MASTER THESIS

Titel der Master Thesis / Title of the Master's Thesis

„Defense is the Best Offense: A Realist Theory of Cyberwar "

verfasst von / submitted by

Gabriel Stewart. Wells

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Advanced International Studies (M.A.I.S.)

Wien 2017 / Vienna 2017

Studienkennzahl lt. Studienblatt Postgraduate programme code as it appears on the student record sheet:	A 992 940
Universitätslehrgang lt. Studienblatt Postgraduate programme as it appears on the student record sheet:	Internationale Studien / International Studies
Betreut von / Supervisor:	Professor Markus Kornprobst, PhD

diplomatische akademie wien
Vienna School of International Studies
École des Hautes Études Internationales de Vienne

(Page intentionally left blank)

Table of Contents

(Page intentionally left blank)

Abstract

This paper addresses the growing threat of cyberattacks on nation-states, the current lack of extensive literature which applies international relations theory to cyberattack cases, and an attempt to apply such theory to some of the available cases in hopes of analyzing how states respond when confronted with cyberattacks. Cyberattacks against states are on the rise, and the true threat level is not known. This paper begins with an explanation of the state of cybersecurity literature, and then proposes that a theory for explaining the responses of states after cyberattacks must be developed and tested. Following this, the IR schools of Liberalism, Constructivism and Realism are assessed for their potential explanatory power in regards to currently available cyberattack cases. A form of Defensive Realism is selected as the basis for developing a theory of cyberconflict, which will seek to explain the response behavior of states after cyberattacks. The theory predicts that, due to the inherent instability and anonymity of the cyber-sphere, states will tend toward defensive behavior when it comes to cyberconflict. The four cases which the theory will be applied to are: the Estonian attacks in 2007, the Stuxnet attack in 2010, the Shamoon Attacks, starting in 2012, and the Ukrainian Power Grid Hacks in 2015 and 2016. The conclusion of the analysis is that Defensive Relist application to cyberattack cases is mostly helpful in predicting and explaining state response behavior, but falls short in some areas, due to lack of access to information, inability to incorporate other possible variables such as national culture or history, and other factors. In the conclusion, shortcomings are discussed, and suggestions for further research in IR applied to cyber are proposed.

Abstrakt

Dieses Papier befasst sich mit den wachsenden Bedrohungen von Cyberattacken auf Nationalstaaten, die aktuellen Sorten umfangreicher Literaturstaaten reagieren, wenn sie mit Cyberattacken konfrontiert werden. Cyberattacken sind auf dem Vormarsch, und die wahre Bedrohungsstufe ist nicht bekannt. Diese Cybersecurity-Literatur, und schlägt dann eine Theorie für die Erklärung der Reaktionen der Staaten nach Cyberattacken müssen entwickelt und getestet werden. Danach werden die IR-Schulen des Liberalismus, des Konstruktivismus und des Realismus auf ihre potentielle Erklärungskraft in Bezug auf derzeit verfügbare Cyberattacke-Fälle hin untersucht. Eine Form des defensiven Realismus wird als Grundlage für die Entwicklung einer Theorie des Cyber-Konflikts ausgewählt, die das Antwortverhalten der Staaten nach Cyberattacken erklären wird. Die Theorie sagt voraus, dass die Staaten aufgrund der inhärenten Instabilität und Anonymität der Cyber-Sphäre in Richtung Verteidigungsverhalten neigen, wenn es um Cyber-Konflikte geht. Die Ergebnisse der Analyse sind die gleichen wie die Ergebnisse der Analyse der Ergebnisse der Analyse Defensive Relay Anwendung auf Cyberattacke ist in der Regel hilfreich bei der Vorhersage und Erklärung der staatlichen Reaktion Verhalten, aber wenn kurz in einigen Bereichen, aufgrund mangelnder Zugang Auf Informationen, Unfähigkeit, andere mögliche Variablen wie Nationalität oder Geschichte und andere Faktoren zu integrieren, In der Schlussfolgerung werden Mängel diskutiert.

Introduction

Cybersecurity is a wicked problem. It defies definitions, protocol, discourse, and international cooperation with its inherent instability and ambiguity. Cyber security is something of a wild west in the international arena. The very concept cannot be agreed upon between many nations. Tackling issues of cybersecurity at the national and international level is an immensely complicated undertaking, hindered by the fact that so little is known concretely about the true level of the threat. (Clemente 2011) The exact capabilities of cyber-actors are not concretely known, nor are the identities and whereabouts of these actors. Adding to this problem is the fact that cybersecurity itself is still quite new as a concept. Cyber as a serious vector for conflict between states and other actors has existed for hardly three decades. But as the digital age swiftly progresses, cyberattacks and cyberconflict are becoming more frequent, and understanding the nature of these attacks and their effects is critical.

While conventional forms of conflict between nation-states and their effects have been analyzed in depth for many decades, and this analysis has drawn on millennia of state interaction, cybersecurity and cyberconflict have existed for only a short time. Considering its guaranteed importance today and for the foreseeable future, cyberconflict and its effects must be observed and analyzed.

I believe the most important aspect of understanding cyberattacks is their ramifications for states and the international system. Specifically, it is paramount to understand how states respond to cyberattacks. As we will see, cyberattacks are fundamentally different in critical ways from conventional warfare or other previously established methods of negative state interaction, and therefore require a fresh look. Responses to cyberattacks are the most empirically available element of cyberattacks from a research standpoint, and therefore they will be chosen as the point of focus for this research. In this paper, I argue that the inherent anonymity of actors in cyber encourages states to act defensively when confronted with cyberconflict. The specifics of what this defensive behavior looks like and how it will be defined by International Relations theory will follow in the next chapter.

Some important terms and concepts will first be addressed: For the purposes of this paper, I will define a cyberattack as the following:

"Any type of offensive maneuver employed by nation-states, individuals, groups, or organizations that targets computer information systems, infrastructures, computer networks, and/or personal computer devices by various means of malicious acts usually originating from an anonymous source that either steals, alters, or destroys a specified target by hacking into a susceptible system" (Lin 2016)

While definitions of cyberattacks vary from author to author, the definition used here will encompass all malicious, intentional digital activity, and will not differentiate between cybercrime, cyber espionage, cyber terrorism, cyber conflict, cyberwar and other "cyber-isms".

It is important to attempt to understand how cyberattacks between states may translate into real-world outcomes and consequences, so we may better predict and plan for the future. This understanding can likely be enhanced and assisted by the use of some International Relations theory. Theory which seeks to explain the behavior of states and the tendencies of the international system needs to be applied to this relatively new field of international relations, in the hopes that it will further enhance our powers of prediction regarding cyberattacks and their consequences. The theory used in this thesis, which will be explained in depth later, will be mapped onto several cyberattack cases, and the explanatory power of the theory will be tested. We will be able to apply theory where it is relevant to the events, and assess whether it represents a best fit explanation for those events. Strong correlations between the claims and predictions of the theory and the actual events as they unfolded will signal strong explanatory power for that theory in explaining state and international behavior after cyberattacks. To clarify: the "data" being assessed will be limited to the responses of states and international bodies immediately or shortly following a cyberattack. This is because the attack behaviors are far less overt and available in comparison to the response behaviors. Rarely, if ever, do states announce that they will conduct a cyberattack. They do, as we will see, often announce their responses after an attack has been committed on them.

The literature surrounding the current array of cyberattacks is limited in comparison to earlier, more conventional topics and events of IR such as World War II or the fall of the Soviet Union. However, what has been written in regards to the descriptive observations and details of many cyberattacks will be covered in detail in the following section. Information on these attacks and responses to them comes from a variety of sources, primarily journalists, but includes a handful of books, scholarly articles, and the official statements of international organizations and nation-states, as well as leaks or the opinions of analysts and experts when no officially state or IO-sanctioned information can be found.

The literature also lacks works which address cyberattacks and state responses through IR theory. This may be due to several reasons: The phenomenon of cyberattacks in general is comparatively new when contrasted with the extensive history of conventional attacks and state responses that IR scholars have at their disposal. The very concept of a cyberattack is perhaps barely 30 years old, and therefore, there hasn't been sufficient time for a wealth of cases to develop, nor time for scholars of IR to apply preexisting IR theory to the events in cyber or to develop new theories to explain these events. In light of this, the literature is separated into two mostly disparate bubbles: the body of existing IR theory which references conventional, historical events, or is simply abstract theory on one end, and a wealth of policy papers, concrete analysis, or otherwise non-IR works on the other end. The space between these two bubbles, where IR theory is applied to cases in cyber, is largely unpopulated. It is my goal to help bridge this gap through this thesis. This will be done by addressing the available IR theory, assessing which strains and sub-strains will be most applicable, than adopting and developing the theory or combination of theories into a testable framework addressing the topic of cyberattacks and state responses.

The state of the literature will have unfortunate consequences on the possibility to understand and explain future cyberattacks. Without a strong theoretical base of literature to further dialogue and debate, cyber security will continue to be something of an ideological wild west. IR theory allows for better policy, better statecraft, and more informed decision making across all areas of foreign policy. The developments in cyber on an international level as a tool for states to influence and affect one another are growing

steadily. If we desire that leaders make more informed decisions in the short and long term in regards to cyber, it is imperative that IR discourse absorb cybersecurity as a field for present and future academic exploration. The gap in the literature can only be bridged in this way.

Through the following analysis, I argue that the prescriptions of Defensive Realism provide the most plausible and credible explanation for the responses to cyberattacks included in the analysis. Defensive Realism is the strand of Realism best suited for this task, and will provide the most explanatory power when assessing the outcomes of cyberattacks in terms of state responses. While Liberalism and Constructivism are less well-suited for application to cybersecurity, as their predictions do not gracefully apply to the newer realm of cyber, Defensive Realism will be shown to be a strong fit when seeking to explain the actions of states after cyberattacks, due to state's tendencies towards defensive behavior overall.

This thesis will be structured in four sections: one for each case. For each case, the context surrounding the cyberattack, the technical details of the attack and other relevant information will be presented. Alongside this, the responses of the relevant states in the wake of the attack will be described. Responses could be short or long-term. Actors disucssed will primarily be limited to nation-state governments, although information on non-state groups, international actors or IO's will occasionally be included. Responses included in the analysis will usually be limited to those which can be confidently determined as having been related to the cyberattack in question. This will often take the form of explicit statements on the part of state governments or leaders referencing the attack, but when responses are included which are not explicitly linked to the original cyberattack, a brief argument for their inclusion will be included. These responses will be compared to the likely predictions of Defensive Realism, and the results of this comparison will be discussed. If the theory's predictions for how states tend to behave closely match the events which transpired after a given cyberattack, the theory can be said to be a good fit for explaining cyberattack response behavior. Conversely, a set of outcomes where the actors in question did not behave in a way predicted by Defensive Realism would count as a weak case in favor

of the explanatory power of the theory. It must also be kept in mind, however, that cases which do not support the theory may not be reason to reject the theory wholesale, because, as Popper tells us, a theory which explains everything and leaves no possibility for counterevidence and falsifiability makes for a weak and unscientific theory. The accumulation of evidence in support of the theory, or otherwise, will be assessed after the four cases.

Chapter 1: Theory

Literature Review

When constructing the theory to be used in this thesis, the first decision was from which school of IR thought should the theory originate? I reject Constructivism and Liberalism as ideal baselines from which to build a theory of cyber response, and believe Realism to be a better fit for the type of analysis attempted here.

