UNIVERSITÀ DEGLI STUDI DI MILANO

Graduate School in Social and Political Sciences

Dipartimento di Studi Sociali e Politici

Doctoral Programme in
POLITICAL STUDIES
XIX Cohort

Explaining State Behaviour During Cyber Disputes

SPS/04 – Scienza Politica

Candidate
Alessandro Fasani

Supervisor
Prof Andrea Locatelli

PhD Programme Director
Prof Francesco Zucchini

A.A. 2016 / 2017

Table of Contents

Introduction

Importance Of The Selected Topic

Technology has always affected the way wars are fought. From the introduction, and the subsequent evolution, of firepower, through naval- and airpower, to the use of space and cyberspace as auxiliaries in military operations (Van Creveld, 1991; Krepinevich, 1994; Boot, 2006; Locatelli, 2010). Cyberspace as the fifth domain of warfare is a relative new scenario. Indeed, it should be noted that, before the end of the '90s, computer technology was intended mainly as a communication and informative support for military operations (Krepinevich, 1994). Indeed, starting from 1999, the world experience the advent of aggressive Computer Network Operations, or CNOs. Starting in 1999 with the espionage operation dubbed Moonlight Maze, states discovered that system could be penetrated in order to steal sensitive information, in this particular case the US and Russia. Moonlight Maze was followed by Titan Rain and Operation Aurora, in 2003 and 2009 respectively, other espionage operations conducted by China against the US. In 2007 and 2008, Russia conductive hostile and disruptive CNOs against Estonia and Georgia respectively, using Distributed Denial of Service attacks on a large scale. Then came 2010, the year of the Stuxnet malware. Stuxnet was the real game changer in this field of study: it gave proof that a computer code is able to produce disruption in the real world (Falliere et al., 2011). However, other means to attack networks and infrastructures exist, since malwares are discovered every day, given the fact that the systems are designed by humans, which are not prone to perfection. Many governments today have espionage capabilities, they could potentially cripple economies and businesses, as well hitting civilians, by exploiting the dependency of private and state actors to the global network. The vulnerabilities in modern technology still refer to the warning of Clausewitz, according to which everything is subject to attrition, and if something breaks or doesn't function like its intended to it complicates even the simplest of tasks (Clausewitz, 1832; Giacomello e Badialetti, 2009). Therefore, governments are becoming more aware of their vulnerabilities and abilities in cyberspace, to consequently shape their strategies, due to the fact that today, cyberspace is shaping the national security documents of many states.

As stated above, the advent of cyberspace, or at least its military exploitation to achieve strategic goals, could remind of previous transformations in international security of the likes of the introduction of the airplane and the nuclear bomb, that were innovations that changed strategic thinking and state behaviour (Saltzman, 2013). For this reason, a new approach tailored specifically for cyberspace is needed, as this technological revolution calls for a rethinking - from scholars and policymakers - about force and conflict (Kello, 2013). Warfare in cyberspace could be considered new wine in old bottles, however the modalities through which it is carried out are different compared to the ones in classic domains. The cyber experience, being a brand new one, could require new analyses that classic theories *strictu sensu* could not be able to provide (Kello, 2013). Not only, such a view minimizes the larger scope of implications of hostile uses of cyberspace. The literature on war and power in cyberspace is expanding, but remains mostly an inner-looking field. The technological hurdle that constitutes a prerequisite to begin to understand this new field, and the fact that a cyber conflict did not cause human victims yet, led to a tardiness in the research by scholars and academics (Kello, 2013). Academics play an important and privileged role in resolving cyber strategic problems, but - as for now - the literature in international relations that tackles this problem remains scarce. However, this problem is understudied in the academic world, and it surprising, given the fact that the intelligence community in the United States states that the cyber threat is more dangerous than global terrorism (Kello, 2013), and several experts are warning for an expected huge cyber attack, dubbed “cyber pearl harbour” (Bumiller, 2012; Tadjdeh, 2015) or “doomsday scenario (Drogin, 2010)”. Whether this is scaremongering or not, there is the need to apply classic theories of International Relations which have the potential to be adapted to the virtual domain and that could be able to explain the influence that derives from the exploitation of cyberspace on international security and on the diffusion of war (Saltzman, 2013).

How This Research Fits In The Current Literature

International Relations Theories and Cyberspace

Literature about the virtual realm goes back as early as the ‘70s, addressing the potential of the new ICT technology as a threat to the State (Eriksson and Giacomello, 2007). This topic would be the silver lining of politics literature on cyberspace until today. The first example to be cited is the so-called Tengelin Report to the Swedish government published in at the end of the ‘70s, which firstly addressed the risk of the dependence on network technology as a vulnerability factor for the State (Eriksson and Giacomello, 2007; Braman, 2009). Another report dealing with the possible threats of the diffusion of digital networks is French - published in English in 1980 - and it is the Nora/Minc report which addresses the problem of the American government controlling European networks (Braman, 2009). Publications on the subject went onward through the '80s but it straightforwardly began to flourish during the '90s, on par with the rapid diffusion of ICT and the first attempts of cyber attacks. Much of the literature taken into consideration revolved around the concept that this new networked and networking technology eroded of power of the State, both for the lack of the capability of controlling and managing it as well as the emerging of new players using this same technology very efficiently, such non-state actors, such as firms, transnational organizations and individuals.

Concerning the first point, during the '90s several authors approached the birth of ICT focusing on the fact that controlling the flow of information has always been a strategic asset and a pivotal exercise for the security of nations (Braman, 2009) but somehow this task revealed to be much more difficult with this new technology (Agnew and Corbrige 1995; Anderson 1995; Krasner 1995). The challenge for the State regards their internal sovereignty, that is to say its ability to control its territory and the people it comprises because cyber attacks are detrimental to the values of information, both tangible and intangible, but also the capability of the government to have an effective and total control over the events (Eriksson and Giacomello, 2007). As far as the second point is concerned, the threat of networking technology linking together different actors and threatening the power of the state was put down in writing, for example, in Castells' trilogy in the '90s. The author stressed the fact that global digital connections will give birth to two new threats for the State: the first one would be international criminal networks challenging the *power* of the State as a law enforcer and a security provider; the second one being a new worldwide society, with its own new identity, that would have rendered the *notion* of States less important (Castells, 1996. 1997, 1998).

The new actors in cyberspace on the one hand challenge the sovereignty of the State and its role as a security provider and, on the other hand, they could act as new security providers (Arquilla and Ronfeldt 2001; Nye 2003, Eriksson and Giacomello, 2007). Furthermore, as far as security is concerned, the first pieces of literature on ICT technology did not tackle the security issue, or better, they highlighted the dangers for private actors, such as businesses and the economy instead the ones for the State (Erikson and Giacomello, 2007). On the same page, it is worth mentioning that also Robert Keohane and Joseph Nye amended their work on the complex interdependence by adding cyber threats and how they are capable to influence over international relations (Keohane and Nye 1998; Nye 2003). However, they tackle the problem from an economic perspective only.

One could argue here that is true that many actors could gain the upper hand in the cyber realm, not as competitors of the State but as an obstacle in the regulation of the realm driven by self-interests. It's not, as it was in the early age of the IT society a matter of controlling the flow of information, rather the lack of security imposed on it. It is safe to say that the State remains the main player on the field, despite the presence of new and different actors that use the same methods to have access and operate, even in malicious terms, in the same virtual domain as the State does. What could be said instead is that, given the fact that still plays a significant role compared to the past, new actors (private security companies, for example antivirus providers and also transnational actors that push for better regulations of the internet and cyber realm as a whole) should contribute and have a say – in various degrees – to national and international cybersecurity.

