CCDCOE
NATO COOPERATIVE
CYBER DEFENCE
CENTRE OF EXCELLENCE

Principles of Cyber Deterrence and the Challenges in Developing a Credible Cyber Deterrence Posture

İhsan Burak Tolga
NATO CCDCOE Strategy Branch Researcher

Tallinn 2018

About the author

Ihsan Burak Tolga is a researcher at the Strategy Branch of Cooperative Cyber Defence Centre of Excellence (CCDCOE) in Tallinn, Estonia. Upon graduating from Turkish Naval Academy, he served in different positions across the Turkish Naval Forces. He holds a BSc in computer engineering and MSc in computer science. His scientific interests focus on cyber deterrence and cyber security of combat management systems.

CCDCOE

The NATO Cooperative Cyber Defence Centre of Excellence (CCDCOE) is a NATO-accredited cyber defence hub focusing on research, training and exercises. It represents a community of 21 nations providing a 360-degree look at cyber defence, with expertise in the areas of technology, strategy, operations and law. The heart of the Centre is a diverse group of international experts from military, government, academia and industry backgrounds.

The CCDCOE is home to the Tallinn Manual 2.0, the most comprehensive guide on how International Law applies to cyber operations. The Centre organises the world's largest and most complex international live-fire cyber defence exercise Locked Shields. Every spring the Centre hosts in Tallinn the International Conference on Cyber Conflict, CyCon, a unique event joining key experts and decision-makers of the global cyber defence community. As of January 2018 CCDCOE is responsible for identifying and coordinating education and training solutions in the field of cyber defence operations for all NATO bodies across the Alliance.

The Centre is staffed and financed by its member nations – to this date Austria, Belgium, the Czech Republic, Estonia, Finland, France, Germany, Greece, Hungary, Italy, Latvia, Lithuania, the Netherlands, Poland, Portugal, Slovakia, Spain, Sweden, Turkey, the United Kingdom and the United States. NATO-accredited centres of excellence are not part of the NATO Command Structure.

www.ccdcoe.org
publications@ccdcoe.org

Disclaimer

This publication is a product of the NATO Cooperative Cyber Defence Centre of Excellence (the Centre). It does not necessarily reflect the policy or the opinion of the Centre or NATO. The Centre may not be held responsible for any loss or harm arising from the use of information contained in this publication and is not responsible for the content of the external sources, including external websites referenced in this publication.

Digital or hard copies of this publication may be produced for internal use within NATO and for personal or educational use when for non-profit and non-commercial purpose, provided that copies bear a full citation.

Table of Contents

1. Abstract

The application of deterrence theory in cyberspace focused on how to establish a deterrence posture, what steps should be taken in this process, what practices should be followed and what kinds of effects each one of those practices would have. The introduction of the concept of cyber deterrence has produced substantial knowledge about the definition of deterrence in cyberspace and its methods. However; as Karl Rainud Popper and his theory 'empirical falsification' suggests:

> '...the criterion of the scientific status of a theory is its falsifiability, or refutability, or testability.' (Popper, 1963)

In other words, a claim supported by thousands of positive examples of its benefit is not universally true, and just giving a single counter-proof is sufficient to prove that the claim is not true. Hence, this paper discusses if it is possible to develop a credible deterrence in cyberspace by examining the past cases in which the deterrence efforts have failed. It then objectively analyses and filters out the methods that contributed to those failures, reaching a refined subset. The aim of this paper is to have a set of real-case practices which have practical value in building a credible deterrence posture in cyberspace.

2. Introduction

In maintaining the status quo for any domain, including cyber defence, nations usually have limited resources that need to be consumed in a logical fashion. The nations' understanding of the subject is usually the biggest factor in their success. Nations strive to take the most rational and reasonable approach for producing the most cost-efficient way of implementing their necessary functions.

To develop a credible deterrence posture in 21st century cyberspace actors need to take the most logical measures to keep their operations robust and available. States and organisations like NATO and the EU need to fully integrate cyber deterrence considerations into all aspects, from people's daily lives to critical infrastructure. Conducting a sophisticated cyber attack is expensive (Slayton, 2017) but falling victim to one or many successful above-medium-sized cyber attacks is even more expensive (Chirgwin, 2018). Given this fact, and with respect to the non-lethal characteristics of cyber attacks in recent history, the main goal for a state or an organisation should be decreasing the number of successful cyber attacks targeting them by spending less than the damage that would be inflicted if they sat idle and did nothing (Denning, 2016). Although this seems to be a shift from traditional deterrence theory, it still holds true to the main goal of deterrence which is to minimize harm. Damage refers not only to physical harm to the victim`s IT infrastructure and operations, but also the negative psychological influence on the society.

Since the realization that 'cyber' and the concept of Cold War era deterrence, a wide array of valuable research, analyses and work have been done to assist actors in building a successful deterrence posture applied to cyberspace. The majority of these efforts focused on producing an answer to the question of 'how'. As a result, numerous methods, approaches and real-life examples have been suggested to add more value to nations' deterrence postures. However, these deterrence postures, with reasonable variance between states, seem to be failing. The occurrence of cyber attacks does not mean that the attackers are achieving their goals, but it suggests deterrence, as a full-scale concept, is failing with each minor and major cyber intrusion, malware, cyber espionage, data theft and security breach (CNBC, 2017). These attacks are viewed as acceptable if the sum of the initial calculation still stays on the intended side and the damage prevented exceeds the effort spent on preventing it. Additionally, when we add the damage already being inflicted on top of the deterrence efforts already in place, the effect of a successful cyber attack is even greater and the cost increases with each additional incident. The nature of deterrence is also difficult because motivations vary. Some attacks might be materialistic (i.e. legal bindings) or others may be intangible (morally good and bad acts) Motivations differ for each person and environment. Since successful deterrence is not possible to accurately measure, it is extremely hard to measure and grade deterrence.