At its core, Constructivism is built on the concept that socially and historically constructed norms and rules of behavior guide state action, and that the role of natural human tendencies, or structural pressures, on state behavior is secondary to socially constructed practice (Wendt 1999). The constructivist school of thought strikes me as ill-fitting for a theory of cyberconflict, simply because so few actors and so little time have been involved in the history of cyberconflict thus far. There simply hasn't been enough time or interaction for social and historical norms to develop in the realm of cyber. It is something of a wild-west domain of international relations; therefore I believe it may be unproductive to search for socially or historically constructed values when it comes to the rules of behavior in cyber. We can already see evidence of this through the Stuxnet case. Little in the way of concrete norms of cyber had been established before the covert and insidious Stuxnet attack, and due to the secretive nature of the attack and its origins, and a lack of official acknowledgement on the part of the perpetrators, a discussion of norms and rules of behavior has not properly materialized.

One could say that pre-existing socially constructed norms color the behavior of states when it comes to cyber, but I find this a pithy concept, as the nature of cyberconflict

is simply too different from conventional, more overt foreign policy for the same norms to apply. Additionally, it could be argued that norms have already developed, which suggest a tendency toward covert attacks, the stealing of state secrets, and the targeting of critical infrastructure, but again, due to the extensive anonymity afforded to cyberattackers, and the low threat of public responsibility or discovery, states thus far have shown little qualms about their targets in the way that they typically do in conventional conflict: civilians, critical infrastructure, banks, and government offices have all been hit with cyber attacks. It seems to me that the norms of cyberconflict either do not yet exist or are at least, severely underdeveloped. Therefore, Constructivism seems an unlikely place to start if maximal explanatory power is the end goal.

Similar problems exist with the school of Liberalism when attempting to explain state interactions in cyber. Liberals believe in the role of international organizations, non-governmental organizations, and subsequent mutual international cooperation as a driving force behind state behavior, and reject the inevitability of power politics that Realism often suggests. (Shiraev 2014) I find this branch of IR theory to be ill-fitting for cyberconflict analysis for the same reason as Constructivism: namely, that cyberconflict is a "wild-west" domain. Legislation, definitions, treaties, and conventions dictating or prescribing state behavior in cyber are rather sparse, again, because the very existence of cyber as a domain of interaction is so new. This lack of formal discussion, let alone international agreement, would make Liberalism a weak choice for attempting to explain state behavior in cyber. The few UN, EU, and NATO treaties that have addressed cyber are minimal in comparison to the breadth of those that deal with historical, conventional international relations. The issue of anonymity and secrecy comes into play here: if states can engage in cyberconflict covertly and indirectly, there is little incentive to work toward mutual cooperation through international organizations, as there is no public pressure for states to be held accountable for their behavior. Another issue might be that technology simply evolves too quickly for the clunky, bureaucratic process of international deliberation and agreement to catch up. Up until now, international cooperation on cyber has been thin and sporadic, and formal

agreements underdeveloped. This makes Liberalism in IR a suboptimal choice for theory-construction.

This leaves the Realist school, which will provide the optimal ground on which to build a theory of cyberconflict. Realism is primarily focused on conflict, and seeking to explain that conflict between states, be it because of the inherent nature of humans, or the anarchic pressures of the international system. Political Realism makes a natural choice for a theory of cyberconflict due to its focus on how inherently conflict-oriented and Hobbesian its predictions of human behavior are. The type of power-seeking behavior that Realists identify as the greatest motivator for behavior in international relations be it on the part of individuals or entire nation-states, *increases* in situations where the actor's identity is shielded. Anonymity, afforded by cyber environments, allows actors to behave in accordance to their true nature and goals, unhindered by pressures from public recognition or responsibility. When the opportunity for mostly-anonymous power-maximization at the expense of rivals is presented, and can be achieved via cyber means, it seems likely that states will capitalize on these opportunities. As previously mentioned, the factors which are central to constructivist and liberalist IR analysis are sparse in cyberconflict, rendering Realism the more suitable school to adopt. Multiple differing schools of analysis exist within Realism, each with varying predictions and assumptions. Each of these schools must be addressed, and from them, a theory of cyberconflict can be selected, and subsequently applied to the cases. As Mearsheimer said: "The ultimate test of any theory is how well it explains events in the real world" (Mearsheimer 2001) This overview of the primary strands of realist IR will also serve as a literature review of the abstract theory "bubble" that this thesis will attempt to bridge with the literature "bubble" regarding applied cyber.

Some theories of Realism are better suited to assist in the explanation of foreign policy decisions, while others seek to explain international relations phenomenon. The four general strands addressed here will be Classical Realism, Offensive Realism, Defensive Realism, (both part of Neorealism), and Neoclassical Realism. Addressing the strengths and weaknesses of each strand of Realism will allow us to pick those best suited for cyberconflict analysis.

The oldest and most ambiguous strand is that of Classical Realism, which primarily originates from Hans Morgenthau's "Power Among Nations". Classical Realist analysis can be summed up in Morgenthau's relatively basic question: "How does this policy affect the power of the nation?" (Morgenthau 1948). Classical Realism looks at the actions of nation-states through the lens of power maximization, and addresses how and why states may have taken the actions they did in the goal of power maximization. On the international level, Morgenthau's idea of power interests as a unifying force for nations, or the "surest of bonds", provide an element for analysis. For example, the often-overlapping security goals of Ukraine with the EU and NATO In regards to Russia, or the similar security goals of Saudi Arabia and the United States in regards to Iran. These bonds can be identified through similarities in response behavior, explicit statements of mutual interest and cooperation, and by chronicling cooperation between nations as it appears in the wake of cyberattacks. (Morgenthau 1948)

Classical Realism has been developed extensively over the past several decades into the more nuanced and contemporary strands of Neo and Neoclassical Realism. Classical remains the most imprecise of the realist strands, and a precise definition of what prescriptions Classical Realism makes in regards to state behavior is unresolved. Later-developed schools of realist thought provide more concrete prescriptions for state behavior in the real world, and therefore, will be easier to test. The ambiguity of Classical Realism appears as a limitation to its ability to be tested, and therefore, a more nuanced strand seems preferable.

We can potentially draw our cybertheory from the two strands of Neorealism. While similar in many of their central assumptions, Offensive and Defensive Realism differ on a key assumption which leads to different explanations and predictions for international relations outcomes. My understanding of Offensive and Defensive Realism will come primarily from Jeffrey Taliaferro's article "Security Seeking Under Anarchy: Defensive Realism Revisited", as well as John J. Mearsheimer's "The Tragedy of Great Power Politics. The baseline assumptions relevant to this thesis for both Offensive and Defensive Realism analysis are as follows:

1. The state is the highest key actor, meaning there is no arbiter or authority above states to help resolve disputes. This leads to the inherent anarchy of the international system as described in offensive and Defensive Realism.
2. All states have some level of military capabilities. (In this case the assumption will be that all states have weaponized cyber capabilities)
3. States can never be sure of the intentions of other states in military (or cyber) matters.
4. States rank survival as their ultimate goal.
5. States tend to fear one another.

The causal chain in the lens of Offensive Realism link the inherently anarchic pressure of the international system and the uncertainty at the intentions of other states are translated into foreign policy. When operating from the perspective of Offensive Realism, it will be assumed that power ambition, crucial to survival in an anarchic system, decided on the part of rational state-actors, is the primary motivator for action within the global sphere, and that this action is more likely to lead to conflict and war between great powers as states continually move to reach a satisfactory level of security through advancements in security-enhancing power or resources. (Mearsheimer 2001).

Another central element of Neorealist IR theory is the security spiral phenomenon, posited by Neorealism. The security spiral or security dilemma is an artefact of the inherent uncertainty and anarchy of the neorealist framework, in which states continually move to heighten their security in response to the security actions of other states, which in turn creates tension and furthers the security measures employed by states. This cycle continues regardless of whether it is truly desired by the involved actors. Security dilemmas will no doubt be identifiable in an analysis of state and international cyber activity. Some core assumptions of the security dilemma as posited by a Neorealist framework which will be most relevant for an analysis of cyber activity are as follows:

1. Security Dilemmas are inherent to the anarchical international system. As states seek to “self-help”, they cause other states to feel less secure. This is one way security dilemma cycles can start.

2. Structural modifiers such as geography, access to resources influence the severity of security dilemmas.
3. Domestic politics can sometimes limit the effectiveness of severity of state responses. (Taliaferro 2001)

Defensive Realist analysis will follow largely similar baseline assumptions from Offensive Realism, as detailed previously, but will differ in how the assumptions affect the ultimate outcomes. They also put forth different predictions for the effects and nature of security dilemmas. While Offensive analysis will assume that states always prefer to act in ways that expand their power, Defensive realists assert that this is not necessarily always the case, and that the inherent anarchy of the international system does not by default mean that states will engage in power expansion behaviors, particularly in the form of short-term, military activities, but that they prefer to balance , and will rationally look to take actions which are more likely to promote greater economic prosperity in the long-run. Furthermore, the security dilemma need not always result in conflict, as cooperation is preferred. Indeed, cooperation is preferred by states, even nondemocratic ones, and IOs. (Taliaferro 2001) This is an important departure from Offensive Realism, and we will see what this framework can offer in terms of explaining the actions of states and the international system in the wake of cyberattacks.

Neorealism will likely be the most suitable strand of Realism for developing a theory of cyber. Much of what is possible in terms of cyberconflict is unknown. The inherent anarchy and uncertainty described by Neorealism is amplified further by the uncertainty and anarchy stemming from anonymity and lawlessness within the cyber domain. The hidden nature of the actors, the lack of information regarding true capabilities or possibilities for attacks, and the scarcity of explicit international agreement on cyber will surely increase the likelihood that states will behave in the ways prescribed by Neorealist IR theory. Between the two strands, Offensive and Defensive Realism, I believe Defensive Realism to be the most sensible choice for a theory of cyber. As will be seen in the cases, states prioritize defense and security above all else, even more so when it comes to defense in the cyber realm. Cybersecurity and cyberdefense have become the buzzwords they are

due to their critical and accelerating importance to overall national defense. Therefore, the main tenets of Defensive Realism will be adopted and put to the test by assessing real world state behavior in cyber events. The results will then be assessed in their proximity to the prescriptions of Defensive Realism. This will mean that identification of state behavior which is primarily defensive in nature, especially in the wake of a cyberattack, and offensive only when it is low-risk, will constitute support for the theory.

Finally, we can address the lens of Neoclassical Realism, and seek any worthwhile additions to the theory to boost its explanatory power. I will be drawing my definitions and assessment of Neoclassical Realism from the review article: "Neoclassical Realism and Theories of Foreign Policy" by Gideon Rose, and the review article "Taking Stock of Neoclassical Realism" by Shiping Tang. Neoclassical Realism also seeks to explain state behavior, and combines structural *and* domestic variables in its analysis. Relevant information analysis through this lens will be the grand strategies of the states in question, their military doctrines, and foreign economic activity. Per the theory, on the basic level, relative material capability determines the parameters of a nation's foreign policy. (Rose 1998) Another central element to the neoclassical school of thought will be the concept of the importance of the decision-making leaders in the process of the realization of foreign policy. The perceptions, beliefs, and opinions of the leaders who affect and ultimately authorize state action should be factored in when attempting to explain why a state took a certain course of action. It is ontologically impossible, of course, to know with total certainty what goes on in the mind of a nation-state's leader, therefore, we can only make educated guesses in regards to leader's perceptions, beliefs, and opinions. However, I believe we can know *generally* what the contents of leader's beliefs, perceptions, and biases are, because they often tell us explicitly what these are, and we can then corroborate their speech with their actions, and look for consistency. Ontological debate aside, the influence of the leader on state action is a fundamental idea in Neoclassical Realism. The Neoclassical approach in this thesis will be focused on how systemic incentives influence internal factors, which lead to foreign policy decisions. (Tang 2009) (Rose 1998)

Considering this definition of Neoclassical Realism, it may be suboptimal to include its methods and assumptions while developing a theory of cyberconflict. The central role of leader's perceptions, beliefs and biases are difficult to access even in more transparent domains of foreign policy, and these factors can only faithfully be included where they can be confidently shown to have had an influence on a foreign policy outcome. Within the secretive and covert domain of cybersecurity, however, meaningful and substantial evidence for leader's perceptions, motives and beliefs will prove prohibitively difficult to find. Much of the information regarding cybersecurity is top-secret or unconfirmed, and world leaders rarely, if ever, make their beliefs, motives, or other internal factors explicitly known. Not having access to this set of factors will limit the applicability of Neoclassical Realist analysis. Because of this, Neoclassical analysis which includes analyzing these factors will prove unsuitable for cyberconflict theory at this time.