Realism

The realist approach to the exploitation of cyberspace focused on two points. The first was that some realists considered cyber threats as primarily a problem concerning economies, not governments of states. Sometimes these authors did not even regard cyber threats as security issues (Erikson and Giacomello, 2007). The second matter of interest was how to categorize the new methods of warfare, that is to say the place that cyber warfare should take within the realist discussion. Some realists would argue that cyber methods of warfare could be added to the discussion and also taken into consideration if treated like a new instrument within a classic interstate conflict framework (Lonsdale, 1999). Under this point of view, the cyber realm is considered as a natural continuation of the realm of communications, since the control over information and communications – for example the encryption of messages, or the interception and jamming of communications – has been part of warfare since the dawn of time. In light of this, traditional realists do not consider cyberspace and its military exploitation as a whole new domain of warfare but simply an evolution of a pre-existing domain (Erikson and Giacomello, 2007). Another view regarding cyber concerning the discussion on the evolution of military technology states that cyber means of warfare are not an evolutionary step from the telecommunications realm but solely an addition to classic means of warfare (Lonsdale, 1999, Biddle, 2010). Another issue with the realist approach to cyber security is the view, shared with some liberal authors, that the influence of cyber threat is limited to the economy and not to national security.

It is straightforward to see how this view must be updated in light of recent events starting from the Stuxnet malware, the North Korean attack against Sony (Robb, 2014) and the Russian attack against the Ukrainian electric grid (Kovacs, 2016). These are the prime examples of direct attacks to be mentioned but hostile Computer Network Operations from States against other States became very common, and it is fair to affirm that they are now a matter of national security. Indeed, also Computer Network Exploitation attacks, that is to say indirect attacks, such as espionage operations, have become so intense and large in scale that could not be considered new wine in old bottles anymore. The economy remains one of the preferred target by malicious actors operating in cyberspace, but that concerns more episodes related to cybercriminal activities, today the focus shifted on the State and its security due to attacks against institutions and infrastructures that lie at the core of the functioning of societies.

Realism offers a useful approach to tackle the issue that stems from State and State-sponsored hostile activities in cyberspace. Indeed, despite the cyber domain is a new and different environment in which to “wage war”, classic theories become optimal tools to decipher States’ behaviour. For example, both the concepts of balance of power as well as arms race are definitely applicable to the cyber realm, as thoroughly explained in this dissertation. Namely, considering the freedom and the ability of performing offensive and defensive tasks in cyberspace as a source of power, that we could call cyber power, one could try to assume also the distribution of power and then its balance. This stems from the fact that the balance of the virtual domain shifted towards all those states capable of performing CNOs at first, and then towards all those states that could launch destructive CNAs, such as the US from 2010 – with the Stuxnet malware – and Russia from 2016, which demonstrate to be able to cripple electric grids. However, despite some countries in the world hold more cyber power compared to the others balancing or balancing alliances did not happen. This could be explained by the fact that cyber power is not a fixed parameter like the military one or it is not yet considered a serious threat. However, another option that could be taken into consideration is that this imbalance of power triggers another consequence, due to the appealing characteristics of cyber weapons and the lack of international regulations. Indeed, such a weaponization at state-level of cyber tools produces a security dilemma, and from this insecurity stems what we could consider an arms race. Huntington (1958) proposed a differentiation between qualitative and quantitative arms race. The first kind referred to weapons' technological advancement, the second kind, straightforwardly, to the quantity of military forces. It is obvious, as well as needed, to specify that in the cyber realm we must address qualitative arms race since it is technologically based weaponry. However, that is not the only reason. Military cyber activities are usually secret therefore it is not possible to have precise estimates of the number of people working on the subject as well as the number of cyber weapons possessed by a State. The problem surrounding cyber weapons will be thoroughly treated in chapter II of this dissertation. Nonetheless it must be said that as cyber weapons are not physical objects like, for example, missiles they cannot be stockpiled in the classical sense. What can be "stockpiled" meaning collected and readily available are 0days exploits, that is to say those means to take advantage of unknown vulnerabilities, around which many cyber weapons are built. This is strictly connected to the fact that the absence of perception of another State's capabilities raises the level of insecurity. Another balance that could be taken in consideration in the cyber realm is the offense-defense one. Following offense-defense theory when the attack has the upper hand over the defence, then war is more plausible. In cyberspace, given the intrinsic characteristics of the digital domain, the attacker has more probabilities to strike a successful attack compared to the chances of the defender to fend off hostile actions (Locatelli, 2015). Last but not least, due to the absence of international regulatory framework, cyberspace could be considered anarchic, and therefore fertile ground for realist analyses.

Using realist theories helps explaining *how* States employ cyber means to signal and project their power to other States and how it is perceived by other States thus giving the chance to analyse the destabilizing power of cyberspace and the real potential of cyber weapons as tools for conflicts. In light of this, and since cyber weapons are now powerful tools in the hands of States, realism helps us seeing how classic mechanisms such as deterrence and arms races happen in the cyber realm.

Liberalism

The liberal paradigm, on the other hand, reprises what has been already stated at the beginning of this paragraph, that is to say the emergence of numerous new actors that intervene in the security-building mechanism as far as cyber security is concerned. Not only, for some liberal authors the problem is still seen largely as an economic issue, on par with environmental security. The liberal paradigm's fault is exactly this comparison that is misleading at best. Humans do not have power to directly bend the environment to their will and to use it to harm other states, attacks to the economy are political in nature and are an extension of soft power. On the contrary, actively using hostile CNOs belongs to the military realm *tout court,* due to the fact that the primary users of cyber attacks are armies and groups that could be defined as paramilitary.

The presence of new and multiple actors in the digital realm is a fact and cannot be discussed with. What can be discussed is the role of these new actors. Liberal authors insist on the fact that the power and the role of the State is somehow limited as far as cyber security is concerned (Alberts and Papp, 1997; Eriksson and Giacomello, 2007). However, when stating that the State alone is not sufficient in building security (Eriksson and Giacomello, 2007) the authors do not present an alternative, for example an international institution capable of managing global cyber security. This is for two reason. The first being the centrality of the State as a provider of security which has been outlined above. The second reason is that such a body is impossible to build to begin with. Cyberspace, and its military exploitation, it's still a deeply unregulated environment due to the absence of an international regulatory framework. This situation, that will be discussed in the first chapter of this dissertation, fosters the secrecy of the cyber operations and strategies of the various States, therefore hindering a process of information sharing that is the basis for a hypothetical international institution. However, it is undeniable that many private actors concur for the securitization of cyberspace. On the technical level we could find for example ISPs (internet service providers), software companies, antivirus companies, hardware companies, and all the other actors that participate in the supply chain for ICT systems. On the non-technical one, international and intergovernmental organizations and NGOs work for the proposal of guidelines to better address cybersecurity. This situation depicts the modern power redistribution in the virtual domain, but what the liberal paradigm fails to grasp is that the pillar of security remains the State. All of these actors cannot enforce guidelines only the State can do that, and it is the State that pushes all the technical actors, for example, for an increase of built-in security in their electronic systems due to the fact that security is not the primary focus for many non-state actors operating in the cyber realm.

Due to the presence of many non-state actors that are responsible – in different degrees – for national and international cybersecurity an increase in public-private partnership is a compulsory step for the future. The nature of the actors becomes pointless in two different scenarios. If the attacker is a private party, for example a terrorist or a hacktivist organization and the victim is a critical infrastructure, then it is the State that must intervene in any way possible to this attack. Critical infrastructures are the backbone for society, and even if they are privately owned, the consequences of an attack could easily fall back on the whole society. The second scenario is a State that deploys a cyber weapon against a privately owned infrastructure. This is an attack which the State has to respond to. Again, this is of paramount importance for critical infrastructures, but it also concerns private infrastructures of different nature. For example, the North Korean attack against Sony in 2016, spurred a reaction from the US government, which assisted Sony in the recovery and responded against North Korea with economic sanctions and even a retaliation in the cyber realm, as described in the last chapter of this dissertation.