This picture is a gloomy one, but it also provides a good opportunity to refine the associated cost of cyber deterrence (Libicki, 2009). Instead of juggling different deterrence strategies against evolving cyber attacks from numerous adversaries, the focus can be shifted to examples where deterrence has failed. The noise factors in this paradigm can be filtered out and then root causes can be identified. This paper attempts to answer the question "Why cyber deterrence fails in real life, and how?" by applying Popper's (Popper, 1959) Falsification Theory. These finding can then be examined against the Pareto Principle (Kaplow, 2005) to identify the major causes and characteristics which contribute to failing cyber deterrence. This will provide deterring actors greater flexibility, as well as, a realistic and cost effective target for developing a credible cyber deterrence posture.

The first two sections of this paper discuss cyber deterrence, where it applies and its boundaries. The paper also tries to determine at what point we admit that deterrence has failed, by reviewing two well-studied historic examples of general deterrence theory. The following sections further examine deterrence failures, analysing two recent real-world examples: the Petya/NotPetya and Shamoon cyber

incidents. Later chapters analyse and categorise the chosen attacks, and finally conclusions and recommendations are presented

3. Deterrence posture in cyberspace

Deterrence is: 'a strategy intended to dissuade an adversary from taking an action not yet started, or to prevent them from doing something that another state desires', or '[d]issuading someone from doing something by making them believe that the costs to them will exceed their expected benefit' (Nye, 2017). This consideration should occur in the adversary's mind. Due to its rather abstract nature, making an analogy to the Cold War era's nuclear deterrence posture has been useful to provide a solid reference to discuss this phenomenon. In the nuclear weapons context, deterrence was achieved due to the weapons' immense destructive power. Use of these weapons would cause the targeted nation to respond in kind due to the long transit time of the warhead providing defender the opportunity to counterattack (Second strike chance). This fact negates any possible gain by either nation. The concept of mutually assured destruction rendered any plan involving a decisive first-strike nonsensical. This characteristic is quite different from the instant nature of cyber attacks.

Deterrence theory in cyberspace differs from the classic nuclear deterrence and conventional deterrence in the aspects of actors and means. Cyber deterrence, at its very core, is a result of states' desire to avoid being attacked in or via cyberspace. Potential targets include their military networks, the networks of state or private firms or any element of the state critical infrastructure (industrial systems, finance, publicity, communication lines, power grid and transportation). The state also needs to understand the interdependencies of critical infrastructure to societal continuity and the psychological impact an attack could have on the public psyche as indirect impacts of a cyber attack. Defence of cyber elements should not be treated different to efforts to defend against conventional attacks. The biggest difference when comparing nuclear to cyber deterrence is that the effects of a strike in cyberspace are far from being as absolute as in nuclear warfare (Nye and Winter, 2011) The consequences of a cyber attack are also significantly smaller. There is less will to deter actions in cyber space, causing weakened deterrence. This allows actors to behave more boldly in cyberspace both in peace or war. The immense investment and political will required to acquire and maintain nuclear weapons is also absent in cyberspace and concealing the existence of cyber capabilities is much easier. Classical deterrence theory (Libicki, 2009) research proposes that there are two main methods of cyber deterrence: deterrence-by-denial and deterrence-by-punishment, Although Nye suggests two other methods of deterrence; entanglement and normative taboos (Nye., 2017).

Deterrence-by-denial relies on the principle that: 'If cyber attacks can be conducted with impunity, the attacker has little reason to stop' (Libicki, 2009). In such cases; independent of how much offensive cyber power a nation possesses, the attacker will not be deterred from conducting cyber attacks. If we consider the use of third parties and proxies for such actions, the attribution problem becomes even bigger. To enable deterrence-by-denial efforts to be successful, the main goal should be turning the cost-benefit ratio for the attacker on the prospective cyberattack above one, in other words, convincing the adversaries that the potential benefit they obtain from the damage inflicted or the intelligence they collect will be less than the effort and resources they need to execute the attack. As Philbin put it: '[a] strong defence deters an attack by convincing an attacker there will be no gains commensurate with the cost of attack' (Philbin, 2013)

The second accepted method of deterrence in cyberspace is deterrence-by-punishment, which suggests that:

> 'The aim of deterrence is to create disincentives for starting or carrying out further hostile action. The target threatens to punish bad behaviour but implicitly promises to withhold punishment if there are no bad acts or at least none that meets that threshold. If the attacker can be persuaded to reduce its efforts in the face of punishment, the defender can save

> some of what it would have spent on defence and still achieve the same level of security' (Libicki, 2009)

In this sense, it is similar to nuclear deterrence in which the parties are mutually assured that in case of a nuclear strike, there will be at least an equal and an opposite reaction. This deterrence-by-punishment approach seems like a better and more cost-effective solution, but attribution in the cyber environment is difficult.

> 'If deterrence is to work before the first retaliation takes place, others must have confidence that the deterring state will know who attacked it. Hitting the wrong person back not only weakens the logic of deterrence (if innocence does not matter, why be innocent?) but arguably makes a new enemy […]The defender must not only convince itself but should also convince third parties that the attribution is correct' (Libicki, 2009).'

