little or no human effort (Hansen 2013). Respondents also emphasised that, in order to ensure that SIEMs pick up excessive network utilisation at unusual times of the day, organisations should implement anomalous behaviour detection mechanisms in their SIEMs. As identified by Karlzén (2009), SIEMs collect and aggregate log data from various devices and applications through software known as agents, filtering uninteresting data, normalising the filtered data to a proprietary format, carrying out analysis through correlation (using contextual information), and alerting administrators in case of attacks. Karlzén (2009) further stated that SIEMs ensure that organisations are compliant with regulations that are related to data retention, which can be very helpful during litigation preparation and for forensic investigations. SIEMs are also capable of assisting organisations during network diagnosis.

Training and awareness

Regular employee training and awareness programs can make employees more knowledgeable about the various forms of cybercrimes. Educating employees to be more security conscious and to be proactive was identified as necessary under training and awareness. One finding of this study was that the training programs undertaken by organisations should not simply cover security at an IT or technical level, but should also include educating employees to be security conscious. Also, employees should be able to ask themselves questions such as What would be

the outcome of installing a certain piece of software on the organisation's network? Moreover, what risks or attacks could the software pose? According to McCormac, Parsons and Butavicious (2012), immediately after an individual is hired by an organisation, a considerable amount of time should be dedicated to train and educate the individual on the organisation's security policies and procedures. In relation to the training of employees to think and be proactive, study respondents reported that organisations should teach their employees to ask questions and think about the implications of their actions online. In agreement with this finding, Symantec (2014) stated that regular security-skills assessment and appropriate training of employees would help mitigate various attacks. Furthermore, organisations should constantly enlighten and train users on essential security protocols so as to reduce human errors to the barest minimum.

Measures for managing some common forms of cybercrimes

During the analysis of data, a different set of themes that represents security measures for addressing and managing more specific forms of cybercrimes were identified. These are discussed below.

Phishing

Phishing is a type of social engineering carried out by an attacker (also known as a phisher) in order to deceptively retrieve confidential information from legitimate users by mimicking electronic communications from a reliable or well-known organisation in an automated manner (Shi & Saleem 2012). Study respondents identified two-factor authentication and anti-phishing software as necessary and specific security measures for preventing phishing attacks.

- **Two-factor authentication.** Some respondents stated that two-factor authentication ensures that the true identities of clients are properly verified. Aloul, Zahidi and El-Hajj (2009) and Rathgeb and Uhl (2010) stated that, in order to improve the security capabilities of access control systems, organisations should employ two-factor authentication mechanisms that combine authentication techniques such as passwords and tokens to authenticate a user.
- **Anti-phishing software.** Organisations should ensure that anti-phishing software (that is, antivirus software with built-in firewalls) is installed on their employees' computers. One cyber-security analyst said that antivirus software and email spam filters installed on the Microsoft Exchange server of an organisation and on employees' computers would do a great job of preventing phishing attacks in an organisation.

Denial of Service attacks

According to Burden and Palmer (2003), Denial of Service (DoS) attacks are usually designed to disrupt access and to inhibit the use of specific Internet resources by legitimate users. In order to prevent or mitigate DoS attacks, Prasad, Reddy and Rao (2014) suggest the use of techniques such as statistical methods, machine learning methods, and data mining. The findings from this study revealed that the measures discussed below can also be adopted.

- **Proper patching procedures.** Organisations should regularly apply patches in order to secure the software packages running within the business environment. A red team penetration tester stated that, in order to prevent web server vulnerabilities that arise whenever users request certain web pages, organisations should regularly patch the

software packages running on their web servers. In line with this strategy, Avecto (2014), Olzak (2008), Sophos (2013), Wolfpack (2012), and Sharma, Kumar, and Sharma (2011) stated that organisations should ensure that applications within the business environment are patched (using, for example, Microsoft Office or Adobe Flash Player) and also should ensure that patching is done within a space of two days for high-risk vulnerabilities in applications.