Due to constraints placed by the insecurity in the attribution mechanisms and the lack of international law concerning the word "response" does not necessarily mean a physical response but also the launch of an investigation and assistance in fixing the systems, for example. However, using a liberal paradigm is undoubtedly useful to study and analyse the mechanisms of international cooperation in matters regarding cybersecurity, such as the promotion of guidelines and best-practices, hence governance and also the behaviour of other non-state actors such as NGOs and other private actors, for example the focusing on the consequence of the diffusion of ICT technologies over the citizens of the world, like the aforementioned "network society" (Castells, 1996) or "global civil society" (Lipschutz, 1992).

Constructivism

The constructivist school of International Relations places itself on a completely different level compared to realism and liberalism. The constructivist approach is based on the notion that social reality is a construction made by norms, beliefs, identities and institutions. These concepts are part of dynamic processes hence subject to change in contrast to a material reality that is more or less static. Due to the dynamicity of social reality, also the approach to threats is different. Constructivists focus on what *could become* a threat not considering threat as fixed compared to realists and liberalists. Despite the fact that some constructivists do actually focus on States (Wendt, 1992), their primary focus is on the individual (Adler, 2002) and also the relationship between the individual, in the form of national identity, and national security (Buzan et al., 1998). Due to the dynamicity of the constructivist approach, they take into consideration all kinds of threat ranging from state to non-state actors and from technical errors to environmental disasters (Erikson and Giacomello, 2007). This kind of approach however, does not help in the framing of this dissertation since its focus is on State behaviour during cyber conflicts and disregards completely other threats that do not have their origin in the State.

The most valuable contribution of the constructivist school is the securitization approach (Buzan et al, 1998). Despite the fact that the Copenhagen school did not produce anything relevant about security and the cyber realm, other constructivists studied the securitization process as far as CNOs are concerned (Der Derian, 2000; Everard, 2000; Eriksson and Giacomello, 2007, Eriksson, 2017). The focus of one of these studies was the framing of a particular cyber attack. That is to say that framing an attack as "cybercrime" or "cyberwar" has two different impacts in the perception of such attack, automatically linking one to a criminal, and therefore subject to the work of the police, and the other to another state actor, falling back to the military sphere of competence (Eriksson, 2017). The other studies also focus on the role of perception within the cyber realm, but from two different points of view: one regards cyberspace as a threat to the concept of identity itself and as something that helps constructing new identities (Everard, 2000), the other tackles the shift in the perception of those who attack through cyber means that could are distant from the actual target, both in geographical and "sentimental" terms. (Der Derian, 2000).

The usefulness of the constructivist's paradigm is a more oriented analysis of more social actors in cyberspace, for example collectives such as Anonymous.

Relevance of Studying States Behaviour

Studying the behaviour of States during cyber conflicts today is of paramount importance because cyberspace holds within itself a series of characteristics, already outlined above as well as throughout this dissertation, especially in the last chapter, that prove policymaking difficult. The comparison with cybercrime it is sort of natural. Together with cyber conflicts the difficulties lie in the challenges posed by technology concerning the access to the data useful for any investigation concerning the cyber realm. Data must be intercepted or collected through third parties, most of the times of private nature and sometimes in other states, due to the international nature of cyberspace. The first issue requires the technical and juridical capability of law enforcement bodies to acquire and analyse the data, while the second one a high degree of international cooperation (Eriksson and Giacomello, 2007).

Another similarity with cybercrime arises despite the fact that cybercrime concerns activities that are already considered criminal within a different realm. The problem at hand here is the difference in the various jurisdictions in different sovereign states therefore, without an international consensus on what constitutes an act of war in the cyber realm nor even a hostile attack, every States reacts independently from case to case. It is straightforward that such independent act could constitute rightful retaliation for the State that assumes that has been attacked but an unlawful act for the State that suffers from this retaliation. Indeed, assuming that the State that has been retaliated against did really attack first, it could defend itself using the uncertainty of attribution that, without any sufficient evidence, could not spur consequences apart from public accusations from the attacked State. However, as outlined in the third chapter of this thesis, evidences that can back retaliation, such as economic sanctions, can be produced but through a tricky process, such as penetrating the enemy's systems first.

Literature has suggested that international solutions and cooperation among states is the best way to tackle contemporary problems such as the consequences caused by cyber attacks (Slaughter, 1997), even when these solutions go beyond the power of the State in order to limit its exercise of power (Ikenberry, 1996). On the one hand, international institutions include a great number of States including the ones that are beginning to acquire cyber capabilities and those which do not possess any. For this reason, international cooperation useful for reducing costs in the process of developing norms, and assures compliance and transparency and completeness of information (Eriksson and Giacomello, 2007). The problem of the lack of information is especially stressed by the exchange of relations in the digital era, and for this reason is international cooperation is needed to draft and propose norms and regulations, and principles and guidelines of behaviour (*Ibidem*). On the other hand, however, international cooperation in cyberspace has to tackle a great issue. The participation to international cooperation and the building of international institutions are still subject to domestic policy and competition among States and these could use international institutions as means to their ends when useful but avoid referring to them when their objectives in national policy are different (Goldstein, 1996, Slaughter, 2000). This issue is of paramount importance because if international cooperation and international institutions are needed, there is still the doubt that States, or some States – namely the most active and prolific in the cyber realm - are content with the lack of international regulations. In light of this, cyberspace as an unregulated realm in which to project power and cause damage without international accusations and with minor State-to-State consequences is a comfortable option.

Hopefully waiting for some international regulating framework, for now it is necessary to study what is the current framework in which States operate, their behaviour in cyberspace - understood as means and ends - and the perilous consequences that these means could produce, in order to better grasp the situation and be able to tackle this issue more carefully, despite the low number of case studies at hand.

Relevance of Studying and Emerging Field

By now, stating that technology has changed our daily habits and the fact the we live in an interconnected world is a double-edged feature that brings a lot of advantages as well as brand new types of security issues, seems like a consolidated notion.

However, reality seems to tell us the contrary. Despite many countries are establishing new military cybersecurity strategies, general awareness on the real risks that come from the lack of cybersecurity measures is still low. The reason behind it could be that the progresses made at government, military and infrastructure levels are too fast for society to follow, despite it massively enjoys the fruit of such increasing dependence from the IT infrastructure. Many economies are becoming ever-dependent from cyberspace because today is a pivotal element to exponentially better businesses, economy, military, social and political life in a short period. Companies are now connected at national, regional and international levels more than ever, thus opening to new markets and possibilities that they couldn't reach before; the military exploitation of cyberspace is confirmed as a mean to project deterrent force postures, and it is useful also for weak states, because it enables them to fight an asymmetric war against countries that are military (in the kinetic sense) stronger; it's easier for governments to reach a larger portion of the population, for example rural areas. However, with a scarce culture of security, the major the internet penetration in civil society, private sector, government, critical infrastructures, businesses and military, the major the holes that enemies could exploit in order to cripple those emerging economies, steal military secrets and personal data, and try to cause malfunctions in critical infrastructures, posing a risk for all society.