Defensive Theory

Therefore, the theory that will be adopted for this analysis will largely follow the assumptions of Defensive Realism. The assumptions regarding state behavior, and its predictions for how states will act given these assumptions cast Defensive Realism as an optimal framework for seeking to explain the inherently anarchic and uncertain domain of cyberconflict. I will also add that states tend towards international cooperation behaviors in the immediate aftermath of cyber attacks. Defensive behavior in this analysis will be identified as instances of state behavior which emphasizes security against future attacks, support from allies in defending against further attacks, or simply not engaging in offensive retaliatory behavior. Support for the theory in the cases will be identified as a prevalence of defensive behavior, such as the creation of a cyber defense group, as well as engagement with international allies and organizations after attacks. Evidence not in support of the theory will be identified primarily in the form of retaliatory cyberattacks or other offensive measures, or a distancing from international cooperation. Instances of behavior where states place a premium on revenge and retaliation as opposed to defensive measures and longer-term stability will provide counter-examples to the defensive theory predictions.

Methodology

Even with the limited cases of the cybersecurity field, we can examine some patterns of response behavior by means of assessing primary and secondary sources, and identifying patterns across the case selection. The methodology which will be employed to address the multi-level responses to cyberattacks through a realist lens will be a case-study, highlighting four high-profile cyberattacks. The four cases selected address a diverse range of circumstances and patterns of response behavior. An overview of key technical details surrounding each case will be explored further in the empirical chapters alongside some important points of analysis.

This framework of four cases has been adopted with the intention of providing sufficient empirical depth and variety for testing the hypothesis. The inclusion of four cases of diverse circumstances allows for broader application of the theory, and therefore allows for more credible results. The more cyberattacks the theory is tested against, and the more factors it is faced with, the more robust the result of the analysis will beIf the theory is not well-suited to explain outcomes accurately, it can be said to have been tested more thoroughly. If the theory can explain outcomes, it will be more thoroughly so, having been applied to a range of cases and circumstances. The case study approach is the preferable way to test IR theory in cyber. It allows us to easily and clearly generate predictions based on the theory and what it purports, then test those predictions against the facts of each case (what really happened) this will allow for straightforward appraisal of the theory and its predictions. A brief introduction to the four cases follows.

In April 2007, Estonia was hit with a wave of cyberattacks directed at websites of newspapers, broadcasters, the Parliament, various ministries, and banking infrastructure. The attacks were primarily distributed denial of service (DDoS) attacks, which utilized many botnet systems. The DDoS attacks were supplemented by the defacing and spamming of several other websites of news outlets and a political party. The attacks took place amidst the controversial relocation of a Soviet war memorial in Estonia. Stephen Herzog, writing for the Journal of Strategic Security, notes that this attack was likely perpetuated by Russian diaspora, in conjunction with Russian minorities within Estonia, who felt disenfranchised by

the removal of the Bronze Solider statue, viewing it as an attack on their identity (Herzog 2011). The Estonian government holds The Kremlin responsible for the attack, although evidence conclusively determining Russian involvement is scarce and the Russian state has officially denied any responsibility. The attack was one of the largest of its scale ever seen, and caused significant economic damage to the highly-connected Estonian state. Estonian official responses included various increases in cyber security measures, a further souring of relations with Moscow, the creation of a volunteer cyber security force, and the cooperation with NATO, the United States, and other international actors for the purposes of cyber-securitization. (Gjelten 2011)

Initially identified in 2010, a targeted cyberattack was executed on the Iranian nuclear program by means of what would be termed the Stuxnet virus. Stuxnet is deemed to be one of, if not the most sophisticated malicious pieces of computer code ever developed. The virus was developed with a rootkit component, rendering it nearly invisible, and although no state or institution has officially taken credit for developing the virus, it is widely believed to have been developed as a joint US-Israeli project during the Bush administration. (Gibney 2016) Official Iranian responses varied, ranging from a downplaying of the damage caused by the virus, to different levels of admission that damage had been caused by the virus, and that a large cyber team had been assembled to combat and clean up the virus. Various changes in the Iranian nuclear program, including the capacity and number of centrifuges in the country, also occurred.

The state-owned Saudi Aramco oil company was hit by a cyberattack in August 2012, which was purportedly able to inflict 30,000 Aramco workstations. The malware employed, referred to as Shamoon, erased data on infected machines, replacing the files deleted with a jpg, and rendering them unusable. The first wave of Shamoon attacks used a picture of a burning US flag, while a late 2016 resurgence of Shamoon attacks on the General Authority of Civil Aviation, and other important Saudi infrastructure targets utilized the infamous photo of the three-year-old Syrian boy Alan Kurdi. Although a group named “Cutting Sword of Justice” claimed the attack, it is believed to have been the doing of Iran in retaliation for the Stuxnet virus cyberattack. (Gibney 2016) The company claimed that no oil production

had been inhibited by the attack. No individual or group has thus far claimed responsibility for the second wave of Shamoon attacks, but the Saudi government and businesses are on high alert for more attacks of this kind. (Seals 2016)

In December of 2015, a massive cyberattack was carried out on a power grid in Ukraine. Multiple energy companies were affected in the attack, and nearly 230,000 people were without electricity for several hours. (Aitel 2016) The foreign attacker was able to remotely control and disable SCADA systems at a power plant in Western Ukraine. (Sans ISC 2016) Reuters wrote that the "Ukrainian regional power company Prykarpattyaoblenergo reported an outage, saying the area affected included the regional capital Ivano-Frankivsk." (Zinets 2017) This attack marked the first known successful cyberattack on a power grid. Like other high-profile attacks, no state or other actor officially claimed responsibility. Russian IP addresses were purportedly identified as origins of some of the attacks, and the Russian Federation is widely believed by many Western analysts and security experts to have orchestrated the attack, although the Kremlin denies any involvement. According to Reuters, "Oleksandr Tkachuk, Ukraine's security service chief of staff, said at a press conference that the attacks were orchestrated by the Russian security service with help from private software firms and criminal hackers…" (2017) This was in response to a second attack on Ukrainian infrastructure, wherein a "power distribution station near Kiev unexpectedly switched off early on Sunday, leaving the northern part of the capital without electricity". These two series of attacks were part of the on-going Russian-Ukrainian conflict.

Chapter 2: Estonian Infrastructure Attack (The Bronze Soldier)

The Estonian infrastructure attack provides overall solid support for the defensive realist theory of cyberconflict. Estonia responded to an unprecedented and significant cyberattack on its institutions and infrastructure in a way that eschewed offensive tactics, and relied solely on bolstering its own defenses and coordinating those defensive measures with its close allies. The extent to which Estonia relied heavily on the international system in

its response also provides support for the second prediction of the theory, which is that states tend toward international assistance and cooperation following cyberattacks.

The earliest of the four cases and one of the most transformative in terms of its international ramifications is the series of Estonian infrastructure attacks which took place in late April, 2007. The Estonian government had announced preparations to relocate a Soviet war memorial, The Bronze Soldier, to a less central location within the Estonian capital of Tallinn. The statue was initially unveiled in September, 1947 to mark the victory of the Red Army over the Nazis. While the statue was initially located in a park in central Tallinn, it was moved to the Tallinn military cemetery. As many other Balkan and Eastern European states have done over the years, Estonia was, through the removal of the statue, seeking to shed its Soviet past. The proposed relocation sparked massive controversy. The monument had been a socially divisive structure for some time, symbolizing Soviet occupation and oppression to many ethnic Estonians, and a constant reminder of their unlawful occupation of the country post-WWII. For the Russian minority population of Estonia, the Bronze Soldier was something of a symbol of pride, and a reminder of their rights within the country. This deep cleavage within Estonia came to a climax on April 26th, when rioting broke out across the capital. Russian ethnic minorities were outraged at this perceived attack on their identity. After the first night of riots, the government moved to transfer the statue immediately. Overall, over 1000 people were arrested, one Russian rioter was killed, and Tallinn suffered widespread arson and property damage. Concurrently, on April 27th, large swaths of Estonian infrastructure were hit with an unprecedented series of cyberattacks. Using DDoS attacks and a swarm of botnet machines, the still-indeterminate number of hackers assaulted websites and digital infrastructure of many Estonian businesses and groups, including several large banks, the website of the Reform Party of Prime Minister Andrus Ansip, several newspapers, the parliament, and other government branches. The effects of the attack were a general loss of service and functionality for two weeks, and a halt on most bank activity in the highly-connected country. Users were not able to use their bank accounts, and activity to newspapers and government websites was redirected to a varying set of pictures and quotes mocking or

threatening the Estonian government and decrying its actions. About $750 million was projected to be lost during the period of the attacks.

After the initial shock, and the eventual recovery of the afflicted sites, with the help of Finnish, German, Slovene and Israeli technical assistance, the Estonian state took a series of actions domestically and internationally in response to the attacks. One of the first things was to issue blame at Russia, for either perpetuating or assisting in the execution of the wave of cyberattacks. The Estonian Defense Minister illustrated that successful cyberattacks could be equated to conventional maritime attacks; essentially that Estonian "cyber ports" had been blockaded. (Herzog 2011) The outrage of the Russian minority population of Estonia, (fueled from abroad) the chaos of the Bronze Nights riots, and the barrage of cyberattacks were labeled as examples or extensions of Russia's well-known campaigns of influence, information warfare, and power projection on its smaller neighbor. (Herzog 2011) Estonian officials pointed to the hostile tone the Russian state media and politicians directed toward the Estonian state and ethnic Estonians, and their attempts to accentuate rumors of ethnic clashes, hatred, and oppression of Russian-speaking populations within the country. It is also worth mentioning the Russian state's history of encouraging and organizing patriotic hackers. The FSB would have no doubt been aware of such a large-scale disturbance originating from Russian cyberspace. This, coupled with the synchronized beginning and ending of the attacks, points to the likelihood that Russian government resources and coordination were utilized.