Victims of cyber attack need to assure third-party nations about the true source of this attack in order to justify their reaction. Attribution of a sophisticated attack cyberspace can take a long time with current forensics techniques and capabilities; by assuming it is a sophisticated attack. Besides the fact that it is difficult to accurately attribute an attack, there is another drawback with the deterrence-by-punishment approach: the need for the deterring actor's credibility of a punishment, as for the action must occur in a timely fashion. (Libicki, 2009) There is a need to convince a possible adversary that a cyber attack will be addressed quickly after it has been accurately attributed. The deterring actor must convey their will and determination to act in response to a cyber attack. Broadcasting an assurance beforehand about a swift punishment in case of a cyber attack to take place is a deterrence strategy. However, not delivering punishment in a timely fashion for any reason (i.e. attribution problems) will seriously harm the deterring actor's credibility and with each following cyber attack, getting closer to zero (Iasiello, 2014). It is also worth mentioning that the deterrence-by-punishment efforts also usually have an opportunity cost, a cost that is forfeited in cases that the associated efforts do not meet their aim to prevent a potential adversary to conduct offensive behaviours in cyberspace.

By a rough comparison, deterrence-by-denial seems to be a more practical approach, although both methods gain more weight when they are put into action in harmony. Deterrence-by-denial makes the task of an attacker more difficult. Deterrence by punishment has a multiplier by raising the potential cost of the attack even though the odds of getting caught remain the same. Hence the combination of the two methods still play an important role for deterring the possible attacker.

There is much research and analysis of the key differences and how to establish a credible deterrence posture in cyberspace against state actors, non-state actors, hacktivists and cyber terrorists. As a result, very thorough and standard practices have been established. Yet the number of cyber incidents has not decreased. It seems that recommended deterrence theories are not working as designed in real life. Successful examples of deterrence in cyberspace are hard to accurately count, because it is almost impossible to assess a deterrence practice as successful in a discrete time frame due to the need of making assumptions about given circumstances (Lonsdale, 2017). Almost every day multiple cyber incidents take place either against individuals, organisations or states, which shows these methods are not deterring hostile actions in cyberspace.

To address this fact, this paper will identify causes that, in the past, have become examples of deterrence failure. The paper will apply universal Pareto dispersion to the study which states: 20% of the efforts to establish a strong deterrence should cover 80% of the success zone. With that percentage in mind and the theory of 'empirical falsification', causes of failed deterrence will be identified. A single example of failure, as the empirical falsification method suggests, will identify methods that do not work. This will leave a smaller set of refined and proven methods for building a credible deterrence posture which will save a nation's / organisation's valuable resources. The main principle to be followed is that

sufficient justification for the deterrence method needs to be provided before tagging as a failure for each instance. This evaluation also concedes that not all actors respond to the same deterrence efforts in the same fashion. At the end of this analysis, the excess efforts saved from unnecessary deterrence actions can be spent on the most-critical or mission-specific aspects of cyber defence.

4. Deterrence failures in recent history

One of the best-known cases in which deterrence efforts failed in recent history is the Chinese entry into the Korean War in 1950:

> 'If we designate the Chinese, as does Lebow, (Lebow, 1981) as the deterring party, then the United States' decision to cross the 38th parallel and proceed toward the Yalu would be the failure of deterrence, brought on, according to Lebow's theory, by a failure of rationality caused by intense domestic pressure to unify Korea' (Orme, 1987).

A Chinese official stated on 3 October 1950 that China would send its military forces to defend North Korea if any military forces apart from South Korea's crossed the 38th parallel. A warning by Chinese Ministry of Foreign Affairs stated the following:

> 'Now that the American forces are attempting to cross the 38th parallel on a larger scale, the Chinese people cannot stand idly by with regard to such a serious situation' (Whitting, 1960).

Detailed inspection of Chinese official statements show they did not communicate clearly what China wanted to deter. The US interpretation was that China's main concerns were about the availability of power plants located south of the Yalu which powered Chinese settlements in Manchuria. (Appleman, 2012) The reality of the situation was China did not desire the complete destruction of the North Korean state. A unified Korea would be heavily influenced by the US. The Chinese saw this as a possible hostile power with direct access to its border (Whitting, 1960). This did not directly correlate with US or UN troops' crossing of 38th parallel. Had the U.S. realized the actual Chinese concern regarding the situation, the decision makers may have reconsidered their actions with respect to the Chinese military capabilities. Following the transaction of statements and initial actions, the Chinese acted covertly. They purposely understated the size of their troops (approximately 300.000) headed towards the Korea border. The troops marched mostly during the night and they addressed their units with terms referring smaller sizes in their communications, caused a big misunderstanding in the opposite side. (Orme Spring 1987) Aware that the US had complete air superiority China decided not to risk exposing their troops to air strikes. Instead China focused on tactical surprise. The US was not aware of the Chinese intention. Had military leaders understood the size of the force China was sending to Korea they might have been deterred from their chosen course of action (Appleman 2012). As this example demonstrates, deterrence can only be effectively achieved if both actors understand the situation from the same point of reference.

The Cuban missile crisis is another example of failed deterrence. The Soviet Union underestimated the resolve the US would display towards limiting Soviet offensive military equipment so close to the US mainland. The Soviet Union was over-confident that by the time the US noticed the covert transportation of nuclear missiles to Cuba, the US would have no choice but to accept these actions. (Glenn, 2017). Statements by the US indicated there was no indication of any offensive ground-to-ground missiles or any other offensive capability by the Soviets, yet the White House issued a statement on September 4th 1962 that: '[w]ere it to be otherwise the gravest issues would arise' (Naftali, 2001), which conveys a clear enough commitment from the US. It is also important to note that US had a distinct military superiority compared to the Soviet Union at the time, both conventionally and with respect to its nuclear arsenal. (Glenn 2017) From the Soviet perspective; their missiles were already en route to Cuba by that time, a decision that was not influenced by the White House statement.