Given the fact that cyber conflicts could be considered a new way of waging warfare, the implications when a new kind of conflict arises are many and worthy of attention. The focus of the research is placed on what could be considered an understudied aspect of this particular subject, that is the behaviour of states that engage in cyber disputes, meaning how they retaliate and, possibly, escalate. After the end of the Cold War, when escalation was a central topic for both the academic and the policymaking environment, the importance of this subject met a sort of a downfall. During the past twenty years, new conditions began to affect the same principles and the same issues around which the "classical" concepts of escalation revolved around. The rise of new actors, new security problems, and new methods to wage war brought the discussion back to the table (Manzo, 2011). Technology, in the form of exploitation of cyberspace for hostile purposes, could be considered the driving factor for a renewed need to study escalation. Cyber operations do indeed affect the modern security environment being disruptive enough to be an opportunity as well as a concern for leaders and policy makers. The opportunity consists in a new, less destructive way for waging limited disputes, and for weaker states this is also an opportunity to develop cyber capabilities in order to being able to retaliate and escalate if they are attacked, and also to amplify their chance of deterrence. In a spectrum of potential weapons to use during an escalation, at one end we would surely have nuclear weapons and at the other end would find cyber weapons, that are definitely less apocalyptic but also less costly, and therefore more attractive, given also the high level of dependence from IT system of the majority of countries around the world (Morgan, 2008). This dependence is perceived as a weakness but also as an opportunity to attack, retaliate and also to escalate in a conflict.

Another reason for adding bricks to the literature on cybersecurity is the rapid pace with which this field is expanding. New malwares, new more complex cyber weapons and new State actors appear on the news more rapidly than before. On the one hand, it means that new States are acquiring cyber capabilities or that States which their cyber capabilities were already common knowledge acquired new cyber tools. On the other hand, it means that academia has to analyse these events swiftly in order to keep up the pace. This necessity helps also policy maker and analysts for the development of better solutions and measures against these hostile cyber operations. Furthermore, it helps other scholars having a foundation upon which building further studies. For this reason, some scholars tend to be more conservative than others. This tendency could be found in the most general debate that is the existence or not of what many call "cyber war" (Liff, 2012; McGraw, 2012; Rid, 2012; Stone, 2012; Junio, 2013) but also regarding more specific aspect of conflicts in the digital realm, such as cyber weapons (Collins and McCombie, 2012; Farwell and Rohozinski, 2012; Rid and McBurney, 2012; Lindsay, 2013). As we have seen, the problem of cybersecurity is not limited to the last ten years, but taking a look to the literature on "cyber war" and "cyber weapons" it is straightforward that what made a lot of questions and the desire to debate over possible answers arise was indeed Stuxnet. Stuxnet was described as a revolutionary cyber weapon, the next step in the revolution of military affairs, or RMA (Collins and McCombie, 2012; Lindsay, 2013). Indeed, Stuxnet was the first malware that was not bound to cyberspace alone but was able to provoke consequences in the physical realm and this characteristic made it a game changer. Not only, it was proof that a State was able to acquire a capability that was, until that moment, a product of science fiction. However, the reality of the new kind of weapon failed to stick completely and for many conservatives Stuxnet remained a one of a kind tool, dubbing cyber attacks to industrial control systems (ICS) as "provocative claims" and attacks by State-sponsored groups as "a nightmare scenario" (Lindsay, 2013). Such view failed to grasp the reality and the pace of the evolution of cyber tools, both horizontally (number of States) as well as vertically (weapons sophistication), and was entirely disproven in only three years with the Russian CNA against the Ukrainian electric grid (Lee, Assante, Conway; 2016) and the rise in the number of State-sponsored cyber attacks in the last years. It is fair to say also that those proponents of the cyber revolution fearing a "Cyber Pearl Harbor" or a "digital 9/11" (Lindsay, 2013) are not exaggerating their claims as far as the current situation is concerned. That is to say that the possibility of a highly destructive CNA in terms of current capabilities of States, due to the fact that today there are at least two States capable of causing physical consequences through cyber means should not be dismissed. On the other hand, the probability of such an attack are low because it is more likely that such destruction could be seen as an act of war, and such a sophisticated attack could be more easily pinpointed to a particular State. What this means is that, even though such an attack is unlikely it doesn't hurt to use that scenario as something to fear and protect against. The real mistake is – on the other end of the scope – downplaying the perception of the damage caused by other kinds of CNOs. Conservative thinking could lead to underestimate the damage also of less sophisticated cyber operations for example by assuming that not every intrusion could lead to a huge loss of valuable data (Lindsay, 2013). It is easy to formulate such thinking but truthfully reality tells us another thing, meaning that, taking the Chinese incursions as an example we could see that the information has been more than valuable due to the fact that they were able to start a project on their fifth generation jet fighter. It is very easy to underestimate or even downplay the threat of cyber weapons but this is wrong for two main reasons. First, it's a strategic error. That is to say underestimating a potential enemy. Second, strictly linked to the first one, is that such an underestimation could lead to a lack of incentives as far as security is concerned.

The fear of a powerful enemy that could potentially acquire and use the information for malicious purposes should give birth to a feeling of need for better national cybersecurity. For these reasons, this dissertation will also provide a comprehensive overview of cyber weapons, in order to assess how they work and which consequences, even potential ones, they could produce. This provides a foundation to further studies avoiding exaggerating claims as well as conservative ones.

Having assessed the issue of cyber weapons and their diffusion, using the realist paradigm, we are able to tackle problems such as the relationship between offense and defence and deterrence. Indeed, with the spread of States possessing cyber capabilities and without regulating frameworks, States have to provide to defend themselves and to deter other State actors from attacking. Defending in cyberspace is not an easy task mainly because it costs more both in technical as well as financial terms because if an attacker has to focus on one – maybe unknown – vulnerability the defender has to perform a full-range of activities to protect its systems (Libicki, 2009; Kesan & Hayes, 2011, Lindsay, 2013). Furthermore, defence is a coordination game that encompasses software developers, vendors, private actors owning critical infrastructures and lack of public-private partnership (Lindsay, 2013). Even the literature on cyber deterrence is united in saying that cyber deterrence is a difficult task mainly because of the attribution problem (Libicki, 2009; Goodman, 2010; Lynn, 2010; Kesan & Hayes, 2011; Lindsay, 2013). The choice of using the State and not other actors as the unit of analysis for this dissertation connects directly to the fact that proponents of the cyber revolutions always refer to "three conventional wisdoms" that are asymmetry, offense-dominance, and deterrence failure (Lindsay, 2013). If offense dominance and deterrence are tackled in the first chapters of the thesis, the last one is all about asymmetry. Asymmetry influences all cyber conflicts whether it is present or it is not. Literature is almost unanimous in saying that cyber weapons are weapons of weaker States, because thanks to the nature of ICT technology and the problem of attribution it is inferred that exploiting systems' weaknesses and the powerful States' dependency from ICT could balance conventional military power (Chilcoat, 1998; Arquilla and Ronfeldt, 2001; Cordesman, 2002; Erikson and Giacomello, 2007; Cornish et al, 2010; Lynn, 2010; Geers, 2011; Philips, 2012; Clarke and Knake, 2014). Reality, however, is different and this thesis challenges this assertion directly. On the one hand it is true that cyber capabilities have been acquired or are bound to be acquired by an increasing number of States and that a smaller state, for example North Korea, is able to attack a more powerful State, for example the United States. Going even lower, brief incursions to steal sensitive data, or even reconnaissance incursion to analyse the digital environment of one particular State systems are becoming common practices. However, very powerful cyber attacks are prerogative of certain States only, for example Stuxnet or CrashOverride[1], but also attacks with "less" impact but still significant, like Shamoon or WannaCry. The reality of asymmetry does not end here, because if on the one hand only some States acquired a certain level of cyber capabilities, on the other hand within this "cyber club" there are differences in power. Literature suggests that there are barriers to the acquisition of cyber weapons for weaker actors and if they use them successfully nonetheless, are prone to face more cyber damage in case of retaliation from a more powerful actor (Lindsay, 2013). Furthermore, cyber attacks have not only a military nature but also a political one, as outlined in the dissertation, both as a signal of power in the attack, as well as a display of defensive and deterring capabilities, also through the political discourse, reducing threats of aggression and further retaliation (*ibidem*). This thesis tackles exactly this problem, asking what is the behaviour, what are the patterns during cyber disputes between States, that could lead to retaliations and escalations.