- **Regular communication with Internet Service Providers (ISPs).** In order to mitigate more sophisticated forms of DoS attacks, such as Distributed Denial of Service (DDoS), organisations should constantly communicate with and establish a healthy relationship with their respective ISPs. An external network and web application tester stated that ISPs are usually in a strong position to mitigate most DDoS attacks. Hence, organisations need the influence of ISPs that control their underlying Internet infrastructure in order to effectively prevent DDoS attacks. Most organisations are often vulnerable to DDoS attacks because they do not have access to the underlying Internet infrastructure to mitigate them. Beitollahi and Deconinck (2012) stated that organisations should enhance cooperation with their respective ISPs and domain providers in order to properly manage sophisticated forms of cybercrimes and also to enhance the response time in trying to mitigate cybercriminal activities against them.
- **Application hardening with intelligence component.** In order to secure the applications/IT infrastructure of organisations from very sophisticated forms of DoS attacks, such as DDoS, a form of application hardening needs to be adopted. An interview respondent advised that applications that are susceptible to application bug-level attacks must be hardened, need to be built to withstand a lot of traffic, and should not accept millions of requests at a time. During application bug-level attacks, an attacker ensures that Internet resources are not easily accessible to users by interfering with system configurations or by rendering an application unavailable. This is done by transmitting packets that exploit application flaws in the target machine (Beitollahi & Deconinck 2012). Also, organisations that develop mission critical applications need to integrate a built-in intelligence component into their applications that possess the capability to make informed decisions in order to mitigate the effects of DDoS attacks. Furthermore, application-hardening techniques that can be used against application bug-level attacks (like ping-of-death) as recommended by respondents include IPSs, IDSs, and proper patching procedures (long-term). As identified by Kumarasamy and Asokan (2011), some forms of DoS attacks, such as application bug-level attacks, can be mitigated by ensuring that organisations enforce tight security policies and employ parameter defences such as firewalls, IDSs, IPSs, and vendor recommended patches for applications.
- **Multifaceted security measures.** To secure information systems from infrastructure-level attacks, one interview respondent advised that organisations should ensure that services such as Arbor networks are employed to black hole DoS traffic with their routing fabrics. The security measures against infrastructure-level attacks must be multi-faceted and should involve the regular review and implementation of, for example, security policies, staff education, and firewalls. Multi-faceted security measures can also involve the configuration of IT infrastructures in a manner that minimises the exposure of the infrastructures to potential attacks. This strategy prevents information systems from getting bombarded with too many requests. Beitollahi and Deconinck (2012) stated that organisations should employ distributed defence techniques that require the combination

of filtering and rate-limiting mechanisms for monitoring network traffic. The distributed defence techniques should also include improved cooperation between organisations and their respective ISPs or administrative domains to mitigate infrastructure-level attacks.

Attacks on mobile platforms

Malicious mobile applications have become powerful tools as cybercriminals increase their distribution of malware under the guise of authentic applications (RSA Research 2014). Interview respondents consistently suggested that applications designed for mobile platforms be regulated before distribution. In order to mitigate attacks on mobile platforms, mobile device platforms should be developed in a way that does not easily allow individual applications to get complete access to and control of the platforms they are running on. Respondents also advised that, in order to prevent mobile applications from HTML injection attacks, organisations (especially banks) should ensure that they properly regulate, monitor, and manage their applications on the Google Play Store. Respondents further stated that organisations that use the Apple App Store environment to run their applications are less susceptible to HTML injection attacks because of proper regulation/vetting of the apps on the store, the flexibility/security of Apple Cloud, and Apple's new mobility innovations, virtualisation technologies, and boot camp. One respondent stated that, generally, malicious applications running on the Google Play Store are a problem for most organisations because of poor monitoring of the digital media store. Malware authors take advantage of the poor monitoring of the Google Play Store by publishing counterfeit applications that look like legitimate apps available at the site. Some banks, for instance, have faced challenges with their mobile banking applications as a result of the increasing number of counterfeit applications at the Google Play Store. RSA Research (2014) reported that HTML injection techniques are used by cybercriminals to send users to a hyperlink that enables them to download malicious applications. During the installation of the malicious application, various permissions are requested in order to gain super user privileges. The installation of the malicious application allows full access to the mobile device's functionality and features and also prevents the owner from deleting the application.