For deterrence to be affective, the actor conducting the deterring actions needs to understand the rational motivations of his adversary. According to Khrushchev's memoir, he assumed the US would realise it was not to the US advantage to wage war against the Soviet Union, so it would accept the Soviet's bold military initiative:

> 'But he (Kennedy) understood that the socialist camp had gained such economic and cultural might – and was in possession of so much scientific and technical knowledge, including the means of war – that the United States and its allies could no longer seriously consider going to war against us, I will always respect him for that.' (Khrushchev, 1976)

Khrushchev's statements during the Vienna Talks supported this conclusion. (Khrushchev 1976) US attempts to deter the Soviets from deploying their missiles in Cuba failed for several reasons. Khrushchev believed Kennedy lacked credibility. Khrushchev also wrongly interpreted Kennedy's performance during Vienna, the Bay of Pigs and his liberal background which led Soviet decision-makers to the false conclusion that he would not react decisively and firmly to Soviet actions.

A very difficult aspect to successful deterrence is actors tend to assume their adversaries have the same, or at least similar, logical paths to reach conclusions and make decisions. They assume actors are provided with adequate information about the other side's capabilities and motivations. Tangible metrics to measure the cost and benefits of an attack are crucial to assess the situation correctly; which also make the basis of game theory and zero-sum game. Yet when the states' perception of their reputation or their concerns about other intangible factors are introduced to the equation, the margin of error expands and sometimes exceeds the limits of the calculated reactions.

5. Petya, Notpetya attack

In 2016, the malware dubbed Petya after the destructive satellite in the 1995 James Bond movie GoldenEye, affected a limited number of computers running Windows operating system across the globe. The immediate analyses (Malware Tech, 2017) classified it as a new evolutionary phase of malware. Petya basically works as a Trojan horse malware triggered by user actions like opening an email attachment. It infects the master boot record to execute its source code, which encrypts the files on the infected computer's hard drive. When the user tries to turn on the computer, instead of the opening screen, a ransom note is shown which requests the user to wire a specific amount of money in bitcoins to get the key for decrypting the file system (Symantec, 2017).

Ukraine celebrated Constitution Day on 27 June 2017. It was a national holiday as well as the beginning of a major, global-scale cyber attack using an evolved version of Petya known as NotPetya. It was dubbed by Kaspersky due to its similarities to Petya. (New Petya / NotPetya / ExPetr ransomware outbreak 2017) Like the WannaCry malware used earlier the same year, NotPetya used EternalBlue (Grossman, 2017), a vulnerability in Windows' ServerMessageBlock (SMB) protocol. The initial attacks seemed to target Ukrainian companies (Nicole Perlroth, 2017) but similar cyber attacks occurred mainly in Germany, Belgium, Brazil, Russia, the United Kingdom and the United States (Microsoft, 2017). The malware was different from its predecessor. Instead of demanding a ransom, it was designed to spread quickly and target mostly energy infrastructure like power stations, gas stations, power grid, banks and airports (LogRhythm, 2017). The initial inspection of the malware indicated that NotPetya was using the software update mechanism of M.E.Doc – a de facto Ukrainian tax preparation program which had a security backdoor. The program was used extensively by the malware to speed up its spreading (ESET, 2017). Although the analysis report indicates the backdoor was a thoroughly well-planned and well-executed operation, (Mike Oppenheim 2017) the developers of the program denied the allegations and stated that they were also affected by the malware. Further analysis of the servers of the tax preparation service revealed there was no software updates since 2013. There was also evidence TeleBots, a cyber-espionage group, had access to a server from a compromised employee account. (Cimpanu 2017)

The damage not only affected Ukrainian critical infrastructure, but for instance TNT Express / FedEx also reported that the malware caused $300 million worth of damage in their corporate computer systems. The shipping company Maersk reported similar effects (Chirgwin, 2018), all probably due to inherent vulnerabilities in their systems. At the state level, the results were also psychological, as some cities in western Ukraine suffered a power outage for several hours due to the malfunction in the computers of the company running the power grid in Ukraine (Greenberg, 2017). It was speculated by ESET (ESET North America 2017) that strained relations with Russia caused Ukraine to be targeted with these attacks.

Given the difficulties of attribution, the actors behind the Petya and NotPetya attacks probably thought it would take months if not years for forensics teams to identify them. It also seemed highly unlikely that the related adversary state would prosecute them over other states' demands, even if they accepted the accusations.

To develop malware like NotPetya, developers need access to publicly unavailable vulnerabilities such as EternalBlue, and an analytical thinking process to involve catalysers such as M.E.Doc to multiply the inflicted damage on their target. The path to effect successful deterrence efforts forks at the ability of the deterring agent to understand the attackers' real intention. If their motivation is financial gain through mandating victims to pay ransom for decrypting their file system, the chances that they can be deterred is most likely small. Typically they represent a small hacker group, concealed behind a shroud of remote networks in multiple countries. There is no easy method for either giving them a clear deterrent message, or convincing them that the punishment will be far more costly than the possible gain because they have little to lose. However, if the actor is part of a government or closely tied to one, actions such

as a public statement detailing the capabilities of the perpetrator could convince the adversary to refrain from future actions. Deterrence-by-punishment seems to have some weight against a state actors. Yet, as both parties would be aware, the forensic activities to enable a valid attribution would take a good deal of time (Libicki, 2009), and it decays the effectiveness of deterring action. This leads to the conclusion that deterrence-by-denial methods are more promising against a NotPetya style of attack. As this method suggests, instead of adding an extra cost to attackers' attempts, the goal here should be cutting down their benefit.