Domestically, Estonia took steps to secure itself against such attacks in the future, and has established itself as a model for how to defend against cyber threats with the formation of the Estonian Cyber Defense League: a team of cybersecurity experts, programmers, software engineers, and computer scientists who form a volunteer cyber army, which would operate under military command in the case of an emergency. The Estonian state may even implement a cyber conscription service for its cyber Defense League, to ensure that valuable cyber skills and abilities are available if an emergency arises. (Gjelten 2011)

The immediate actions of the European Union and NATO were to examine the claims of Russian involvement. A Russian cyberattack on a NATO member would constitute a serious and dramatic response from the other members of the alliance. Although neither the EU nor NATO have been able to find any concrete evidence of Russian state involvement, the likelihood of Russian involvement via proxy actors, especially in the realm of cyber, where anonymity remains an alluring factor, is not impossible to rule out. (Herzog 2011) As mentioned before, assistance from Estonian allies in the region was swift in order to restore operations to the country's infrastructure. The EU and NATO engaged in significant intelligence sharing at this point regarding the likely perpetrator(s) of the attacks and the damage caused. What followed was a significant reconsideration of cyber security on the part of the EU, NATO, and their respective members. The European Union released its revised Internal Security Strategy in November 2010, which called for integrated cyber security responses from its member states. (Herzog 2011) NATO, as part of the Bucharest summit in 2008, adopted its Policy on Cyber Defense, and created the Cyber Management Authority. In a move which placed further emphasis on this recommitment to cyber security, stationed the new headquarters of the NATO Cooperative Cyber Defense Centre of Excellence in Tallinn. While this facility was already planned before the attack, its importance was only emphasized, and NATO's commitment to the plan intensified. (Herzog 2011)

It is also worth noting the immediate actions of Russia during and after the Estonian cyberattacks. In addition to providing vocal support for the Russian diaspora living in Estonia, and their professions of solidarity with the minority, the Kremlin was quick to label the arrests and injuries of the rioters as ethnic oppression, and human rights violations against Russians, utilizing the term "Russophobia". They further demanded the resignation of the Estonian Prime Minister in the aftermath of the Bronze Night. (Herzog 2011) Further down the timeline, some experts have pointed out that Russia also took measures to increase its cyber security defenses by directing resources and manpower. China is also reported to have followed suit. Both countries seemingly perceived a change in the wind and responded accordingly. (Herzog 2011) Although Estonia as a NATO member could have

swiftly dragged its strategic allies into a conflict, retaliating against the Russians, it did not engage in a tit-for-tat power politics strategy. While Russia's power politics cyberattack could have been utilized as justification for such a response, Estonia instead moved to open international discussion and cooperation with its allies within the EU and NATO.

The responses of the Estonian state overall support the claim that states in cyber do not engage in offensive responses against their rivals, and tend to favor defensive, security measures. They did not respond with an aggressive power politics strategy, appealing to their NATO allies in a counter attack against Russia, but instead sought to engage in defensive strategies domestically and through international influence because of the attack. Estonia, by opening the debate on cyber cooperation in the West and among NATO, and leading with its example of the potential future attacks could hold, moved to take an important role in leadership on cyber security. Its creation of the Cyber Defense League as a direct response to the attack provides additional evidence for the defensive nature of their response to this attack. This new resource will readily be deployed in the future for any Estonian cyberattacks. Estonia acted rather prudently, and worked toward the inclusion of NATO and the EU in the formation of new cyber discussions and cooperation which further isolated Russia.

Estonia responded to the external threats against its infrastructure, along with its internal factors of uneasiness and turmoil (Estonian interests against Russian outrage) by seeking help from other ally states, who share a common rival. Realism tells us that the strongest determinant of shared interests between states is the existence of a common enemy. (Tang 2009) NATO and the EU offered their support and resources, further isolating Russia on an international level. Estonia's importance as a leader in the West in cyber defense was solidified with the creation of the CCDCE. While we cannot always adopt a neoclassical lens and include the perceptions and beliefs of individual actors, as they are difficult to obtain with certainty in many cases, in the Estonia case, it seems illogical to disregard their evidence as influential factors in Estonian foreign policy. Estonian president Toomas Hendrick Ilves made clear some of his beliefs and perceptions with his recent call to NATO, in which he urged Estonia's allies to send troops to the Estonian border, claiming

that the treat of a Russian incursion is real indeed. (Blair 2015) Ilves' perceptions of Russia as a grave threat and the enemy of the West no doubt influenced his decisions in regards to Estonia's responses after the 2007 attack. The Cyber Defense League and the cooperation and discussion with NATO were at least partly influenced by Ilves' perceptions and beliefs, as well as the external factors of Russian threats. Ilves represents a widespread opinion and historical perception present in Estonia, which considers Russia as the successor to Soviet annexation, domination, and oppression. This framework colors the ways in which Estonia engages with Russia, and it colors the entirety of dialogue between leaders of the two states. This is one perception whose importance in Estonian foreign policy decisions cannot be ignored, and it almost certainly played a role in the aftermath of this attack.

Through NATO and Western dialogue, Estonia set off a wave of discussion and reform that changed the environment of cyber security and cyber defense in its own country and abroad, while Estonia itself is placed at the forefront of the most cyber-advanced and well-defended Western states. Additionally, the magnitude and ambition of Estonia's action was limited by their relative material power resources. Being a small country of barely 1.23 million people, Estonia cannot attempt to change its environment or exert control with its foreign policy with the ambition that a larger nation such as the United States could. Therefore, it acted in a way relative to its power capabilities: establishing a domestic force which seeks to utilize the best domestic resources (personnel) possible, and enlist the assistance and cooperation of the allies within its Western sphere.

We can see that Estonia did in fact, *not* engage in tit-for-tat, aggressive, or offensive responses against Russia, nor did it try to expand in the region because of the attack. There are some pragmatic factors as to why this is, namely that the country is simply too small in relation to its massively powerful neighbor for any direct unilateral response to be meaningful or plausible. Instead, it worked with the international systems in place to raise its security to further attacks, and establish institutions alongside international institutions which will likely deter future attacks from its enemies. Even though Estonia could have engaged in a counter attack or sought to expand utilizing the military power available to it through its NATO allies, it did not and instead used international dialogue to establish

defensive measures for the future. In this series of events, we see the lack of an aggressive, offensive response, when one could have certainly been possible through NATO alliances.

In order to see support of our defensive theory of cyberconflict, we can assess the scenario by asking the question "Did Estonia engage in "self-help" behaviors within the anarchic international system, which resulted in a possible security spiral?" The underlying primary assumptions of Neorealism that hold true in the Estonia cyberattack case: The system is anarchic and unpredictable, states prioritize survival above all else, and seek defensive maneuvers most of the time over offensive maneuvers. Our theory also says that states engage in "self-help" because of the inherent chaos and uncertainty of the international system and in doing so, they tend to create security spirals. Although the answer to whether Estonia engaged in behavior which created or facilitated a security spiral seems to be no, Estonia did proceed in an almost exclusively "defensive" capacity, particularly in the creation of its Cyber Defense League. It did not engage in extensive "self-help" behaviors, which are actions likely to cause fear in a state's neighbors. Estonia worked with the international system to great extent, utilizing the cyberattack from Russia to highlight the importance of cyber defense to its allies, to establish a cyberbase of sorts in its capital, and largely avoided individual power politics moves, or self-help measures. Our theory would additionally predict that states are not so short-term and reactionary as Offensive Realism would suggest. Estonia's lack of tit-for-tat response shows that they had the long-term in mind, and future economic and political prosperity seem more likely the driving motivation for their actions, rather than short-term military preparedness and retaliation. We can see in this case significant support for the predictions of our Defensive cyberconflict theory: the creation of an exclusively defensive cyberprotection agency (CDL), and the coordination with allies for defensive measures. Possible criticism here could include the fact that Estonia's behavior was largely guided by, and directed toward, the international organizations it is a part of. This coordination suggests support for a liberalist theory of cyberconflict, that when a state is under threat from cyberconflict, it will engage in international solutions. However, Estonia's size was likely the most important factor in its

decisions to defend itself: it requires the protection and resources of its international allies in order to stave off the threat from its massive neighbor, Russia.

Chapter 3: Iranian Centrifuge Attack (Stuxnet)

Stuxnet provides lower evidence than the previous case for the defensive realist theory of cyber. After the identity of the attacker was revealed to the world, Iran engaged in primarily offensive behavior as a state, favoring retaliation of similar magnitude against rivals domestic and abroad. The prevalence of offensive responses and the lack of international cooperation after the attack overshadow any instances of defensive behavior. Multiple factors not contained within the scope of the theory could be the cause of this difference.

Much has been written about this, perhaps the most important and consequential cyberattack perpetuated. Common descriptions of the Stuxnet cyberattack include "unprecedented", "revolutionary", or "paradigm-shifting". The extent to which these descriptors accurately reflect reality is still debated, but for this analysis, the context and details of the event require unpacking. It is widely accepted that the Stuxnet virus, named by the Symantec technicians who first analyzed the malware's code, was the result of a joint US-Israeli operation titled "Olympic Games". (Gibney 2016) The goal of the malware was to halt, delay, or otherwise damage Iran's nuclear facilities, and stifle their capabilities. This attack differs from the previous case in that the level of sophistication far exceeds that of the Bronze Soldier attacks. While the Estonian attacks used relatively standard and well-understood methods of cyber harassment, (DDoS and botnets), Stuxnet was a precision-crafted piece of software with hyper-specific purpose and function, a function which it executed spectacularly. Stuxnet worked through a combination of two functions once it had infected the Iranian SCADA systems controlling the centrifuges: It would simultaneously cause destructive alterations in the rotational speed of the centrifuges, causing them to be destroyed or severely damaged, while displaying false data to the scientists in the Natanz lab indicating that there was nothing wrong. (Gibney 2016) This strategy of destruction through software, combined with obfuscation of data, was unprecedented in cyber. In fact,

the only way it was discovered was its eventual spread to other machines outside of the Natanz lab, which triggered other groups to investigate the malware and thereby unearth its origins. (Gibney 2016) Some speculated that the reason the virus was eventually discovered was because it had been modified several times in successive waves, and new, more aggressive functions had been added, causing it to spread to unintended machines, and catch unwanted attention.

The question of effectiveness of Stuxnet as an attack does not have a clear answer. While serious damage was indeed caused to the Iranian centrifuges, the lasting damage, or the true costs of the initial damage are hard to measure. Some US officials claim the Iranian program was set back by 5 years, others think that any time delay is closer to one year. Once the Iranians discovered that it was outside interference that was causing their centrifuges to fail, and that they could rule out technical deficiencies in the equipment, or incompetence on the parts of the scientists and technicians, they ramped up production, attempting to compensate for the damage caused initially. Now the Iranian nuclear facilities are larger than ever. Some argue that the discovery of the Stuxnet virus only further emboldened the Iranians to ramp up their nuclear program. (Gibney 2016) Iran claims that no lasting damage or hindrance to their program occurred because of Stuxnet.