Botnets

According to Bleaken (2010), botnets are distributed networks of personal computers infected with 'zombies' or 'bots' which enable them to be effectively controlled by cybercriminals. The aim behind the use of botnets is to carry out DDoS attacks, to distribute large volumes of spam and other malware infections via file distribution networks, and to email attachments, links to infected websites, and links to peer-to-peer (P2P) networks. Study respondents recommended the use of multiple antivirus engines and behavioural monitoring tools as security measures for preventing botnet attacks.

- **Multiple antivirus engines.** In order to effectively mitigate botnet attacks, multiple antivirus engines should be employed. One respondent stated that organisations should avoid using the same antivirus engine on their web gateways and endpoints. Organisations need to have diverse antivirus engines securing their web gateways and endpoints to be able to prevent bot attacks. Robust antivirus engines should preferably be employed on the critical points on an organisation's network. However, organisations should be very careful of a potential challenge that often involves multiple antivirus packages trying to compete and disrupt each other's processes during the discovery and

quarantine of malware. This is usually as a result of antivirus programs seeing other antivirus programs as a virus or threat. Most organisations are hardly faced with this challenge because they have antivirus programs handling several separate points on their network. (This ensures that they hardly interfere with each other.) In line with this finding, Oberheide, Cooke and Jahanian (2007) stated that, instead of running just a single/specific antivirus package on various points of an organisation's network, organisations should run a wide range of or multiple antivirus programs in parallel on an organisation's IT infrastructure. The integration of multiple/diverse antivirus packages on an organisation's network allows for resilience against attacks. Oberheide, Cooke, and Jahanian (2007) further stated that running multiple antivirus programs on several points on an organisation's network helps increase the detection rate of malicious programs.

- **Behavioural monitoring tools.** Behavioural monitoring tools (for example, IDSs) would enable organisations to monitor traffic on their networks and also identify communications made by botnets with C&C servers. One interview respondent stated that, whenever a payload targets an organisation's network, an attacker's bot is going to constantly send communications to a C&C server in order to get instructions that would enable the bot to execute the payload. Results of this study indicate that, for organisations to easily identify bots on their network, they should employ behavioural monitoring tools, such as FireEye.

Some respondents identified fast-flux hosting as a botnet attack technique employed against organisations' networks. Fast-flux hosting involves the dynamic distribution of network resources across an ever-changing range of IP addresses (Bleaken 2010). Security measures that can be employed against this include

- **Proactively registering fast-flux domains.** This method for mitigating fast-flux hosting is reactive in nature. Organisations should be quick to report an attacker's fast-flux domain botnets to antivirus firms in order to close down the domain and also to prevent the propagation of the fast-flux domain botnets on other organisations' networks. One interviewee stated that, by proactively registering an attacker's fast-flux domain before the attacker makes his/her fast-flux domain, botnets will become more robust and difficult to detect. Thus, organisations could more effectively manage fast-flux hosting. This measure enables antivirus firms to reverse-engineer the algorithms of the botnets, determine what domains the malware is going to use, register the domain, report the fast-flux domain botnet, and possibly shut it down. Organisations (such as F-Secure and Microsoft) and security agencies (such as the Dutch police) have actively employed these steps.
- **Regular self-education.** Organisations should constantly educate their employees on the dangers of fast-flux hosting techniques employed by bot masters. One respondent stated that constant training and awareness programs should be conducted by organisations in order to prevent fast-flux hosting attacks. He further stated that regular self-education conducted by organisations allows them to become more savvy by knowing which security measures work and which do not work with fast-flux hosting attacks.