Cyberwar is a new event and cyberspace is a new domain of conducting warfare, it is still in its inception and it is constantly changing, so there are few elements to build up a brand new theory. Some critics still argue that cyber warfare never happened before and that is not going to happen (Rid, 2012; Liff, 2012). That could be true, but the main problem is that these same critics do not rely on a theoretical framework that seeks to find and explain causal mechanisms of warfare (Junio, 2013), in order to evaluate their absence, and neither they propose one to back their statements. Evaluating those mechanism is of paramount importance in order to dissipate the "fog of war" created by the uncertainty of attribution and the difficulty of calculating the exact behaviour of cyber attacks, that could lead to inadvertent cyberwar by retaliating against the wrong countries or erroneous calculations in the costs and consequences of an attack (Junio, 2013). Defying to acknowledge cyber conflicts and disputes as a new influencer on the war and peace scale only because it is different from classic interstate conflict risks to avoid the analysis of an important factor for international security (Kello, 2013).

The importance of cyberspace as a new dimension of warfare where to wage war is given mainly by three factors, that are namely the power of cyber weapons, the difficulties incurred in cyber defence, and the issues that stem from strategic instability (Kello, 2013).

[1] See chapter III

The increasing numbers of cyber arsenals around the world is faster than the number of doctrines produced, limiting the analysis of the threat (Kello, 2013). For example, in 2015 the US increased fivefold its budget on cyber expenditure (from 1 billion dollars to 5 billions), China increased its budget for cyber operation in 20-30%. Russia in 2014 began a hunt for IT experts in order to create a dedicated, national cyber army with an initial investment of 500 million dollars (Gerden, 2014) and Iran, after the events of Stuxnet that will be described in the third chapter, increased its cyber expenditure twelvefold since 2013, giving the Islamic Revolutionary Guard Corps (IRGC) an annual budget, for cyber operations only, of almost $20 million (Paganini, 2015; Bertrand, 2015). Furthermore, the military exploitation of cyberspace fits both the descriptions of the two triggers that cause a technological shift given by Lieber (Lieber, 2005) given the fact that technology advancement brought to the discovery of a new, less costly weapon, namely cyber attacks, and also it betters already existing Command and Control systems and synergy with the other domains of warfare. However, this research will focus on disputes in cyberspace treated as a domain per se, detached and independent from the other domains of warfare.

Research Questions and Research Design

This research aims at researching dynamics in states' behaviour during cyber conflicts, concentrating on those elements that bring state to deploy cyber weapons against other states, and the mechanisms that bring attacked states to respond, and how. The research questions of this dissertation that wants to investigate in the aforementioned matter are two and tackle directly *when* and *how* states employ cyber weapons and engage in cyber disputes. The questions are the following:

Q1) Is there a particular context within which cyber disputes take place?

Q2) What are the elements that influence the mechanisms of retaliation, and possible escalation, during cyber disputes?

The hypotheses that this research brings forward in order to answer these questions are:

H1) Cyber disputes are more likely to begin and end within contexts of political, military, diplomatic tension between states as an alternative mean to signal power and force posture without recurring to physical measures, intended as political accusations, economic sanctions and military interventions.

H2) Cyber disputes are likely to be influenced by two different elements, the first being symmetry in power between states engaged in cyber conflicts, the second being self-restraining mechanisms which take place in substitution for the absence of international law applicable to cyber operation, as well as enduring uncertainty around the deployment of cyber weapons.

In situations of asymmetry the most powerful state, both as an attacker as a responder to a previous attack, will show enough power to acquire a position of escalation dominance. Against this position the counterpart will be very likely to retaliate but in a less powerful way, de-escalating the intensity of the conflict, or surrender to the attack and avoid retaliating. In a situation of symmetry, the dynamics of attack are expected to follow a tit-for-tat movement, without increasing the intensity of the conflict. This, supposedly, to avoid an escalation into the physical realm that would easily create an impasse or a prolonged crisis, given the equality in power.

In order to test the hypotheses it is necessary to establish the variables to analyse. To test the first hypothesis the dependent variable is the presence of a cyber dispute while the independent variable is the presence of a situation of political, military or diplomatic tension. To test the second hypothesis, the dependent variable is retaliation against cyber attacks tested against two independent variables, namely symmetry in power[2] and self-restraint.

To further corroborate these hypotheses, the dissertation is divided in three chapters, the first two are theoretical chapters and the third one utilises the qualitative method of case study in order to test empirically whether the hypotheses are true or false.

The first chapter is a reorganisation of classic International Relations concepts and literature that adds important elements to the research. The first section of the first chapter describes and justifies the use of the state as unit of analysis. This is important because, compared to physical military disputes, hostile operations in cyberspace could also be

[2] The variable of power, in order to assess symmetry or asymmetry is derived from the 2017 Global Fire Power index 2017 which is a power index assessing world’s militaries based on more than 50 factors. Indeed, following the index US, Russia and China could be considered as symmetric, and the US military is definitely asymmetric compared to Iran and South Korea.

conducted by private individuals and private institutions. The first sections explains also why states employ cyber weapons as tools of statecraft. Connected to this, the main question to be asked is when hostile actions in cyberspace become state activities. The case studies analysed within this research and also future analysis could rest on the fact that when one of the two actors involved in a cyber dispute is a state, then the event requires nonetheless a state response and could be subject to international law. Even in those cases when CNOs are perpetrated by individuals, whether they are state-sponsored or not, the normative notion of due diligence requires state to assume responsibility, due to its accountability for the actions of its citizens and the obligation to avoid that private actors could cause harm to another country. Then the second sections explores how the increasing sophistication and number of CNOs, the increase in the number of states adopting or developing cyber weapons as a military tool and the general difficult in employing cyber defence worsens the renewed concept of the security dilemma. The fourth and the fifth sections of the chapter tackle the classic concepts of retaliation, escalation and deterrence in cyberspace, declining them into cyberspace. These section helps in the empirical analysis of case studies allowing a comparison to the mechanisms of kinetic retaliation, escalation and deterrence to their counterparts in cyberspace. Being able to trace elements of difference and similarities between the two domain helps in better understanding the mechanics of cyberspace, allowing a solid theoretical base for contemporary as well future analyses. Important elements tackled by this chapter are the importance of surpassing the problems in the attribution process, also as a mean for deterrence, and also the important notion of the lack of an international normative framework to regulate state's behaviour, accountability and responsibility in cyberspace. For this reason, policy suggestions are presented to help policymakers fill a normative void in this matter. Furthermore, the lack of a normative framework in this case provides help also in defining why states resort to self-restraining mechanisms.

The second chapter enlists thoroughly and also technically the options that state could possess to conduct CNOs as well as responding to them. Due to the fact that the object of the research are states engaging in cyber dispute, there could be no hostile cyber activity without cyber weapons. The first section tackles the problem of defining a cyber weapon using the Tallinn Manual as the main support, to which add technical literature as well as all the literature in International Relations on cyber weapon. The second section describes how the functioning of cyber weapons could be divided in three main elements: the presence of vulnerabilities, the exploitation of given vulnerabilities and the subsequent propagation into the systems, and the delivery of the payload. The third section will outline what are the objectives that the employment of cyber weapons could achieve and how different objectives call for two different kind of CNOS, namely Computer Network Attacks, CNAs, and Computer Network Exploitations, CNEs. Understanding this difference is essential in order to analyse the intensity of attacks and retaliations in cyber disputes. The fourth section provides a taxonomy of cyber weapons, distinguished in two families: malwares which include viruses, worms, trojan horses and spywares, and blended threats, such as Denial of Service and Distributed Denial of Service attacks, and Advanced Persistent Threats. The chapter is concluded with a final section on suggested policies that call for a regulation of cyber weapons due to the fact that, as of today, it remains an improbable exercise. Also in this case, the lack of a regulating framework as far as cyber weapons are concerned added further motivation to states resorting to self-restraining mechanisms.