Ukraine had a Cyber Incident Response Team and networks designed to defend against a cyber attack (Borys, 2017). The real damage caused by this incident came from elsewhere. Extensively used computer programs, much of which were comprised of third party software, had security vulnerabilities. This caused wide exposure to the many terminals these software running on. Programs that required user privileges to operate in particular (such as antivirus software) acted as the bottleneck in the overall security of their networks. They were invaluable to their operators, but they were only a single-layer barrier between the hackers and the vulnerable area behind them. Although the power outage was isolated to a specific area, there was breached linking point between the business network and the production network of the power grid of Ukraine. This resulted in serious negative effects on public psychology. The results could have been worse if not for an existing isolation mechanisms for ICS networks across the country (SANS ICS, 2016). The non-availability of bank services or a ministry has some effect on the public, but it is far from being comparable to a major power outage.

Among many other smaller-scale incidents, Danish shipping giant Maersk experienced the heaviest damage from NotPetya. The attack spread quickly in Maersk. Malware reached 45,000 PCs, 4,000 servers and 2,500 applications, effectively their whole infrastructure. It resulted in a 20% drop in volumes of their business operation (Chirgwin, 2018). According to Maersk's official statement, the damage to the company from NotPetya was collateral, and they tackled the situation by reverting back to manual systems, which can also be considered as a mitigation plan. The primary reason behind the devastating damage to their revenue appeared to be that their network's topology. It did not have practical layering and malware protection, thus once the malware was in, it was a matter of hours for the infection to spread. The second major cause appears to be there was no effective isolation mechanism between the business network at their corporate headquarters and the APM terminals. Maersk's port operators, which resulted in the infection of the APM terminals, went out of service (Gronholt-Pedersen, 2017).

Private companies do not have the luxury of enacting deterrence-by-punishment methods against a potential adversary. The best defence for private companies is a deterrence-by-denial method to protect against the damage inflicted by a NotPetya style attack. More sophisticated layering in the company's network topology and daily updates to the OS running on their business PCs (especially on the ones where the infection started) would stop the malware from spreading so quickly. Segregation of their networks would have provided an effective barrier to stop the malware from also infecting the port operator APM terminals which eliminated Maersk's ability to mitigate the effects of the malware and continue business functions.

6. Shamoon cyber incident

The Shamoon cyber attack was a unique and sophisticated malware attack targeting the computers running Microsoft Windows operating system with 32-bit NT kernels (Symantec, 2012). Shamoon stands out from other malware because of its highly destructive capabilities, the total damage it inflicted, its lack of financial motivation and the precise targeting of its attack. It was an act of cyberwarfare.

Once Shamoon reached a single computer in the targeted network, it copied itself into other terminals and started gathering files across different machines in the network, erasing them and, in the final stage, it overwrote the master boot record of the infected computer to make it unusable. Analyses (ICS-CERT 2012), particularly the ones inspecting the variants of the malware which resurfaced during late 2016, showed that the attackers had performed extensive preparatory work prior to the attack itself, including obtaining the credentials of some of the target company's (Saudi Aramco) workers. This information was embedded in the malware, but it is still unknown how they acquired the passwords (Symantec, 2016).

According to the immediate outcomes and the following analyses, (ICS-CERT 2012) Saudi oil company Aramco and Qatar's RasGas (Harper, 2012) were the only major targets that suffered significant losses. According to an online notice prior to the incident (Perlroth, 2012), a group called 'Cutting Sword of Justice' claimed responsibility for the act stating they were opposed to recent actions of the Saudi government. They targeted Aramco because they represented the largest financial source for the Al-Saud regime. On 15 August 2012, the master boot records of over 30,000 Aramco business computers were overwritten and failed to operate. All the data on those computers was overwritten with an image file depicting a US flag (Symantec, 2012). The company confirmed the attack on its social media accounts and stayed offline for 10 days. They issued a statement that their business was not affected but a leaked photograph showed hundreds of oil trucks parked and waiting idly near the company's oil stations due to its hampered business operations on 1 September (Pagliery, 2015).

Expert reports pointed towards the Iran as the most likely perpetrator of the Shamoon malware (Perlroth, 2012). Forensic analysis showed Shamoon attack was specifically targeted and, given the political atmosphere at the time. It was believed it was retaliation for the US and Israel's possible development of the Stuxnet attack. However, no Iranian official has accepted or affirmed the allegations.

In the aftermath of such a complex and damaging attack, the reasons why Saudi Arabia or Qatar were not successful in deterring have little to do with deterrence-by-punishment. The real aim of deterrence is to prevent the damage from being done to one's information infrastructure and operations. Prior to the Shamoon attack, the Saudi state had not stated their intent of reactive punishment against a possible cyber attack on its infrastructure. This essentially negates the essence of deterrence-by-punishment. One might argue at this stage that any possible state-level attacker probably had more information regarding the Saudis' offensive cyber capabilities than was publicly available, but the extreme complexity of the malware also suggests that the actors behind the attack were capable of covering their tracks. This would make attribution and the follow on retribution actives very difficult. Given these facts, deterrence-by-denial seems to have been a better tactic. This allows the reasons behind deterrence failure to become clearer.