The responses of the Iranians covered the breadth of implicit and explicit, overt and covert (or semi-covert). Similar to the Estonian response of the creation of the Cyber Defense League, Iran founded its own cybersecurity force, the Iranian Cyber Army, a group which pledges its loyalty to the Supreme Leader of Iran, but which is not officially recognized by the Iranian government. (Reuters 2013) As with Estonia, the group is likely comprised of Iranian cybersecurity experts, civilian computer scientists, hackers, and others, but differs in that it is not an overt government organization, but a covert hacker collective only known through its claimed responsibility of hacks and attacks abroad. (Eisenstadt 2017) The Iranian response included public statements and threats in the wake of Stuxnet's discovery, including Khameni's declaration that "The United States will be taught the mother of all lessons." (Business Insider 2013) Iranian spending for cyber defense increased 1,200%, and their nuclear capabilities, as mentioned, have climbed to historic highs. The

Atlantic Council, in a 2015 report, notes that Iran is no longer a Tier 3 cyber power. (Sen 2015)

The most significant responses, and most pertinent to this analysis, are the handful of recorded cyberattacks against Iran's enemies, all of which followed the revelation of Stuxnet and its origins. Attacks on Capital One, Chase, and Bank of America occurred. (Eisenstadt 2017) Similar attacks on Las Vegas Sands Corp. and Qatari RasGas occurred in 2014. (Capaccio, Lerman and Strohm 2015) Security analysts note that Iran is a power which, like Russia, tends to rely on proxy groups for its attacks, be they cyber or political. Russia utilizing its patriotic diaspora, and Iran employing the Iranian Cyber Army, or the self-proclaimed "Cutting Sword of Justice" (Perlroth 2012) Iran has developed a very strong cyber presence in the region, and its use of ambiguous proxy groups seems to have emboldened its responses against its enemies. Apart from the attacks on Qatar and the United States, Iranian proxy groups conducted another large-scale attack, which will be the subject of the following case. Although it will be explored in detail then, the basics will be examined for analysis of Iranian responses. The attack on Saudi Aramco was conducted two years after Stuxnet, and was conducted via a malware which wiped the hard drives of some 30,000 machines at Saudi Aramco. (BBC Technology 2012) The files were all replaced with images of burning American flags, with the Iranian proxy group "Cutting Sword of Justice" claiming responsibility. (Perlroth 2012)It is widely believed that the cyberattack was in direct retaliation for Stuxnet, and a reason for its target, a key US ally in the region. (Fitzsimmons 2012)

Since the election of President Rouhani, and the subsequent conciliatory nature of Iran's recent interactions with the international community, a decline in the severity of Iran's cyber and nuclear capabilities and desires is clear, but with shifting administrations in both spheres, this has the potential to change soon.

Numerous Western and international security experts and analysts have characterized the strategies and preferences of the Iranians regarding their foreign policy and covert and overt responses, conventional or cyber-based. Iran is frequently accused of sponsoring and supporting foreign terrorist groups whose ideology or interests align with

those of Iran and its hardline Islamic government. Iran's foreign actions are often characterized by subterfuge, indirection, ambiguity, and proxy groups, which it supplies, advises, trains, or funds. (Eisenstadt 2017) Hamas, Hezbollah and Al-Qaeda are among the groups provided as examples of those supported by Iran in the past. This indirect connection is also identifiable in the way Iran has conducted its responses to the Stuxnet cyberattack. The numerous retaliatory attacks attributed to Iran in the years following Stuxnet were conducted through ambiguous or anonymous proxy hacking groups, which the Iranian government claims it does not support. This maximum-profit, minimal-risk approach is in line with the established character of Iran's foreign policy in other areas: favoring proxy groups and covert action. Iran is pragmatic in its responses to threats in cyber, and its responses to its adversaries were marked by prudence and calculated assessment of risk. Any overt attack or one that could be easily attributed to the Iranian government would meet strong international condemnation, and likely illicit retaliation from Iran's much-stronger adversaries. If indeed the Iranian government supports the Iranian Hacking Army, which is likely, it would be a sensible and low-exposure way of projecting power in the region through targeted, ambiguous cyberattacks against rival powers. The increased cyberdefense spending, invested personnel and resources into cyber programs, as well as the ramping up of nuclear production are all among the observable Iranian responses. Iran sought to show that it will not be intimidated or bullied by larger powers, nor will its nuclear program be stopped, and responded in kind.

The verbal threats to teach the US "the mother of all lessons" are another element of Iranian response strategy. These kinds of threats, particularly when cyberattacks are in mind, rouse popular support domestically, contribute to a sense of revenge against the nation's enemies, and thereby spurn further civilian support for Iranian developments in cyber, as well as the under-fire nuclear program. This has taken the form of protests declaring the "human right" of nuclear power in Iran must be guaranteed. (Gibney 2016)

Equipped with our cybertheory of Defensive Realism, we can assess the attacks and other foreign moves of the Iranian state and see how these predictions hold up. Iran engaged in numerous retaliatory attacks in the immediate years following the revelation of

the Stuxnet attack and its perpetrators. Significant cyberattacks are largely believed to have been conducted on American, Qatari, and Saudi targets, foreshadowed with the Supreme Leader's ominous threat of retaliation. Iran's actions can be interpreted as a manifestation of several Offensive Realist prescriptions. Iran, knowing its best chance for survival and seeking to ensure that survival, was to strengthen its cyber capabilities, likely assisted in the creation and support of the Iranian Hacking Army. This move echoes the creation of the Cyber Defense League by Estonia after it suffered a serious attack; an organized cyber army will assist in protection and even deterrence against further cyberattacks. Differently from Estonia, Iran behaves in a way which appears to provide evidence for the prescriptions of Offensive Realism, as opposed to Defensive: Its political strategy and its attempts to strengthen its capabilities in both cyber, nuclear, and conventional means point to its desire to become a regional hegemon, and its primary rivals in the region, Israel and Saudi Arabia, are the main impediments to this goal. Iran's actions in the region, cyber or otherwise, can be seen as manifestations of their desire to prevent either of these rival powers from obtaining regional hegemony. The series of Iranian cyberattacks following Stuxnet against targets in the United States and its ally nations point to the state behavior put forth by Offensive Realism; Iran is relying on aggressive and offensive tactics to increase its power in cyber, and ideally establish itself as a more powerful regional nation. Engaging in constant security competition "with war in the background" via cyber warfare is a recognizable feature of the Iranian response. (Mearsheimer 2001)

Defensive Realist predictions point to the tendency of states to seek long-term economic gain over short-term victories, and we can see that at least in the immediate several years after Stuxnet, Iran did not behave in this way. Defensive Realism would also suggest that states do not always engage in aggressive, offensive tactics, when possible, nor do they engage in tit-for-tat retaliatory measures, instead preferring long-term security. Iran in the wake of Stuxnet does not seem to have taken this kind of measured or defensive approach, instead focusing on executing many retaliatory, aggressive cyberattacks against its rivals and those it deems responsible for the Stuxnet attack. Defensive Realism, also claiming that states engage in "self-help" to defend themselves in the anarchic system,

shows some explanatory power in this scenario. The massive increase in cyber research and resources post-Stuxnet is an entirely appropriate response. Iranian security increases as a result of Stuxnet can be described as the kind of self-help behavior outlined by Defensive Realism: working towards security and defense unilaterally, and outside the bounds of any international cooperation or agreement. A variety of security spiral is potentially at work here; as the Iranians have ramped up their cyber capabilities, so too have their rivals, in what many have called a "cyber-arms race" in the "post-Stuxnet world". (Gibney 2016) Iranian actions do make sense when including additional views of Defensive Realism regarding the creation of strong central institutions. Iran has very strong and wealthy regional neighbors, some of whom are its fierce rivals (Saudi Arabia, Israel). This, combined with its relatively weak borders, and the general instability of the region, contributes to the creation of its strong military ambitions, and likely to the creation of the Iranian Hacking Army.

The support for a theory of Defensive Realism in Cyber appears mixed in the Stuxnet case. Iranian responses were comprised primarily of offensive maneuvers, counter-attacks, and support for proxy hacker groups. Iran's preference in this case for short-term retaliation against its rivals superseded its interest in primarily defensive measures. This is not to say that there is no support at all for defensive realist cybertheory: Iranian spending and resource allocation for cyberdefense and research increased significantly, but the primacy of offensive, short-term tactics means that evidence for an offensive theory is more present than for a defensive one. There are many possible factors that could influence this which are not necessarily captured by a theory of Defensive Realism.

The grand strategy of the Iranian Islamic Republic has been consistent throughout the decades since its foundation. The cultural and political identity of the state pits its interests against the West, and particularly the United States and its allies. The consistency of Iranian ideology and attitude toward the United States, and the West in general can be identified in the frequent and explicit statements made by Iranian leaders and government figures. Until recently, with the election of the moderate president, Hassan Rouhani, Iranian political and cultural rhetoric towards the West and Israel has been as toxic as any

discussion in international relations in the modern era. Supreme Leader Khomeini famously declared United States "The Great Satan". This factor of traditional rivalry against the West, the Saudis, and the Israelis, could play a role in why Iran's response was so aggressive. Since we are able to compare the rhetoric of Rouhani to former president Ahmadinejad, we can notice a distinct shift between the two: Ahmadinejad's fiery, anti-Western stance could easily result in different foreign policy outcomes than the more conciliatory and open Rouhani. It's possible these factors of cultural antagonism, rivalries, and leader rhetoric can explain some of the shortcomings of the defensive realist theory, and the prevalence of offensive responses in this case. The most effective of these response attacks, and the one with the most significant cyber security ramifications, was the attack launched on Saudi Aramco, the state-run oil company, an attack commonly referred to by the virus' name: "Shamoon".

Chapter 4: Saudi Aramco Attacks (Shamoon)

The Shamoon cases are more difficult to analyze because there is less concrete information on responses when compared to the three other cases used in this thesis. Very little information regarding the Saudi's responses has been made public. Apart from issuing warnings of the Shamoon virus, the Saudi state has said very little openly. Literature addressing this case is minimal, and the few sources that do address it conclude that there is not much to analyze in terms of response behaviors. This will impede analysis somewhat, but the behaviors that can be identified will be addressed, as well as the possibility that Saudi Arabia may have covertly launched a retaliatory attack. Overall, there is marginal evidence for a defensive theory of cyber, and the Saudis did not engage extensively with the international community, but it's difficult to conclude one way or the other given the scant information. Some additional factors for the nature of the Saudi responses (or lack of) are addressed as well.

On August 15th 2012, the state-sponsored oil company Saudi Aramco, the largest privately owned oil company in the world, was subject to a zero-day exploit cyberattack which wiped out the hard drives of around 30,000 machines. The malware was named

Shamoon after a name found embedded in the code. The malware functioned through a combination of separate modules. After acquiring login credentials, Shamoon, in both the 2012 and 2016 versions, activated a “dropper” component, allowing it to remain present on infected systems, followed by the “wiper” component, which drops a driver, allowing it to execute its function of wiping the system’s hard drive and replace data with a jpg image. The first Shamoon attack utilized a picture of a burning American flag, and the second wave of Shamoon attacks swapped this image for the infamous photo of drowned 3-year old Syrian refugee Alan Kurdi. Finally, a “reporter” component establishes communications with the attacker’s command and control server, and sends information to the server reporting that the target system’s hard drive has been successfully wiped. (Symantec Security Response 2012) (Symantec Security Response 2016)

The Aramco computers were not exactly “destroyed”, as many analysts and news outlets reported. Within ten days, the company had restored normal services and operations and replaced the unusable machines. Even the substantial cost of replacing and restoring 30,000 computers to operation, including any personnel expense was likely dwarfed by the staggering $1 billion daily profit of Saudi Aramco. Official statements from the company claimed no lasting damage of any kind occurred, and production was not hindered in any way. As the malware affected Saudi Aramco’s internal network only, the malware must have been introduced by a compromised piece of equipment or personnel working for the perpetrators of the attack. This attack is notable for the lack of real damage it caused in terms of any lasting or physical effects, but also because of the disproportionately alarmed response from many news outlets, analysts, and commentators. As US Defense Secretary Leon Panetta put it: “the most destructive cyberattack on the private sector to date” (Reuters 2012) A massively wealthy oil company was hit with an event which, in relation to the actual damage caused and costs incurred, could be called a minor inconvenience.