The third chapter is the empirical analysis of case studies, which sum up all the theoretical parts in concrete cases. The unit of analysis are dyads of major cyber-capable States which are in different context of symmetry. The choice of qualitative case studies stems from the fact that CNOs remain covert operations and quantitative information available stem from secondary sources bringing the risks of not being representative enough, lowering the level of confidence by doing a large-N analysis. Data completeness is a necessary requirement for doing a quantitative analysis, and this requirement could not be fulfilled with the information available to general public about cyber activities. Trying to use what the data available by transforming it into a database poses severe limitation as far as reliability is concerned, invalidating the analysis. Furthermore, among cyber weapons it is difficult to establish a base for an equivalence principle (intentions, sophistication, consequences, and how to measure them individually) and grading them on Likert scales, for example, and the same applies to targets. Given these methodological shortcomings, case studies better suit the analysis of cyber disputes. Furthermore, aware of the low number of case studies at hand, and the largely understudied characteristic of phenomena like retaliation and escalation during cyber conflicts the precondition for this study is that it can be considered as a plausibility probe (Eckstein, 1975). The main goal of the thesis is to probe the theory in order to assess that the theoretical construct taken into consideration is worth studying both now as well in the future. To establish the usefulness of this approach, empirical instances with enough consistency of data and reasoning must be found in order to be justifiable in a precise form and thorough testing (Levy, 2008). To do that, the method and logic of "structured, focused comparison" is chosen (George, 1979; George and Bennett, 2005). The reason behind this choice is that historically this methodology was used to produce knowledge of important foreign policy problems such as deterrence, in such way that the cases explanations could be drawn into a broader and more complex theory (George and Bennett, 2005). This method is regarded as structured because the research questions reflect the objective through a systematic comparison of the findings of the cases and it is focused because it takes into consideration only certain aspects of the cases at hand (*Ibidem*). Furthermore, this methodology combines well with the plausibility probe as more cases can be added to further the study.

The case studies are three, the first two analyse asymmetry, the third symmetry. Every chapter would provide a description of the geopolitical context, the actions carried out, the retaliation and the analysis thereof. Namely, if it was vertical, horizontal, de-escalating, stable on the same level of intensity of if the state surrendered. The first one describes a powerful state attacking through cyber means a less powerful states, namely the Stuxnet case, involving the US and Iran. The second one concerned the opposite scenario, a low power state against a powerful state, that is to say the Sony Hack, involving North Korea and the US. The third one is a situation of asymmetry and the cases involve the three major powers in the world, US, China and Russia, divided into two dyads US and China, and US and Russia. For further completeness of information, and not falling into a case-selection bias, other cases are tackled by this research, such as Russia versus Estonia 2007, Georgia vs Russia 2008. Also a recent case such as Russia versus Ukraine is mentioned as an asymmetry case in a peculiar geopolitical context, but given the incompleteness of data due to the event being very recent, it was not chosen as a case study, but it's briefly analysed nonetheless.

The aim of this research is to pushing forward the current state of the studies on International cyber security. The main force driving behind dissertation is the will to go beyond the same questions that literature on cyber security tackled by an international relations point of view is stuck in. The impossibility of attribution, the semantic explanations behind the Clausewitzian concepts of "war" and the whether scholars should use the word "weapon" or "tool" when referring to malwares employed by states, are important but secondary questions. States are acquiring sophisticated cyber capabilities that could be offensive and also disruptive. In order to be ready to analyse those event, but more importantly to try to prevent them and pushing towards a legally binding normative framework for state cyber activities, an analysis of states behaviour during cyber disputes and policy suggestion like the ones contained in this dissertation could be considered an important added value to the field, serving also as a basis for future researchers, given the increasing importance that cyberspace is gaining.

Coming to acknowledgements, I would like to thank, first and foremost, my supervisor Prof Andrea Locatelli who has been such a great mentor to me. I would like to thank him for supporting and encouraging my work and for allowing me to grow as a researcher. His patience, knowledge and his advice on both my research as well as on my career have been crucial and priceless. Besides my supervisor, I would like to thank the two reviewers of the first draft of this dissertation, Dr Francesco Moro and Prof Giampiero Giacomello for their insightful comments and encouragement. This work has also benefited from a visiting research period at the Department of Politics and International Relations' Cyber Studies Programme of the University of Oxford. Therefore, I would like to thank my academic mentor across the Channel, Dr Lucas Kello, as well as Jamie Collier, Florian Egloff, and Max Smeets for their invaluable feedbacks, insights and engaging discussions.

As great works cannot exist without great support, I am deeply grateful to Dr Stefano Costalli who encouraged me in pursuing my research interests and without whom I would not have started my scientific journey. I would also like to thank all the research group working with the International Relations chair of the Università Cattolica del Sacro Cuore, namely Prof Vittorio Emanuele Parsi, and Dr Enrico Fassi and Dr Antonio Zotti for having shared with me their office and their time, tolerating my presence for three long years.

Undoubtedly, these years would have been so much harder without my fellow PhD students. A special mention goes to Federico Bianchi, Iraklis Dimitriadis and Gianluca Pozzoni for our discussions, for sharing with me this journey - also in difficult times - and for all the fun we had in the past years as my friends, more than colleagues. Support and friendship are not to be found only within academia but also outside of it. For this reason, I would like to thank Ivan Latagliata and Ewan Sirtori for sharing laughters, music, and coping with me on a daily basis. I am also greatly thankful to Francesco Bonaccorso and Marco De Benedictis with whom I share roads, Negroni, sadness, and happiness. Thank you for being always there for me.

A special thanks goes to Roberta, who lovingly encouraged me, supported me and put up with me in the most difficult period of my research. I cannot stress enough how her presence beside me was invaluable for helping me going through all the hardships of the past year, not only research-related.

Last but not least, I would like to thank my family. Words cannot express how grateful I am to my mother and my father for all of the sacrifices that they have made for me, for the continuous support and encouragement in every aspect of my life that made me who I am today. Of course this work is dedicated to my grandmother who fostered my curiosity and knowledge since an early age.

Chapter I: Conflicts in Cyberspace: State Responsibility, Escalation and Deterrence

Describing the new domain of warfare: Cyberspace

Cyber insecurity has become a global and growing problem that affects all layers of society, from the government and military apparatus to the civilian common user. This happens due to the diffusion of IT technology that has been weaponised in the last stage of its history. It is interesting to note how “the internet” and all the IT-related technologies that are now of common use and that we can call, collectively, *cyberspace* began as military projects. Indeed, the internet was born in the United States as a reliable military infrastructure in case of a nuclear attack coming from the USSR (Leiner et al, 1997). But it isn’t limited to that, let’s think about, for example, to GPS (Global Positioning System) or ACARS (Aircraft Communications Addressing and Reporting System), among other communication technologies that rely on the electromagnetic spectrum and, therefore, belong to the cyberspace (Libicki, 2009). It should also be pointed out how there is no official definition of what constitutes cyberspace, and indeed many definitions could be found, also in different times. This does not constitute a problem, on the contrary, reflects the constant evolution of how cyberspace is perceived by different actors, such the general, public, policymakers, militaries and scholars.