Reports after the incident (Kubecka, 2015) show Saudi Aramco's business model dedicated a large portion of its security budget on protection of its ICS systems which were responsible for oil production and logistics operations. Top Aramco officials stated that their business computer network and the production computer networks were completely isolated from each other. This allowed the oil production facilities and industrial components to suffer no damage at all. This practice, which may have been a lesson learned from the Stuxnet incident, played a crucial role in their continuity of operation.

There were several reasons that the Shamoon attack caused significant damage to Saudi Aramco. Firstly, their risk assessment plan severely underestimated the cost of a temporary unavailability of their business network. (It is important to note here that Aramco held 10% of the world's entire oil production in 2012.) The case shows that, despite the isolation of their business network and production systems, there was still a relationship between them which clearly hampered production. Secondly; the malware was successful because it had the credentials of some corporate personnel with the authorization required for the initial penetration of the network. This suggests that there were shortcomings in the security of the passwords and authorization layering mechanism. There may have been an insider that provided information as well. Another major factor in the scale of the damage was there was not an auxiliary system which could act as a back-up. The costs of having a back-up system that could take over with a minimum delay of operations were not accurately calculated during their risk assessment.

From the deterrence perspective, Aramco did not have sufficient measures in place to effectively execute a deterrence-by-denial approach. Even though Aramco was a government sponsored corporation, deterrence-by-punishment methods from a private company do not seem feasible. And Aramco's deterrence posture did not benefit from their association with the host state.

7. Reality versus concept – deterrence fails

These two example show cyber attacks take place despite the deterring actor feeling assured that necessary measures are in place. The adversary does not perceive the associated costs outweigh the potential benefits.

In both cases there was a substantial lack of understanding about the adversary's (be it a state or a group) capabilities and real intentions. It is practically impossible to simulate the same cases with different factors and precise results. What we know is the adversary's actions had success in both cases. The extent of their ambition is unknown. No matter what their actual aim was, the results were at least partially successful.

In the NotPetya incident, the major shortcomings appear to have been vulnerabilities to third-party software, lack of layered and sophisticated topology in corporate and critical infrastructure networks, not applying security patches and not investing sufficient effort and attention into plans for resilience. Also the lack of isolation mechanisms to quarantine the intra-business and production network greatly enhanced the damage caused by the cyber attacks on the Ukrainian power grid and Maersk's intra-business and production network.

The Shamoon incident demonstrated there was no relationship between the size of the operation, such as Saudi Aramco, and their preparedness for a possible cyber attack, nor for cyber warfare considering it was a crucial component of the country's critical infrastructure. Additionally, poor isolation mechanisms inside its business network enabled the attackers to gain access to almost the entire network once they breached the first line of defence. The incident also highlighted the importance of user training and cyber hygiene as the attackers were able to use seized passwords and known vulnerabilities to access the companies network.

Cases of obvious cyber warfare, where the actors are distinguished and bilaterally or multilaterally known, deterrence-by-punishment methods are crucial. This response can be enhanced with the use of other instruments in other domains to defend against a cyber attack. Deterrence-by-denial works best by advertising capabilities. The aim is to convince the attacker of increased costs of a potential cyber attack as well as the probability of failure and associated risks. These effects can be achieved by building resilient cyber defence systems that includes all hardware, software, policy and human factors.

Giving precise percentages on how much weight the factors involved in these two real deterrence failure cases possess is not possible. Yet, it is not unfair to estimate that approximately 80% of the cumulative weight of the results of the entire results set caused by those major factors. Conversely, a few correctly implemented points among a wide array of procedures of cyber deterrence contain weigh slightly less than 20%. (Kaplow 2005) It is important to note for the sake of obtaining meaningful results that every single practice or action in cyber deterrence is unique in its own sense and environment. Thus, there must be a level of abstraction involved in this matter.

According to falsification theory (Popper, 1959), the actors in the two examples provided did not fully implement correct deterrence methods. Efforts in these two cases serves as indirect proof of their success claim. Had they applied their deterrence methods correctly and still failed to deter an actor, their attempt would be classified as a failure. Their deterrence efforts represent more than 80% of the negative outcomes the deterring actors tried to avoid. Deterrence methods have the greatest efficiency when the true motivation behind each actor's actions are known.

Once the most effective deterrence methods are identified which seem to have a major impact on maintaining a robust deterrence posture with respect to impracticality of attribution and other problems associated with deterrence-by-punishment; keeping these deterrence methods with the biggest impact for cyber defence is always desired. The main philosophy should be having an effective first line of

defence, accepting that there is a possibility it could be breached at any time. In addition, your deterrence posture should have fast-acting resilience mechanisms if your first line of defence gets breached (Jaatun, 2009).

These cases also show that demonstrating a credible impression that the defending actor has significant capabilities to punish the attacker does not successfully deter actors in the cyberspace domain. Hostile actions still occur. In other words, they represent the 80% side of the Pareto chart's cause (outcome) area.

Going into further details, we look at the first line of defence. Deterring parties need to protect credentials and any other items which can cause a vulnerability. They need to make clear and concise statements about the strength of their defensive cyber capabilities while abstracting the inner workings. Designating specific entry points and peripherals to their internal network is also useful in avoiding diverting their defensive efforts among numerous risks. Saudi Aramco, in particular, would have fared much better had they maintained a good first line of defence against cyber attacks and network penetration. The cost of such measures is relatively low. Lastly among the major factors, isolation between business and production networks, or incorporating closely controlled bridges if the former method is not possible, would keep the possible damage contained, and hence easier to manage (Jaatun, 2009).