Shamoon disappeared for four years, until on November 17th, 2016, the malware resurfaced against Saudi targets, this time it struck the Saudi Aviation Agency. (Chan 2016) Following a warning from the Saudi government, the Ministry of Labour and Social

development and the Sadara Chemical Company (a joint venture of Saudi Aramco and Dow Chemical) were hit by Shamoon attacks. (Chan 2016)This second wave was described by Symantec analysts as “exceedingly similar”, and almost instantly familiar to any who had analyzed the first wave of Shamoon four years prior. (Alperovitch 2016) (Symantec Security Response 2016) The most immediate difference being the update of the jpg used to replace wiped files. John DiMaggio of Symantec, who responded to the initial cyberattack as well as the resurgence, said that “When malware is discovered in a targeted attack and documented publicly, the group behind the malware usually modifies the malware to avoid detection and/or burns infrastructure in an attempt to disappear…” He also commented that the disappearance and resurgence of the same malware (Shamoon) “…is almost unheard of and certainly not the norm.” He concluded with: “The only theory we can provide is that the attacker in both the 2012 attacks and 2016/17 attacks is primarily interested in Saudi Arabia.” (Williams 2017)

Following both waves of the Shamoon virus, fingers began pointing toward Iran. The possible motives for an organization, state or otherwise to execute an attack against Saudi Aramco and other important Saudi targets pointed toward one state. Many factors influence this conclusion; Iran and Saudi Arabia have been fierce regional adversaries since the late 1970s, with each competing for regional hegemony as Islamic leader nations. Their sectarian conflict has manifested in a multi-layered rivalry, as well as their recent backing of opposing sides in the war in Yemen. Iran’s Shiite revolutionary government is in direct ideological and political conflict with the Wahhabist-Sunni theocracy of the Saudis. The ongoing oil embargo on Iran no doubt fueled this animosity, as customers and currency fled Iran in the wake of crippling sanctions, while their regional rival was unaffected in this way, and likely even prospered as a result. As mentioned earlier, it is also more than likely that Iran, in direct retaliation for the attack against its nuclear centrifuges, struck back at a US proxy ally in the region with a cyberattack of its own. What better target than a long-standing regional rival, an ally of those behind Stuxnet, and a competitor in the oil market? It’s clear the intended goal was to cause damage which would stifle or hinder oil production for the Saudis, but that did not occur.

The Saudi response in the aftermath of both Shamoon waves was almost nonexistent. Perhaps considering the damage (or lack thereof), the Saudi state decided it best (or simply realistic) to note that the attack had no effect, and they would take no action. Apart from a likely restructuring of cyber defense and data backup at Saudi Aramco, nearly no formal overt foreign policy actions were taken by the Saudi state in retaliation for the attack. What is noteworthy is how the United States responded: a declaration that this kind of attack against its allies would not be tolerated and that such attacks would be subject to retaliation in the future. (Valeriano and Maness 2015) The Saudi government, apart from a warning to Saudi organizations and companies to be on alert for Shamoon cyberattacks in early 2017, seemed relatively unfazed by Shamoon in any of its iterations.

The Saudi response may, however, not be as nonexistent as it overtly and officially appears. Possible evidence for this idea comes in the form of a large-scale series of cyberattacks against the Iranian Oil Ministry in April, 2012. Iranian officials disconnected entire facilities from the internet, to try and contain an attack which utilized similar wiper components to the Shamoon attacks, erasing critical data and attempting to affect production. (Erdbrink 2012) After the attack, Iranian officials claimed that no production had been affected, and that the only delays had been administrative in nature. These responses were similar to those given by Saudi officials immediately after Shamoon. Although the perpetrator was not clearly identified, as is the case in most cyberattacks, possible perpetrators include Saudi Arabia or the United States. Some have called this series of attacks between Saudi and Iranian targets “tit-for-tat cyberwar “and this would suggest that Saudi Arabia is in fact engaging in responses to retaliate for the Shamoon attacks, themselves something of a retaliation for Stuxnet. (Chan 2016)

The responses on the part of the Saudi government in the wake of the Shamoon attack are noteworthy for the fact that virtually no significant overt response took place. Apart from the official company statement which reported that no significant structural damage had occurred and that oil production had not been affected, any other direct response, in the form of a threat, attack, or other diplomatic or foreign policy action is hard to confirm. The possibility of the Oil Ministry Attacks in April 2012 being a Saudi project is

possible, and would provide an enticing option for the Saudis, coupling well with their overt response: unfazed dismissal of the Shamoon attacks. A prudent and calculating attack retaliating in the same industry with the same goal (affecting oil production), would send the message that the Saudis are not to be messed with, and have no problem using the same methods against their attackers. Again, the anonymity and plausible deniability that cyberattacks offer would make it a prudent choice for the Saudi government.

In the interest of displaying power over its regional rival, and likely cyberattacker, Iran, it seems logical that the Saudis would remain relatively mute in the aftermath of the Shamoon attack. They already held the upper hand over Iran in nearly every possible way; they weren't under any embargo or sanctions from the West, and they sustained no damage or production problems. To issue no statement against Iranian action, even when evidence and circumstance points to them as the most likely perpetrator, likely portrayed a sense that the Saudis were hardly inconvenienced by the attack, did not view it very seriously, and in this way, made light of Iran's attempts to cause damage to their oil production. This could have been done in the hopes of humiliating or angering Iran after they executed a serious and no-doubt expensive operation.

It should be again noted that the lack of significant data in the form of identifiable responses poses a challenge to significant and detailed analysis. Even more so than the Iranians before them, the Saudis really left the majority of their responses to inference and conjecture. One reason the amount of information available for research on this case is so sparse is simply that very little has been made public. The internal affairs of the Saudi state in the aftermath of the Shamoon attacks are not widely available, and if they have taken a plan of action, they have not announced it to the world or their allies, as in the case of the Estonians and their Cyber Defense League.

Keeping this scarcity of concrete information in mind, we can address our analysis of the scenario with our potential defensive realist cybertheory. The Saudi response was overtly more defensive in nature than offensive. Their muted non-action, coupled with their statements of the negligible damage of the attack characterizes a measured and withdrawn response to such a large-scale and aggressive cyberattack. The company, after restoring

service and replacing the wiped systems, likely beefed up its security and, if it hadn't already, adopted more rigorous data backup measures, but evidence of self-help as a result of an anarchic international system is lacking in the immediate aftermath of the Shamoon attack. Of course, the possibility of the Saudis engaging in frequent, covert cyber war with the Iranians is possible, and more than likely the case. The mutual ramping up of cyber capabilities occurring between these back-and -forth cyberattacks on critical infrastructure is certainly possible to define as a security spiral; this could be in turn escalating hyperbolic discussion among analysts and media outlets about the destructive capability of cyber, and the dangers it poses to the private sector. Dramatic news overstating the damage of cyber and whipping up fears can only contribute to further insecurity.

Offensive Realism's predictions of aggressive and opportunistic state behavior do not seem to fit this scenario, unless they are indeed responsible for the retaliatory attacks on the Oil Ministry in April 2012. It would have been possible for the Saudi government to issue scathing condemnations or seek to take overt action against Iran, backed up by the US. As cyberattacks are harder than conventional attacks when it comes to procuring tangible proof, it may also be the case that the Saudis could not have made an aggressive move due to any obvious incrimination of the Iranian state, as this would have caused more international concern than the initial cyberattack. If, however, the Saudis were responsible for the Oil Ministry Attacks, or any other covert cyberattacks on Iran, these could be viewed as the kind of aggressive, offensive responses that the states in the Offensive Realist framework would undertake. An opportunity or justification for attempting to damage a rival's critical infrastructure would be acted upon in this framework.

Overall, the non-action of the Saudi state is itself an action, and a largely defensive one. Whether this provides strong evidence for the defensive theory of cyber is uncertain. There were likely other contributing factors to the Saudi's non-action, perhaps including the fact that the cyberattack may not have been serious enough to warrant a strong response. Whether this is justification for little to no response after a cyberattack committed by a longtime regional rival is open to interpretation, as far stronger responses and retaliations have been carried out by other states after far less serious initial attacks. To conclude this

case, we can classify the muted, measured action of the Saudis as primarily defensive behavior, and the lack of serious offensive, short-term measures signals that the Saudis are likely keeping long-term security and economic stability in mind when making foreign policy decisions.

Chapter 5: Ukrainian Power Grid Hacks

This case largely provides support for a defensive theory of cyber: Ukraine refrained from retaliatory attacks, similarly to Estonia. The element of our theory which suggests that states will seek the help of the international network also holds true here. Ukraine's appeals to the West and to the United States for continued support provide evidence for this tendency. Similarly to Estonia, the involvement with the international system in response to cyberattacks is sensible for states attempting to move toward, or remain in, the Westerns sphere of power, and who may not be large enough or strong enough to defend themselves from cyberattacks alone. While the exact nature of why Ukraine's response was defensive is unclear, the fact that it was of such a defensive nature, and involved calling on international allies and organizations for support provides moderate support for the theory.

On December 23rd near 3:30 PM, a power control center of the Prykarpattyaoblenergo firm in the Ivano-Frankivsk region was hacked. The worker monitoring one substation stared in disbelief as the cursor "went skittering across the screen". (Zetter 2016) The invisible hacker had taken remote control of the operator machine and was taking substations offline. By the end of the attack, nearly 60 substations had been disabled, with light and heat shut off for an estimated 230,000 residents. (Zetter 2016)

Although it seemed this attack sprang from nowhere, it had in fact been brewing for over half a year. The attack was conducted in several stages, the first of which was a spear-phishing campaign, targeting workers at the power station. Emails were sent with a corrupted word file attached, which, if opened, and macros enabled, would create a back door for hackers, utilizing the BlackEnergy3 program. The hackers were then able to conduct reconnaissance on the infiltrated machines and their users. To infiltrate the SCADA system, user credentials were harvested during the months leading up to the attack itself.

The hackers tampered with the uninterruptible power supply (UPS), to ensure the power would be far more difficult to restore. Additionally, the attackers studied the individual firmware of every substation that had been infiltrated and wrote new malicious firmware that would prevent operators from turning the power back on. After these months of extensive background recon and programming, the group unleashed a highly coordinated and well-timed assault on the substations. (Zetter 2016)

First, the group logged into the substation systems through hijacked VPNs. Secondly, they disabled the UPS systems (backup power), and followed this with their shutdown of substation power. In a brilliant example of contingency planning, the attackers bombarded the power company call centers with a TDos attack (telephone denial-of-service): Since the phonelines were overwhelmed with fake calls, residents were unable to report outages to the company. Finally, the hackers used the program KillDIsk to completely wipe operator systems. Power was out for one to six hours in various areas, and it was two months before operations were fully restored. (Zetter 2016)

Robert M. Lee of Drogos Security called the attack “sophisticated” (Zetter 2016). The attackers had to have been well-trained and well-funded to execute such an extensively planned, multi-faceted and intricate attack. The synchronization of the final phase of the attack also points to the workings of a well-funded and highly trained group of hackers, and not a lone black hat or loose coalition of hackers.