To exemplify the degree of difference that sits between definitions of cyberspace, the Oxford dictionary defines cyberspace as "the notional environment in which communication over computer networks occurs” (Oxford English Dictionary, 2017). The US Department of Homeland Security gives the following definition: “the interdependent network of information technology infrastructures, that includes the Internet, telecommunications networks, computer systems, and embedded processors and controllers” (DHS U.S., 2017). For the United States’ Department of Defence Dictionary of Military and Associated Terms instead: “a global domain within the information environment consisting of the interdependent network of information technology infrastructures and resident data, including the Internet, telecommunications networks, computer systems, and embedded processors and controllers.” (DoD U.S., 2017). One of the most recent definition, unofficial as it comes from scholars, that could be considered the most encompassing is the following:

"Cyberspace is a global and dynamic domain (subject to constant change) characterised by the combined use of electrons and electromagnetic spectrum, whose purpose is to create, store, modify, exchange, share and extract, use, eliminate information and disrupt physical resources. Cyberspace includes: a) physical infrastructures and telecommunications devices that allow for the connection of technological and communication system networks, understood in the broadest sense (SCADA devices, smartphones/tablets, computers, servers, etc.); b) computer systems (see point a) and the related (sometimes embedded) software that guarantee the domain's basic operational functioning and connectivity; c) networks between computer systems; d) networks of networks that connect computer systems (the distinction between networks and networks of networks is mainly organisational); e) the access nodes of users and intermediaries routing nodes; f) constituent data (or resident data). Often, in common parlance (and sometimes in commercial language), networks of networks are called Internet (with a lowercase i), while networks between computers are called intranet. Internet (with a capital I, in journalistic language sometimes called the Net) can be considered a part of the system a). A distinctive and constitutive feature of cyberspace is that no central entity exercises control over all the networks that make up this new domain." (Mayer et al, 2014).

While the composition of cyberspace (points "a" through "f") is as correct as it can get, the opening lines raise some doubts at least from a terminological point of view. What has been listed, including "eliminate information and disrupt physical resources" is not one of the *purposes* of cyberspace, but it is surely one of its possible applications. And here we arrive at the point of the weaponisation of cyberspace. Cyberspace is flawed, since it is constituted by man-made hardware and man-scripted software it presents vulnerabilities. These vulnerabilities in cyberspace have been exploited since its inception, from the hacking of the wireless telegraphy transmitter of Marconi in 1903 (Davis, 2015), through John Draper hacked the American phone company AT&T systems through a plastic whistle found in a package of cereals (hence the name Captain/Cap'n Crunch) in 1971 (Hafner and Markoff, 1995) and the MorrisWorm in 1988 (Davis, 2015) until the recent WannaCry ransomware in 2017. The exploitation of vulnerabilities has been used primarily for criminal purposes, meaning private actors searching for personal gain. Due to the increase in global connectivity and all infrastructures and service companies connecting to cyberspace, adapting the words of Willie Sutton, crime moved to the internet because that's where the money is. What changed is that at a certain point in history, states began using exploiting the same vulnerabilities (sometimes new and unknown ones, called *0days*) with better tools to reach political objectives. These political objectives could be: gain a strategic advantage, by hacking into servers where military operations plans are stored; gain a technological advance, by hacking into research and development agencies that work in close contact with the military and the government of a certain state, furthermore this could be used also to gain an economic advantage; gain a political advantage by hacking into personal devices and try to influence political elections by exposing confidential information about politicians.

Therefore, we could see how the evolution of cyberspace and - on par - its exploitation has completed a sort of "circle", started as a matter of state, then passed to private actors, and then again back to be a matter of state. This, however, does not imply that the threat of private actors has vanished, on the contrary, it has constantly risen. Indeed, cybercriminal activities is still today a great threat, but it is a threat that concerns law enforcement agencies and, albeit there are spill-overs between the "cyber crime field" and the "cyber war" one, it is necessary to draw a line between the two. This, however, it is not a simple task. In this dissertation the subject of research is the state and its actions but given the complexity and the multitude of actors that inhabit and exploit the vulnerabilities of this environment it is difficult to address and to isolate the issues that concern states, and that's what the next paragraph is about.

What Does The State Do?

Cyber attacks between countries have increased in number in the past years (Symantec, 2017; Trend Micro, 2017; PwC, 2017). Nation states rarely use cyber attacks to physically disrupt machinery, the only recorded event was Stuxnet, discovered in 2010. Nevertheless, it proved that it could be done, and the possibility of similar future events is not to be excluded.

However, nation states generally use cyber attacks, or better, Computer Network Attacks (hereafter CNA) to carry out other types of Computer Network Operations (CNO), a category that also includes Computer Network Defence (CND) and Computer Network Exploitation (CNE). Given the fact that most of important information is conveyed through the aid of computer networks, and given the fact that knowing sensible information of an enemy, or being capable of degrading its network is relevant strategically because they tend to give information superiority (by denying it to the enemy), and consists a form of power, CNO are increasingly used by nation states.

CNO are not to be imagined only as attacks directed to the principal centres of power, i.e. the Pentagon, reality tells us that nowadays the main targets are "secondary" centres of power, governmental offices, if not corporations. A nation state actor may opt to target a local government network as opposed to that of a central government entity as the local network poses an easier and less complex target. Local governments likely lack the resources for stringent network security and monitoring, making them a technically easier target for threat actors. However, despite the relatively lax network security, local government networks also likely contain potentially valuable information for nation state threat actors, including insight into major industries operating within their jurisdictions, as well as personnel and financial data.

One issue to underline and explain is the one about the nature of different networked attacks. Many authors (Morgan et al, 2008; Kramer, Starr and Wentz, 2009; Libicki, 2009) tend to distinguish among the different categories of threats depending on the motivation behind the attacks, since nation states can engage in cyber espionage, as well as in information war, as well as in hacktivism. All of these different categories have different ends but one common actor that is to say that between one attacker and one defender they involve at least one state For this reason the main question here is: when could we classify a hostile act in cyberspace as a matter of state?

State Responsibility

In order to address the question of escalation, deterrence, and the applicability of international norms, we should define the cases in which a state is responsible of an hostile act carried in cyberspace.

Pinpointing responsibility in cyberspace is difficult for various reasons, the first one is the problem of attribution. This issue could be defined as the absence of certainty in the identification the culprit of a given CNO. Indeed, hostile cyber activities do not travel in a visible way like ICBMS or the movement of troops, but at the same time, like with kinetic weapons, intelligence and forensic units exist in order to mitigate the problem. It is important to underline at this point that while some techniques are available to circumvent the problem, none proved successful in solving it, as outlined later. The second problem is defining state responsibility in a solid framework. As for now, the start of hostile CNOs has not been declared, ever. They still are covert operations that rely on the absence of a normative framework and lurk in the grey area between simple espionage activities and hostile acts, not to mention acts of war.