The organizations' mind set must be breaches and attacks will keep occurring. Fast-acting resilience mechanisms need to be in place. The goal for organizations needs to be that, for a prospective cyber attack to be successful at least two consecutive failures must take place in the defending actor's deterrence posture. Furthermore, segregation of networks can prevent an infection from spreading rapidly. Requiring different authentication methods for different access levels can also make any attacker's job more difficult. A layered network topology, maintaining backup systems and hot-redundancy help continuity of normal operations, in addition to serving as an additional deterrence-by-denial. These measures can be advertised to the potential attackers with sufficient publicity by the nation and its public and private bodies.

These methods are proven factors in deterrence posture. They should be implemented into general and flexible policies. Organizations should accept the fact that minor or moderate scale cyber attacks will occur independent of the publicly known deterrence-by-punishment capabilities. The overall goal of the deterrence methods chosen by an organization should be to minimize loss. The cost and benefit of the deterrence methods should be constantly reviewed since they will dynamically change as each action is taken by an actor. The weight of a policy applied at any point of building and maintaining the deterrence posture fluctuates, as does the end result depending on the time frames in which these policies and actions occur. Therefore, every organization will evaluate their most appropriate first line of defence and resilience mechanisms based on the cost. This will be the bedrock for developing an overall cyber deterrence strategy and help deriving implication policies in this regard.

8. Conclusion

In the history of cyber deterrence, although there is no absolute way to measure this, a lot more has happened than has been deterred. Accepting the problems associated with this reality is a big step towards solving it. Starting from the Pareto principle and relying on falsification theory, major causes of deterrence failure were examined. The cases in which deterrence failed were also stripped out of the noise from other possible causes. This research suggests that the biggest positive change among all endeavours from nations in building their deterrence posture would come from these smaller, already-proven sets of actions. This refined subset of deterrence practices requires skilful attention and to be fitted to each nation's unique structure. This is a wise approach for using the resources of states and organisations in the best possible way, delivering the biggest possible impact in their deterrence posture.

Falsification theory was used during the analysis to eliminate the deterrence methods which failed to provide the desired effects. This smaller set of principles and the Pareto principle need tailoring for their target environment and the characteristics of each actor or entity in cyberspace before incorporating them into the overall deterrence mechanism. Otherwise, actors are again faced with a situation of fighting in an environment which contains too many assumptions and an exponentially bigger set of outcomes resulting from the decisions made on these assumptions. Defending actors in cyberspace always need to evaluate potential goals and motivations of an adversary (i.e. human factors, technical infrastructure) to adjust to different cases. Therefore, the expected consequences of hostile actions in cyberspace should be stated as clearly as possible by taking the distinct motivations of each actor into consideration. The required modifications for unique cases is a topic for another research. Additionally, accurate estimates for deterrence can only be made for rational adversaries. Actors in cyberspace cannot always deter irrational actors who do not place much importance on the cost of their cyber attacks and the possible outcomes. Albeit assumptions are a natural part of this intangible deterrence concept, minimizing the possible cost through spending limited resources on the most promising methods appears to be the best choice.

Unfortunately, complete deterrence in cyberspace will not exist. One does not have complete control over the adversaries' actions. The goal of a reliable deterrence posture is to minimize the harm or damage inflicted on defended systems and infrastructure. As discussed in this paper, under the umbrella of the actors' overall deterrence posture, they adopt an array of deterrence methods categorised as deter-by-punishment and deter-by-denial. This goal can be achieved by modifying deterrence efforts to the most efficient level in the continuously changing and flexible cyber environment.

9. References

Appleman, Roy Edgar. 2012. *United States Army in the Korean War. South to the Naktong, North to the Yalu.* Washington D.C.: Center of Military History.

Borys, Christian. 2017. *The day a mysterious cyber-attack crippled Ukraine.* 4 July. http://www.bbc.com/future/story/20170704-the-day-a-mysterious-cyber-attack-crippled-ukraine.

Chirgwin, Richard. 2018. *IT 'heroes' saved Maersk from NotPetya with ten-day reinstallation bliz.* 25 January. https://www.theregister.co.uk/2018/01/25/after_notpetya_maersk_replaced_everything/.

Cimpanu, Catalin. 2017. *BleepingComputer.* 6 July. https://www.bleepingcomputer.com/news/security/m-e-doc-software-was-backdoored-3-times-servers-left-without-updates-since-2013/.

CNBC. 2017. *There are 20 billion cyber attacks every day: Cisco.* 11 May. https://www.cnbc.com/video/2017/05/11/there-are-20-billion-cyber-attacks-every-day-cisco-.html.

Denning, Dorothy. 2016. “Cybersecurity's Next Phase: Cyber Deterrence.” *The Conversation*, 13 December.

ESET North America. 2017. *“Petya” Ransomware: What we know now.* 27 June. https://www.eset.com/us/about/newsroom/corporate-blog/petya-ransomware-what-we-know-now-3/.

Glenn, Marcus. 2017. *Failure of Nuclear Deterrence in the Cuban Missile Crisis.* Montgomery: Air War College, Air University.

Greenberg, Andy. 2017. *How an Entire Nation Became Russia's Test Lab for Cyberwar.* 2017 June. https://www.wired.com/story/russian-hackers-attack-ukraine/.

Gronholt-Pedersen, Jacob. 2017. *Maersk says global IT breakdown caused by cyber attack.* 27 June. https://www.reuters.com/article/us-cyber-attack-maersk/maersk-says-global-it-breakdown-caused-by-cyber-attack-idUSKBN19I1NO.

Grossman, Nadav. 2017. *EternalBlue – Everything There Is To Know.* 29 September. https://research.checkpoint.com/eternalblue-everything-know/.