Almost exactly a year later, a highly similar attack took place at the Pivnichna transmission facility, outside Kiev. Less extensive and less damaging than the previous attack, power was cut for one hour for the equivalent of $1/5^{th}$ of Kiev’s power. Ukrainian cyber analyst company ISSP notes that the attack was conducted in a very similar manner, and likely by the same group. David Emm of Kaspersky labs said in an interview that “If a system is porous, it is likely to encourage further attacks.” A similar spear-phishing campaign was conducted, utilizing the same macro exploits of the previous attack, albeit in a much more refined capacity. The hacker’s tactics had evolved, even though the attack itself was much less extensive. (Zetter 2017) Instead of bricking systems as in the previous year’s attack, the attackers merely shut them down remotely. The same credential

harvesting and backdoor reconnaissance methods were used. One analyst said it appeared that Ukraine was a “hacker training ground” and that this attack seemed more of a “demonstration of capabilities”, as the damage that could have been done was far greater than what occurred. (Zetter 2017)

These attacks took place amidst a catastrophic deterioration of Russian and Ukrainian relations, following the 2014 annexation of Crimea. Several other key events point to the likelihood that this series of attacks has its origins with Russian actors and interests. After the Russian annexation of Crimea, Crimean authorities nationalized Ukrainian energy companies stationed there. In Ukraine, the parliament was set to consider the nationalization of private power companies in the country. Some of these companies are owned by a powerful Russian oligarch with close ties to Putin. Additionally, shortly before the attack, a group of Ukrainian activists destroyed power substations in Crimea after the Russian annexation, which cut power to 2 million residents. (Zetter 2016) Although the first attack on Ukrainian power substations occurred shortly after this destruction, hackers had already infiltrated the Ukrainian systems via phishing months earlier, and the Crimean station attack may have only served as a catalyst to hasten the launch of the already-planned attack.

The Ukrainian government and numerous Ukrainian officials hold Russia responsible for, or at least complicit in, the attacks. Ukrainian president Poroshenko claimed that Russia attacked Ukraine via cyber means nearly 6,500 times in the last two months of 2016, and went on to tell Reuters that Russia is “at cyberwar with the whole world”. (Zinets 2017) Russian officials have denied any responsibility for the attacks, and decried these and other accusations as a “tiresome witch hunt”. (Huggler 2017) Other Western voices have added their condemnation of the attacks, and of Russian cyberaggression. German head of the Bundesamt für Verfassungsschutz (BfV), Hans-Georg Maasen said that Russian cyberaggression posed a threat to German elections. (Huggler 2017) In the United States, cyberintelligence firm iSight identified the actors involved in both attacks on Ukrainian energy infrastructure as an APT known by the name “Sandworm”, because of references to Frank Herbert’s *Dune* novels embedded in the code. The use of BlackENergy3 and KillDIsk

are some of the identifiers of this actor. Sandworm is, according to US intelligence, a cyber actor with close ties to the Russian state, and their actions consistently align with the goals and interests of the state. (BBC Technology 2016)

An additional motive to the TDoS tactic in the 2015 cyberattack was likely to sow further distrust of the Ukrainian companies and government in the residents of the affected area. This motive would align closely to Russian state interests. Whether Sandworm is a proxy for the Russian government, or is indirectly supported or funded by powerful Russian interests has yet to be proven beyond all reasonable doubt.

Ukraine seems to be attempting to garner support for a multi-national, Western-driven cause against Russian cyberaggression. Russian-Ukrainian relations are at a low point since the annexation of Crimea, and these cyberattacks have no doubt compounded the fears and uncertainties surrounding Russian action. By publicly decrying Russia's alleged role in the attacks, Ukraine seeks to utilize international support to protect themselves against future attacks.

When analyzing the response behavior in this case, we will include responses after both power grid attacks. The concrete, observable actions of Ukraine and its highest officials included issuances of condemnation for the attacks, identifying Russia as the culprit, and warning Western allies that Russia poses a serious continued cyber threat on an international scale. Ukraine has not engaged in any aggressive responses or retaliatory attacks and, despite a seriously threatening and uncertain future, has not responded in any kind of offensive regard towards Russia, who it deems responsible for the power grid hacks. Offensive measures are lacking in this case even in the sense of the previous "Shamoon" case, where a state is perhaps believed to be engaging in retaliatory cyberattacks, but the evidence is inconclusive. Ukraine has, at least, not been accused of any serious cyberattacks by its rivals, therefore we will conclude for the case analysis that it has not engaged in offense-oriented cyber behavior.

Ukraine's responses to the series of power grid attacks were mostly defensive and withdrawn. They acted in restoring functions which had been hindered by the attacks to normal and increased defenses in cybersecurity. Ukraine also implored other nation-states

to help it further its security goals. The majority of the responses that are publicly available for research are of this nature. A recent Guardian article regarding Ukrainian president Poroschenko noted that the Ukrainian president is certain that the Trump administration understands the threat that Russia poses, and that the United States will stay "firm in its support for Ukraine in its battle to prevent Russian aggression". (Wintour 2017) In this way, we see Ukraine relying heavily on its allies and the assumption that they have continued shared security interests with regards to Russia. Assuming this, they try and draw as much attention to the Russian cyberthreat as possible, and hopefully mobilize allies to oppose Russia. By doing this, Ukraine will hopefully be better defended from Russian cyberthreats in the future, as the eyes of Ukraine's allies are now watching for a similar attack. The international community knows what they are looking for now.

Conclusion

Here I would like to summarize the results of applying our defensive realist cybertheory to the four cases in cyber security as described by this thesis, and offer some possible conclusions based on the patterns of response behavior in these four cases, as well as whether the theory was largely successful in its explanations or not. Shortcomings of the theory to explain events in the four cases, as well as some potential avenues for further research, will follow.

The case of Estonian infrastructure attacks largely provided support for the two primary aspects of the theory. Defense was the primary goal of Estonia in the immediate aftermath of the attack. The Estonians ramped up cyberdefenses with the creation of the CDL, and have not conducted, nor been blamed for, any retaliatory attacks against Russia or other rivals. Additionally, Estonia sought to achieve these defensive measures through International cooperation, particularly through NATO and the EU. This case therefore, provides support for the theory.

Iran, after Stuxnet, behaved in a way which largely contradicted the predictions of the defensive cybertheory. Very quickly after recovering from the Stuxnet attack, Iran launched multiple counterattacks through proxy groups in an aggressive surge of cyberattack retaliation. Their response was largely not focused on restrained defense, but

on lashing out to multiple rivals, and in that way, focused only on short-term victories over long-term stability. Additionally, Iran eschewed cooperation internationally to address the threat by cyber, and instead, took matters into its own hands. The theory found little support in this case.

In the Saudi Aramco/ "Shamoon" attacks, the defensive theory of cyber found mixed support. As mentioned, the difficulty with assessing the theory in regards to this case was due in part to the scarcity of data. Even so, it can be interpreted that a muted or nonexistent response is still a response, and certainly a more defensive response than a retaliatory attack. The Shamoon case provided little evidence for the idea that states seek the international community's assistance in cases of cyberattacks. Saudi Arabia, similarly to Iran, did not seek the help of IO's, and apart from a statement of support by the United States, the country kept mostly to its own devices. Other factors or more concrete information on Saudi responses may be required for better analysis of the case.

Finally, Ukraine and the energy infrastructure attacks: This scenario provided moderate support for both aspects of the theory. In many ways the Ukrainian response echoed that of Estonia: loud and widespread condemnation of Russia as the agitator, and appeals to international allies and organizations to take notice of the attack and the threat Russia poses. Ukraine was not accused of, or made public, any response attacks via cyber means. Its international focus, largely withdrawn response, and lack of tit-for-tat responses provide good evidence for the theory.

Naturally, there are several limiting factors that inhibit the robustness of this analysis. Several can be identified as shortcomings of the theory itself, while others are simply inherent to any empirical work in cybersecurity. In regards to the theory, the framework cannot explain the Stuxnet responses on the part of Iran. Reasons for this might be that the theory does not take historical and cultural factors into account, and instead makes general prescriptions for state action in a given scenario. The cultural ethos and historical background of Iran as a people and as a political regime are leagues apart from the history and culture of Estonia, and this difference likely plays a role in how the nations respond to being attacked in cyber. The very name of the cyber security groups associated

with each country can provide clues into these differences. Estonia's "Cyber Defense League"-measured, calculated, oriented towards defense and the protection of the country from external threats. Conversely, "The Iranian Hacking Army" or "Cutting Sword of Justice"- vengeful, aggressive, geared towards offense and retribution against perceived injustice. These types of cultural and historical differences, manifested into foreign policy, are not contained within the scope of this defensive theory of cyber, but no doubt have an impact on state responses. There may also have been simple practical reasons why states did not engage in counter attacks, which aren't covered in the defensive realist cybertheory. Estonia worked with the international system in this case, but it's difficult to ignore the fact that it's just as likely they wouldn't have, and instead responded with counter cyberattack, if they were powerful enough. These practical, material capabilities are difficult to determine in cyber conflict, as the true capabilities of any nation-state are not as objectively measurable, or publicly discoverable, as they are with conventional measurements such as number of tanks, aircraft carriers, or missiles.

One question which warrants further exploration is the matter of why some nations engage with the international system and not others? Western-oriented states clearly tend to work with IOs, and at least from our cases, non-western states do not. A splinter analysis could be conducted to shed light on why some states engage with the international system more robustly and reliably than others in the wake of cyberattacks, and what factors determine this. Some elements of liberalist theory could prove useful in explaining the behavior of some states. Estonia and Ukraine's responses, Estonia in particular, were geared toward international cooperation, dialogue, and institutions. This tendency was not true of Saudi Arabia and Iran, but in the Eastern European cases, perhaps more liberalism in theory creation could help explain the international cooperation observed.

Additionally, the analysis is hindered by the short time that cyberconflict has existed. The underdeveloped repertoire of cases means that by default, empirical research is limited.

Another barrier to research in cyber stems from its inherent element of anonymity. There are two secondary factors which affect research because of this. One is the inability

to determine conclusively who has carried out a cyberattack. Due to the nature of cyber, which naturally shields the identity of actors to a degree where government involvement in an attack can be denied, we cannot prove beyond all reasonable doubt, with tangible evidence, that a cyberattacker was Russian, Iranian or American. This inability to prove the identity of the attacker in the same way as conventional attacks may affect the extent of state and IO responses. Actors may not take strong offensive action in retaliation for cyberattacks because the attacker's identity cannot be proven or even argued for in the same way that conventional attackers can. For example, the recent missile launch on a Syrian air field ordered by President Trump was presented as a direct response to a chemical weapons attack which killed dozens of Syrian civilians. The overt nature of the chemical weapons attacks and the missile response were only possible because the identity of the attackers was easily verifiable. The certainty with which American and Western officials claimed the attack was the work of the Syrian government is unheard of in cyberattacks. This uncertainty seems to temper all retaliation measures. (Although it certainly doesn't temper retaliatory rhetoric on the part of state leaders). The secondary effect is that this pervasive culture of secrecy inhibits access to important empirical material for analysis and dialogue to occur. A prime example is the Stuxnet case. Even though it is widely accepted that the Americans and the Israelis were behind the attack, the fact that it is still not officially acknowledged, and most details are hidden from the public view in the way a missile launch could never be, prevents further development of analysis, debate, and theory formation.

This critical difference of identity between conventional and cyberattacks may be one reason why analysis through a defensive realist lens did not always lead to accurate prescriptions for state and international behavior. Another observation is that some states responded to cyberattacks by engaging further with other international actors and organizations, appealing to them for assistance, protection, and further dialogue.

The size of the nation in relation to its targets is another element which was not addressed in this analysis, but could very likely play a role in the perpetuation of cyberattacks. Through these four cases, we find little correlation between the material

capabilities of a nation and their tendency to attack targets smaller or larger than they are. Large countries attack small ones, (Russia->Estonia and USA->Iran) and small countries attack larger ones. Iran had no problem launching a significant attack on United States banks, and they are far less powerful then their target.