Cost to an organisation when implementing security measures

Symantec (2012) estimates that cyber attacks on credit cards and other sensitive customer information cost the global economy $1 trillion dollars a year. The increasing economic costs due to these attacks has forced governments, organisations, and individuals to develop and implement robust cyber-security frameworks, awareness programs, and policies in order to ensure adequate security from stealthier forms of attacks in the future (Dlamini & Modise 2012). Booz Allen Hamilton (2015) reported that the global spending on cyber security by individuals, organisations, and government agencies was estimated at $76.9 billion. Moreover, one cyber-security analyst interviewed for this study noted that the cost incurred in trying to implement robust security measures against cybercrimes varies with organisation and often depends on the size of its security budget. Fielder *et al*. (2015) stated that most Small-to-Medium Enterprises (SMEs) face the challenge of having limited funding to support the implementation of security measures. Hence, most SMEs have to make trade-offs with regard to how their information systems are protected. On the other hand, Bernik (2016) noted that the security budget of most large-scale organisations is, in most cases, sufficient for managing and implementing security measures and tools such as firewalls, spam filters, and antivirus programs. In addition, SMEs usually take more time to implement and manage security measures because of the lack of human resources (Xie 2004). Lack of human resources for SMEs means that most do not have the required data, since collecting data annually by an expert is so expensive. This data, however, is necessary for making informed decisions and planning for the implementation of a less or a more sophisticated form of security measure (Xie 2004).

Conclusion

Cybercrimes are hampering the growth of most organisations. It is, therefore, of utmost necessity to consider security at all levels before a particular project is implemented or before an organisation deploys an information system infrastructure. Cybercriminals have evolved and have adapted continuously to the increasing developments and changes in computing technologies. Most cybercriminals employ complex tools and now use sophisticated attacks to exploit vulnerabilities in organisations' information systems. More aggressive malware is gradually evolving, with more polymorphic algorithms becoming a regular feature in recent attacks. Most organisations, aside from the basic procedures in securing information and information systems, do not really know how to secure their information and information systems from more complex attacks. The security measures discussed in this paper are, however, insufficient; effective policies and standards are necessary to reduce the consequences of future attacks.

References

Ahmad, A, Maynard, SB & Park, S 2014, 'Information security strategies: towards an organizational multi-strategy perspective', *Journal of Intelligent Manufacturing*, vol. 25, no. 2, pp. 357-370, viewed 10 April 2014, <https://pdfs.semanticscholar.org/8183/186dcf4c89e7caf6ee1e657b4f4deeddd960.pdf>.

Aloul, F, Zahidi, S & El-Hajj, W 2009, *Two factor authentication using mobile phones, Proceedings of the 2013 ACS International Conference on Computer Systems and Applications (AICCSA)*, Rabat, Morocco, <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.455.3299&rep=rep1&type=pdf>.

Anderson, R 2007, *Thematic Content Analysis (TCA): descriptive presentation of qualitative data*, Institute of Transpersonal Psychology, CA, U.S.A., viewed 1 July 2015, <https://books.google.com/books?id=Gma9pr7NXCgC&pg=PA320&dq=www.wellknowingconsulting.org/publications/articles.html&hl=en&sa=X&ved=0ahUKEwiS6aaYt9fSAhXI4CYKHf1PBvYQ6AEIITAB#v=onepage&q=www.wellknowingconsulting.org%2Fpublications%2Farticles.html&f=false>.

Australian Signals Directorate 2012, 'Network segmentation and segregation', *Information security advice for all levels of government*, Department of Defence, Australian Government.

Avecto 2014, *Achieving compliance*, white paper, viewed 1 May 2015, <https://www.avecto.com/media/1020/whitepaper-achieving-compliance.pdf>.

Beitollahi, H & Deconinck, G 2012, 'Analyzing well-known countermeasures against Distributed Denial of Service Attacks', *Computer Communications*, vol. 35, no. 11, pp. 1312-32, viewed 5 March 2015, <http://www.sciencedirect.com/science/article/pii/S0140366412001211>.