Harper, Michael. 2012. *Energy Company RasGas Is Infected With Shamoon Virus.* 31 August. http://www.redorbit.com/news/technology/1112685657/shamoon-virus-rasgas-aramco-083112/.

Iasiello, Emilio. Spring 2014. “Is Cyber Deterrence an Illusory Course of Action?” *Journal of Strategic Security* 54-67.

ICS-CERT. 2012. *Joint Security Awareness Report (JSAR-12-241-01B) Shamoon/DistTrack Malware (Update B).* 16 October. https://ics-cert.us-cert.gov/jsar/JSAR-12-241-01B.

Kaplow, Louis. 2005. “Pareto Principle and Competing Principles.” *The Harvard John M. Olin Discussion Paper Series.*

Khrushchev, Nikita. 1976. *Khrushchev Remembers: The Last Testament.* Boston: Bantam.

Kubecka, Chris. 2015. “How to Implement IT Security After a Cyber Meltdown.” *[Slideshow].* 3 August. https://www.blackhat.com/docs/us-15/materials/us-15-Kubecka-How-To-Implement-IT-Security-After-A-Cyber-Meltdown.pdf.

Lebow, Richard Ned. 1981. *Between Peace and War: The Nature of International Crisis.* Baltimore: The Johns Hopkins University Press.

Libicki, Martin C. 2009. “Cyberdeterrence and Cyberwar.” In *Cyberdeterrence and Cyberwar*, by Martin C. Libicki, 27-37. Santa Monica, CA: RAND.

LogRhythm. 2017. *NotPetya Technical Analysis.* Boulder: July.

Lonsdale, David J. 02 February 2017. “Warfighting for Cyber Deterrence: a Strategic and Moral Imperative.” *Springer.*

Malware Tech. 2017. *Petya Ransomware Attack - What's Known.* 27 June. https://www.malwaretech.com/2017/06/petya-ransomware-attack-whats-known.html.

Martin Gilje Jaatun, Maria B Line, Tor Olav Grotan. 2009. “Secure Remote Access to Autonomous Safety Systems: A Good Practice Approach.” *International Journal of Autonomous and Adaptive Communications Systems Vol. 2 No. 3* 297-312.

Microsoft. 2017. *New ransomware, old techniques: Petya adds worm capabilities.* 27 June. https://cloudblogs.microsoft.com/microsoftsecure/2017/06/27/new-ransomware-old-techniques-petya-adds-worm-capabilities/?source=mmpc.

Mike Oppenheim, Steve Stone. 2017. *A ‘Wiper’ in Ransomware Clothing: Global Attacks Intended for Destruction Versus Financial Gain.* 29 June. https://securityintelligence.com/a-wiper-in-ransomware-clothing-global-attacks-intended-for-destruction-versus-financial-gain/.

2017. *New Petya / NotPetya / ExPetr ransomware outbreak.* 27 June. https://www.kaspersky.com/blog/new-ransomware-epidemics/17314/.

Nicole Perlroth, Mark Scott, Sheera Frenkel. 2017. *Cyberattack Hits Ukraine Then Spreads Internationally.* 27 June. https://www.nytimes.com/2017/06/27/technology/ransomware-hackers.html.

Nye, Joseph S. Winter 2011. “Nuclear Lessons for Cyber Security.” *Strategic Studies Quarterly* 18-38.

Nye., Joseph S. 2017. “Deterrence and Dissuasion in Cyberspace.” *International Security, President and Fellows of Harvard College and the Massachusetts Institute of Technology* 44-71.

Orme, John. Spring 1987. “Deterrence Failures: A Second Look.” *International Security* 96-124.

Pagliery, Jose. 2015. *The inside story of the biggest hack in history.* 5 August. http://money.cnn.com/2015/08/05/technology/aramco-hack/index.html.

Panikkar, K M. 1955. *In Two Chinas: Memoirs of a Diplomat.* London: Allen and Unwin.

Perlroth, Nicole. 2012. *In Cyberattack on Saudi Firm, U.S. Sees Iran Firing Back.* 23 October. http://www.nytimes.com/2012/10/24/business/global/cyberattack-on-saudi-oil-firm-disquiets-us.html.

Philbin, Michael J. 2013. *Cyber Deterrence: An Old Concept in a New Domain.* Carlisle, PA, USA: U.S. Army War College.

Popper, Karl. 1963. *Conjectures and Refutations: The Growth of Scientific Knowledge (2002 ed.).* London: Loutredge.

—. 1959. *The Logic of Scientific Discovery.* United Kingdom: Hutchinson & Co.

SANS ICS. 2016. *Analysis of the Cyber Attack on the Ukranian Power Grid.* Washingon DC: SANS.

Slayton, Rebecca. 2017. "Why Cyber Operations Do Not Always Favor the Offense." *International Security, Harvard Kennedy School*, February: 1-3.

Symantec. 2016. *Shamoon: Back from the dead and destructive as ever.* 30 November. https://www.symantec.com/connect/blogs/shamoon-back-dead-and-destructive-ever.

—. 2012. *The Shamoon Attacks.* 16 August. https://www.symantec.com/connect/blogs/shamoon-attacks.

Synmantec. 2017. *Petya ransomware outbreak: Here's what you need to know.* 24 October. https://www.symantec.com/blogs/threat-intelligence/petya-ransomware-wiper.

Timothy Naftali, Philip Zelikow. 2001. *The Presidential Recordings, John F. Kennedy, The Great Crisis Volume II.* New York: W.W. Norton and Company.

Whitting, Allen S. 1960. *China Crosses the Yalu: The Decision to Enter the Korean War.* Stanford: Stanford University Press.