Taking these potential shortcomings into account, it will prove beneficial to adopt these shortcomings into areas of focus and improvement for future analysis. Several paths forward exist for continued cyber research in IR theory. As more material becomes available regarding the internal processes and covert material comes into official status, a neoclassical realist approach may prove beneficial. Combining material capabilities with domestic variables, *and* the perceptions of the individual leaders of a nation provides a more empirically sound and holistic framework for determining state behavior. Interpreting the statements of Iran's Supreme Leader, Russian foreign officials, or the Ukrainian president provides one of the clearest possible methods of gleaning the beliefs and motives of a nation's leaders, and in which direction that nation will most likely act. Neglecting this variable can create confusing prescriptions for behavior, undermining understanding of state motives. Neoclassical ideas seem a strong fit for continued analysis of cyber action and its effects on international relations and state behavior. As the covert nature of many cyber operations leaves uncertainty in variables and evidence which renders other branches of realist analysis coming up short or inconsistent, neoclassical analysis and its inclusion of the leaders stated motives and goals can help to make this type of analysis more robust and convincing.

In time, as, more cyberattacks occur, more international treaties are established, and more norms of cyber-engagement arise, perhaps liberalism and constructivism will play a larger role. The longer that norms of behavior develop, the more fruitful it will be to conduct analysis in this lens. A look at the historical and cultural factors of each nation is also worth considering as a factor in the way states conduct cyberattacks. What is their attitude? What is their historical framework of morality or justice? What from the historical record would promote a state to attack another state, or choose defensive, diplomatic, or

otherwise peaceful measures? The answers to these questions can largely be used to determine how they will respond to attacks in cyber.

It may come about that entirely new schools of International Relations theory arise, as the nature of IR in cyber will be shown, over time, to be somehow fundamentally different in a way that necessitates entirely new theory to be understood. There is no doubt, however, that the moderate success of Defensive Realism as a theory for cyberconflict should point to the bulk of analysis and theory-development that has yet to be done. Such analysis will be indispensable in the future.

Works Cited

Aitel, Dave. "Guest editorial: The DNC hack and dump is what cyberwar looks like." *ars Technica*, June 17, 2016.

Alperovitch, Dmitri. "Shamoon Round 2 or the Power of Machine Learning." *CrowdStrike Blog.* December 1, 2016.

Arquilla, John. "Cyberwar is already upon us." *Foreign Policy*, 2012: 83-84.

Arthur, Charles. "Saudi Aramco hit by computer virus." *The Guardian*, August 16, 2012.

Balzaq, Thierry. "The Three Faces of Securitization: Political Agency, Audience and Context." *European Journal of International Relations*, 2005.

BBC Technology. "Hackers behind Ukraine power cuts, says US report." *BBC News*, February 26, 2016.

—. "Shamoon virus targets energy sector infrastructure." *BBC News*, August 17, 2012.

—. "Ukraine power cut 'was cyber-attack'." *BBC News*, January 11, 2017.

Betz, David J., and Tim Stevens. *Cyber-Power, Cyberspace and the State: Toward a Strategy for Cyber-Power.* London: IISS, 2011.

Blair, David. "Sitting near a nuclear tripwire, Estonia's president urges Nato to send troops to defend his country." *The Telegraph*, April 11, 2015.

Brasso, Bret. "Cyber attacks against critical infrastructure are no longer just theories." *FireEye*, April 29, 2016.

Brenner, Susan W. ""At Light Speed": Attribution and Response to Cybercrime/Terrorism/Warfare." *The Journal of Criminal Law and Criminology*, 2007: 379-475.

Business Insider. "US General: Iranian Cyberattacks Are Retaliation For The Stuxnet Virus." *Business Insider*, January 18, 2013.

Capaccio, Tony, David Lerman, and Chris Strohm. "Iran Behind Cyber-Attack on Adelson's Sands Corp., Clapper Says." *Bloomberg Technology*, February 26, 2015.

CCDCOE. *National Cyber Security: Framework Manual.* Edited by Alexander Klimburg. Tallinn: NATO/OTAN, 2012.

Chan, Sewell. "Cyberattacks Strike Saudi Arabia, Harming Aviation Agency." *The New York Times*, December 1, 2016.

Clemente, Dave. "International Security: Cyber Security As A Wicked Problem." *Chatham House*, 2011: 15-17.

Department of Homeland Security. *What is Critical Infrastructure.* October 14, 2016. https://www.dhs.gov/what-critical-infrastructure (accessed January 25, 2017).

Eisenstadt, Michael. "Cyber: Iran's Weapon of Choice." *The Cipher Brief*, January 29, 2017.

Erdbrink, Thomas. "Facing Cyberattack, Iranian Officials Disconnect Some Oil Terminals From Internet." *The New York Times*, April 23, 2012.

Eriksson, Johan, and Giampiero Giacomello. "The Information Revolution, Security, and International Relations: (IR) Relevant Theory?" *International Political Science Review*, 2006: 221-244.

Finkle, Jim, and Jeremy Wagstaff. "Shamoon virus returns in new Gulf cyber attacks after four-year hiatus." *Reuters*, 12 1, 2016.

Fitzsimmons, Michelle. "Shamoon malware virus swipes and wipes PCs." *Tech Radar.* August 12, 2012.

Zero Days. Directed by Alex Gibney. 2016.

Gjelten, Tom. "FIRST STRIKE: US Cyber Warriors Seize the Offensive." *World Affairs*, 2013: 33-43.

—. "Volunteer Cyber Army Emerges in Estonia." *NPR*, January 11, 2011.

Gjelten, Tom. "SHADOW WARS: Debating Cyber 'Disarmament'." *Sage Publications Inc.*, 2010: 33-42.

Goodin, Dan. "Defects leave critical military, industrial infrastructure open to hacks." *ars Technica*, July 13, 2012.

—. "Hack attack on energy giant highlights threat to critical infrastructure." *ars Technica*, September 26, 2012.

—. "Hackers trigger yet another power outage in Ukraine." *ars Technica*, January 1, 2017.

Hansen, Lene, and Helen Nissenbaum. "Digital Disaster, Cyber Security, and the Copenhagen School." *International Studies Quarterly*, 2009: 1155-1175.

Herz, John H. *Political Realism and Political Idealism.* Chicago: University of Chicago Press, 1959.

Herzog, Stephen. "Revisiting the Estonian Cyber Attacks: Digital threats and Multinational Responses." *Journal of Strategic Security*, 2011: 49-60.

Higgins, Kelly Jackson. "Lessons From The Ukraine Electric Grid Hack." *Dark Reading*, March 18, 2016.

Huggler, Justin. "Germany accuses Russia of cyber attack on Ukraine peace monitors, as Kremlin dismisses US intelligence claims as a 'witch hunt'." *The Telegraph*, January 9, 2017.

Hundley, Richard O., and Robert H. Anderson. "Emerging Challenge: Security and Safety in Cyberspace." *IEEE Technology and Society Magazine*, 1995: 19-27.

info security group. "Disttrack/Shamoon: a new targeted and destructive virus." *info security group*, August 17, 2012.

Jackson, Robert, and Georg Sorenson. *Introduction to International Relations: Theories and Approaches.* Oxford: Oxford University Press, 2016.

Kurkov, Andrei. "What Young Ukrainians Think." *British Council*, June 30, 2015.

Lin, Tom C.W. "Financial Weapons of War." *Minnesota Law Review*, 2016.

Maurer, Tim. "The New Norms: Global Cyber-Security Agreements Face Challenges." *Carnegie Endowment for International Peace*, 2016.

Mearsheimer, John J. *The Tragedy of Great Power Politics.* New York, 2001.

Morgenthau, Hans. *Politics Among Nations.* New York: McGraw-Hill, 1948.

Perlroth, Nicole. "In Cyberattack on Saudi Firm, U.S. Sees Iran Firing Back." *The New York Times*, October 23, 2012.

Platt, Victor. "Still the fire-proof house? An analysis of Canada's cyber security strategy." *International Journal*, 2011: 155-167.

Polityuk, Pavel. "Ukraine investigates suspected cyber attack on Kiev power grid." *Reuters*, December 20, 2016.

Reuters. ""Shamoon" virus most destructive yet for private sector, Panetta says." *Reuters*, October 11, 2012.

—. "Iran strengthened cyber capabilities after Stuxnet: U.S. general." *Reuters*, January 17, 2013.

Rid, Thomas. "Think Again: Cyberwar." *Foreign Policy*, 2012: 80-84.

Rose, Gideon. "Neoclassical Realism and Theories of Foreign Policy." *World Politics*, 1998: 144-172.

Russell, John. "The NSA can now share unfiltered surveillance data with other intelligence agencies." *Tech Crunch.* January 3, 2017. https://techcrunch.com/2017/01/13/the-nsa-can-now-share-unfiltered-surveillance-data-with-other-intelligence-agencies/ (accessed January 23, 2017).

Sandle, Tim. "Shamoon Virus Attacks Saudi Oil Company." *Digital Journal*, August 12, 2012.

Sans ISC. *Analysis of the Cyber Attack on the Ukrainian Power Grid.* Defense Use Case, Washington D.C: E-ISAC, 2016.

Saudi Gazette. "Saudi Arabia warns over cyber attacks as labor ministry hit." *Saudi Gazette*, January 24, 2017.

Seals, Tara. "Shamoon Resurfaces to Attack Important Saudi Targets." *Infosecurity*, 12 2, 2016.

Sen, Ashish Kumar. "Iran's Growing Cyber Capabilities in a Post-Stuxnet Era." *The Atlantic Council*, April 10, 2015.

Shiraev, Eric B. *International Relations.* New York: Oxford University Press, 2014.

Symantec Security Response. "Greenbug cyberespionage group targeting Middle East, possible links to Shamoon." *Symantec Official Blog.* January 23, 2017.

—. "Shamoon: Back from the dead and destructive as ever." *Symantec Official Blog*, November 30, 2016.

—. "The Shamoon Attacks." *Symantec Official Blog.* August 16, 2012.

Taliaferro, Jeffrey W. "Security Seeking Under Anarchy: Deensive Realism Revisited." *International Security*, 2001: 128-161.

Tang, Shiping. "Taking Stock of Neoclassical Realism." *International Studies Review*, 2009: 799-803.

The Culture of National Security. New York: Columbia University Press, 1996.

Valeriano, Brian, and Ryan C. Maness. *Cyber War versus Cyber Realities: Cyber COnflict in the International System.* New York: Oxford University Press, 2015.

Wendt, Alexander. *Social Theory of International Politics.* Cambridge: Cambridge University Press, 1999.

—. *Social Theory of International Politics.* Cambridge: Cambridge University Press, 1999.

Williams, Brad D. "Iran's cyber strategy: A case study in Saudi Arabia." *Fifth Domain Cyber.* February 7, 2017. (accessed March 15, 2017).

Wintour, Patrick. "Ukraine president says sanctions keep Russian tanks out of central Europe." *The Guardian*, April 19, 2017.

Yould, Rachel E. "Beyond the American Fortress: Understanding Homeland Security in the Information Age." *Bombs and Bandwidth: The Emerging Relationship Between Information Technology and Security*, 2003.

Zetter, Kim. "The Ukrainian Power Grid Was Hacked Again." *Motherboard*, January 10, 2017.

—. "Everything We Know About Ukraine's Power Plant Hack." *Wired*, January 20, 2016.

—. "Inside the Cunning, Unprecedented Hack of Ukraine's Power Grid." *Wired*, March 3, 2016.

—. "The NSA Acknowledges What We All Feared: Iran Learns From US Cyberattacks." *Wired*, October 2, 2015.

Zinets, Natalia. "Ukraine charges Russia with new cyber attacks on infrastructure." *Reuters*, February 15, 2017.

On my honour as a student of the Diplomatic Academy of Vienna, I submit this work in good faith and pledge that I have neither given nor received unauthorized assistance on it.