Bernik, I 2016, 'Cybercrime: the cost of investments into protection', *Journal of Criminal Justice and Security*, vol. 2, pp. 105-16, viewed 18 January 2017, <https://www.fvv.um.si/rV/arhiv/2014-2/01_Bernik.pdf>.

Bleaken, D 2010, 'Botwars: the fight against criminal cyber networks', *Computer Fraud & Security*, no. 5, pp. 17-19, viewed 1 April 2015, DOI: 10.1016/S1361-3723(10)70055-5, <https://www.researchgate.net/publication/250726774_Botwars_The_fight_against_criminal_cyber_networks>.

Booz Allen Hamilton 2015, *A practical approach to quantifying the financial benefits of cyber security*, white paper, viewed 17 January 2017, <https://www.boozallen.com/content/dam/boozallen/documents/2015/04/Cyber_ROI_Whitepaper.pdf>.

Boyce, R 2001, *Vulnerability assessments: the pro-active steps to secure your organization*, viewed 4 May 2016 <https://www.sans.org/reading-room/whitepapers/threats/vulnerability-assessments-pro-active-steps-secure-organization-453>.

Brown, TJ & Suter, TA 2012, 'Exploratory, descriptive, and causal research designs', *Marketing research*, Cengage Learning, Oklahoma State University, OK, U.S.A.

Burden, K & Palmer, C 2003 'Internet crime: cyber crime—a new breed of criminal?', *Computer Law & Security Review*, vol. 19, no. 3, pp. 222-27, viewed 5 March 2015, <http://www.sciencedirect.com/science/article/pii/S0267364903003066>.

Casey, E & Stellatos, GJ 2008, 'The impact of full disk encryption on digital forensics', *ACM SIGOPS Operating Systems Review*, vol. 42, no. 3, pp. 93-8.

Cichonski, P, Millar, T, Grance, T & Scarfone, K 2012, *Computer security incident handling guide: recommendations of the National Institute of Standards and Technology*, Special Publication 800-61, rev. 2, Gaithersburg, MD, U.S.A., National Institute of Standards and Technology, viewed 5 March 2015, <http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-61r2.pdf>.

Dimkov, T, van Cleeff, A, Pieters, W & Hartel, P 2010, 'Two methodologies for physical penetration testing using social engineering', *Proceedings of the 26th Annual Computer Security Applications Conference*, 6-10 December 2010, Austin, TX, U.S.A., pp. 399-408.

Dlamini, Z & Modise, M 2012, 'Cyber security awareness initiatives in South Africa: a synergy approach', *7th International Conference on Information Warfare and Security*, University of Washington, Seattle, U.S.A., viewed 12 March 2017, <http://researchspace.csir.co.za/dspace/bitstream/10204/5941/1/Dlamini_2012.pdf>.

Fielder, A, Panaousis, E, Malacaria, P, Hankin, C & Smeraldi, F 2015, 'Comparing decision support approaches for cyber security investment', 19 February, viewed 17 January 2017, <https://arxiv.org/pdf/1502.05532>.

Fortinet 2013, *Fortinet 2013 cybercrime report*, white paper, viewed 13 July 2014, <http://www.fortinet.com/sites/default/files/whitepapers/Cybercrime_Report.pdf>.

Grauman, B 2012, *Cyber-security: the vexed question of global rules*, viewed June 2014, <http://www.friendsofeurope.org/media/uploads/2015/06/SDA-Cyber-report-FINAL.pdf>.

Grobler, M, van Vuuren, JJ & Zaaiman, J 2011, 'Evaluating cyber security awareness in South Africa', Workflow Series 6845, Council for Scientific and Industrial Research, Pretoria, South Africa, viewed 13 July 2014, <http://researchspace.csir.co.za/dspace/handle/10204/5108>.

——, ——& Leenan, L 2012, 'Implementation of a cyber security policy in South Africa: reflection on progress and the way forward', *Proceedings of the 2012 IFIP International Conference on Human Choice and Computers—HCC 2012: ICT critical infrastructures and society*, eds. MD Hercheui, D Whitehouse, W McIver & J Phahlamohlaka, IFIP Advances in Information and Communication Technology, vol. 386. Springer, Berlin, Heidelberg, Germany, viewed 13 July 2014, <http://link.springer.com/chapter/10.1007/978-3-642-33332-3_20>.

Hansen, S 2013, 'Cybercrime security risks and challenges facing business', Symantec, *East Africa Security Conference*, East Africa.

High Representative of the European Union for Foreign Affairs and Security Policy 2013, *Cybersecurity strategy of the European Union : an open, safe and secure cyberspace*, European Commission, JOIN (2013), Brussels, Belgium, viewed 11 March 2017, <http://eeas.europa.eu/archives/docs/policies/eu-cyber-security/cybsec_comm_en.pdf>.

Internet Policy Task Force 2011, *Cybersecurity, innovation and the Internet economy*, U.S. Department of Commerce, viewed 11 March 2017, <https://www.nist.gov/sites/default/files/documents/itl/Cybersecurity_GreenPaper_FinalVersion.pdf>.

Jooste, J 2012, 'South African businesses unprepared for the growing risk of cyber attacks', *FIA Insight*, 4th Quarter, p. 37.

Karlzén, H 2009, *An analysis of security information and event management systems: the use of SIEMs for log collection, management and analysis*, MS thesis, Chalmers University of Technology, University of Gothenburg, Göteborg, Sweden.

Kumarasamy, S & Asokan, R 2011, 'Distributed Denial of Service (DDoS) attacks detection mechanism', *International Journal of Computer Science, Engineering and Information Technology*, vol. 1, no. 5, pp. 39-49.

Lessig, L 1998, 'The laws of cyberspace', *Taiwan Net '98*, March, Taipei, Taiwan, viewed 5 July 2014, <https://cyber.harvard.edu/works/lessig/laws_cyberspace.pdf>.

McCormac, A, Parsons, K & Butavicious, M 2012, *Preventing and profiling malicious insider attacks*, Edinburgh South Australia, Australia, Command, Control, Communications and Intelligence Division, DSTO Defence Science and Technology Organisation, (DSTO-TR-2697).

Norton 2013, *2013 Norton report*, viewed 5 July 2014, <http://www.yle.fi/tvuutiset/uutiset/upics/liitetiedostot/norton_raportti.pdf>.

Oberheide, J, Cooke, E & Jahanian, F 2007, 'Rethinking antivirus: executable analysis in the network cloud', *Proceedings of the 2nd USENIX workshop on Hot Topics in Security—HOTSEC'07*, USENIX Association, Berkeley, CA, U.S.A., viewed 11 March 2017, <https://www.usenix.org/event/hotsec07/tech/full_papers/oberheide/oberheide.pdf>.

Olzak, T, 2008, *Keystroke Logging (Keylogging)*, viewed 9 May 2014, <http://adventuresinsecurity.com/images/Keystroke_Logging.pdf>.

Pilling, R 2013, 'Global threats, cyber-security nightmares and how to protect against them', *Computer Fraud & Security*, no. 9, pp.14-8, DOI: 10.1016/S1361-3723(13)70081-2.

Prasad, KM, Reedy, ARM & Rao, KV 2014, 'DoS and DDos attacks: defense, detection and traceback mechanisms—a survey', *Global Journal of Computer Science and Technology: E-Network, Web & Security*, vol. 14, no. 7, pp.15-32.

PricewaterhouseCoopers (PwC) 2011, *Cybercrime: protecting against the growing threat Global Economic Crime Survey*, viewed 6 July 2014, <http://pwc.blogs.com/files/2011-global-economic-crime-survey-report.pdf>.

Pubal, J 2015, *Web application firewalls: enterprise techniques*, white paper, SANS Institute InfoSec Reading Room, viewed 13 April 2016, <https://www.sans.org/reading-room/whitepapers/application/web-application-firewalls-35817>.

Rathgeb, C & Uhl, A 2010, 'Two-factor authentication or how to potentially counterfeit experimental results in biometric systems', *Proceedings of the 7th international conference on Image Analysis and Recognition, ICIAR'10,* Springer-Verlag Berlin Heidelberg, Germany.

RSA Research 2014, *The current state of cybercrime 2014: an inside look at the changing threat landscape*, EMC Corporation, white paper, viewed 5 February 2015, <https://www.emc.com/collateral/white-paper/rsa-cyber-crime-report-0414.pdf>.

Shafi, MI, Akram, M, Hayat, S & Sohail, I 2010, 'Effectiveness of Intrusion Prevention Systems (IPS) in fast networks', *Journal of Computing*, vol. 2, no. 6, pp. 78-84, viewed 10 April 2016, <https://arxiv.org/abs/1006.4546>.

Sharma, G, Kumar, A, & Sharma, V 2011, 'Windows operating system vulnerabilities', *International Journal of Computing and Corporate Research*, vol. 1, no. 3, viewed 10 April 2016, <http://www.ijccr.com/index.php/papers-selected-november-2011>.

Shi, J & Saleem, S 2012, *Phishing*, viewed 8 April 2015, <https://www2.cs.arizona.edu/~collberg/Teaching/466-566/2013/Resources/presentations/2012/topic5-final/report.pdf>.

Shivayogimath, C 2014, 'An overview of network penetration testing', *International Journal of Research in Engineering and Technology*, vol. 3, no. 7, pp. 408-13, viewed 7 April 2016, <http://esatjournals.net/ijret/2014v03/i07/IJRET20140307070.pdf>.

Sobh, T & Aly, Y 2011, 'Effective and extensive Virtual Private Network', *Journal of Information Security*, vol. 2, pp. 39-49.

Sophos 2013, *Security threat report 2014: smarter, shadier, stealthier malware*, viewed 30 April 2015, <https://www.sophos.com/en-us/medialibrary/PDFs/other/sophos-security-threat-report-2014.pdf>.

Symantec 2012, *The ongoing malware threat: how malware infects websites and harms businesses—and what you can do to stop it*, white paper, viewed 7 May 2015, <https://www.geotrust.com/anti-malware-scan/malware-threat-white-paper.pdf>.

——2014, *Internet security threat report 2014*, vol. 19, viewed 1 March 2015, <http://www.symantec.com/content/en/us/enterprise/other_resources/bistr_main_report_v19_21291018.en-us.pdf>.

U.S. Department of Homeland Security 2011, *Enabling distributed security in cyberspace: building a healthy and resilient cyber ecosystem with automated collective action*, white paper, viewed 12 July 2014, <https://www.dhs.gov/xlibrary/assets/nppd-cyber-ecosystem-white-paper-03-23-2011.pdf>.

Vande Putte, D & Verhelst, M 2013, 'Cyber crime: can a standard risk analysis help in the challenges facing business continuity managers?', *Journal of Business Continuity & Emergency Planning*, vol. 7, no. 2, pp. 126-37.

von Solms, R & van Niekerk, J 2013, 'From information security to cyber security', *Computers & Security*, vol. 38, pp. 97-102, viewed 12 May 2014, DOI:<http://dx.doi.org/10.1016/j.cose.2013.04.004>.

Wolfpack 2012/3, *The South African cyber threat barometer, a strategic public-private partnership initiative to combat cybercrime in South Africa*, British High Commission, viewed 23 March 2014, <https://www.wolfpackrisk.com/wp-content/uploads/2016/03/SA-2012-Cyber-Threat-Barometer_Medium_res.pdf>.

Xie, N 2004, *SQUARE project: cost/benefit analysis framework for information security improvement projects in small companies*, U.S. Department of Defense, Software Engineering Institute, Carnegie Mellon University, viewed 18 January 2017, <https://resources.sei.cmu.edu/asset_files/TechnicalNote/2004_004_001_14360.pdf>.