Furthermore, states are known to have employed nationalists group as workforce to conduct cyber attacks against other countries, for example in the case of the cyber dispute between Russia and Estonia in 2007 and regularly by China (Rid, 2012; Klimburg, 2011;). Both countries possess units within their national armies specifically employed to perform CNOs but due to the lack of international regulations it is difficult to frame those action to elicit an official response. This problem is stressed even further in the case of state-sponsored cyber attacks. For this reason, state responsibility shifts in another normative grey area. Indeed, state-sponsored actions are difficult to frame even when they are kinetic, for example in cases of state-sponsored terrorism. The problem, within the framework of international law, is how to deal with state-sponsored attacks, namely identify cases where to held a particular state responsible. According to existing law literature on the subject, there exist two competing standard on how to deal with state-sponsored cyber attacks in order to determine state responsibility (Kulesza, 2009; Shackleford, 2010; Tsagourias, 2014). The first refers to the International Court of Justice (ICJ) *Nicaragua case* (Nicaragua v. United States, 1986), and to the ICJ *Application of the Genocide Convention* (Bosnia and Herzegovina v. Serbia and Montenegro, 2007) where it was applied what is called *effective control,* that is to say when the State directly controls state actors or official organs (Kulesza, 2009; Shackleford, 2010; Tsagourias, 2014). The second one refers to International Criminal Tribunal for the Former Yugoslavia (ICTY) *Tadic case* (Prosecutor v. Tadic, 1995) where it was put in action another kind of control called *overall control,* which describes a situation where the state has a role in the organisation and the coordination of a group and supports its actions (Kulesza, 2009; Shackleford, 2010; Tsagourias, 2014). Applying the first standard, namely effective control, to CNOs is straightforwardly impossible, this for different reasons: first of all there must be evidences beyond *any* reasonable doubt that a government is linked to an hostile act in cyberspace carried out indirectly by a third party. On the one hand, this could be done in kinetic conflicts, for example when a group is using heavy weaponry that should belong only to national armies or is using weaponry in such quantities possible only through a financial aid coming from a state, or through the interrogation of individuals belonging to this group to exert information linking this group to the government (Tsagourias, 2014); on the other hand, it is very difficult to conduct the same kind of investigation in cyberspace because, with due exceptions (such as Stuxnet or some Advanced Persistent Threats [3]), cyber weapons until a certain threshold of sophistication are available both to the private as well to the public sector. Furthermore, under this standard, these groups under the subordination of the state must be organised in a military-like fashion, that is to say there must be a hierarchy and it must have the capability of launching coordinate attacks (*Ibidem*). If cyber militias have some degree of organisation it is very difficult to prove the existence of a pyramidal hierarchy, due to the high number of members that could vary during the execution of hostile cyber activities. Moreover, proof of a *direct* link between state and group could be easily bypassing in two ways: first, although as it will be explained later it could be done, finding the exact location and the exact identity of a mastermind (or multiple culprits) is very difficult, therefore avoiding the identification of a physical person; second, every digital communication between the group and the state could be easily deleted. The last instance against the application of the effective control in cyberspace is linked to the *Nicaragua* ruling. Here it was made the distinction between "most grave" and "less grave" acts of force that, if applied in cyberspace, would render many cases of state-sponsored hostile CNOs pointless given that DDoS, defacements, and other kinds of malwares do not easily surpass a certain threshold that is temporary disruption (Shackelford, 2010).

Instead, applying the overall control could lower the strict threshold posed by the effective control standard, therefore giving a little bit of more room to allow the accountability of states in cases of state-sponsored CNOs. In this case, it is not necessary that the state employs direct control over the group, but it is sufficient that it exerts a general influence on this group, and this influence could be limited to financial aid and help in coordination and planning (Tsagourias, 2014). In reality, both of the standards would not help in cases such as Estonia 2007 and not even in Georgia 2008.

What is needed – and that is a theoretical construct that is ever-returning in this dissertation and it is recurring in the cyber security environment – is to treat the virtual domain as something different to the kinetic one, because it is subject to internal laws that are, to a certain extent, different to the laws of physics. Data travels in different ways and in different times compared to physical objects but, at the same time, actions performed in the virtual realm have consequences in the physical world, whether direct or indirect, and for this reason they must be subject to the laws of man. To do so, however it is important to understand the "physics of cyberspace" in order to apply international laws and norms. In

[3] Duly described in chapter II.

this case for example, it is necessary to lower the standards of the “burden of proof” and the “standard of proof”. Starting from the latter, it is important to underline that, compared to the jurisprudence of cases that happen in the physical realm, in cyberspace it is difficult to produce evidence of sufficient quality and quantity to make an accusation “beyond reasonable doubt”, but at the same time the standard of proof must be reliable enough to produce an accusation “sufficiently certain” in order to avoid cases of erroneous attribution (Roscini, 2014). Alas, here comes into play a sort of Catch-22 paradox. The technical evidence that should provide proof in court is, in many cases, classified for security reasons when the forensics has been carried out by security agencies, or it is produced by third party security companies, that do not meet international Court favourably (*Ibidem*). Furthermore, it is argued that in cyberspace the burden of proof should shift from the the victim to the country where the attack is supposed to have originated (*Ibidem*), that is to say it is the alleged attacker that should defend itself providing proof that the CNO did not start within its territorial boundaries and that it has no involvement in it. These two proposed standards seem very loose and difficult to apply considering the normal application of international law, but it could be argued that they are based on a realistic assumption that has two precedents in the physical world. Indeed, the state has also responsibility when, within its borders, fruitful condition to carry out hostile acts against another state are created not only when it directly carries out such acts or supports them (Tsagourias, 2014). The first example is the *Iranian hostage case* of 1980 where, albeit there was not a degree of evidences enough to link the action of Iranian citizens to the Iranian government, the latter did not respect the Vienna Convention on Diplomatic Relations of 1961 and the Convention on Consular Relations of 1963 according to which a state as the obligation to protect the embassies and all the individuals within them (Shackelford, 2010). In a situation concerning a cyber attack this would mean that if there is no evidence strong enough to link the attacking group to the government, then the state could be found liable of allowing an attack on another state’s infrastructure, therefore breaking international law (*Ibidem*). The second example is more recent, namely the attacks of 9/11, that spurred the American military action in Afghanistan and Iraq because both states failed in their duty of “due diligence” that is to say the responsibility to impede private actors to conduct actions that belong uniquely to a state (such as carrying out an hostile act against another state) or that go against the obligation imposed by international law to such state (Tsagourias, 2014).

State Power

Adopting the obligation to due diligence would lower the threshold according to which a state could be prosecuted for cyber attacks. Such a measure is strongly needed because, within the current normative state, states have limited room for action, despite their capability to produce evidence. However, this capability is strongly influenced by variable of power of the state, that in this case could be circumscribed to the so called “cyber power”. Defining “cyber power” it is not an easy task and, in reality, is not really important. The most important issue, in this context, it is that the definition of cyber power does not coincide with the variables taken into consideration in the the only Cyber Power Index that exists right now[4], namely the Booz Allen one. It takes into consideration variables that do not influence military use of cyberspace, such as national censorship, access to information and communications technology and its affordability, and economic and social context variables. On the other hand, Kuehl defines it as “the ability to use cyberspace to create advantages and influence events in other operational environments and across the instruments of power.” (Kuehl, 2009), and Nye defines cyber power as “the ability to obtain preferred outcomes through use of the electronically interconnected information resources of the cyber domain.”(Nye, 2010). These are broad definitions that could be considered generally true, but it is Nye that points out that what influences the most the power of a state in cyberspace are “1) the development and support of infrastructure, education, intellectual property; 2) legal and physical coercion of individuals and intermediaries located within borders; 3) Size of market and control of access; eg. EU, China, US; 4) Resources for cyber attack and defence: bureaucracy, budgets, intelligence agencies 5) Provision of public goods, eg. regulations necessary for commerce; 6) Reputation for legitimacy, benignity, competence that produce so power.”. The most important point in this context is the fourth one, that is to say that a state with a bureaucratic infrastructure that permits easily the flow of the command and control chain, capable of allocating enormous amounts of money to inject in research and development of cyber weapons and defence measures, and with specific agencies that could enforce in full this capability. Due to this fact it is easy to see why, in this context, defining clearly what “cyber power” means could be considered as pointless, it simply is military power (with all that it entails) declined in cyberspace. It is important to understand

[4] http://www.boozallen.com/content/dam/boozallen/media/file/Cyber_Power_Index_Findings_and_Methodology.pdf

that the power differential influences also means of escalation and deterrence, because it involves both military power in cyberspace as well as military power in the physical world. As for now, no cyber dispute gave rise to a spill over effect in the physical world - mainly because in that case states will be bound to international law - but, nevertheless, it could constitute a deterrence for hostile actions in cyberspace.

To better understand how could the difference in power come into play in cases where state responsibility is concerned, the aid of three examples could explain better the situation .

Private Actor Vs State Actor: Russian Hackers Vs Estonia

It could be the case that a nationalist group would start hostile actions in cyberspace against the governmental facilities of another state and this group would receive state support only at a later stage. Furthermore, the state - in this case Russia - would also deny cooperation in the investigation to find the individuals responsible of the hostile activity or activities (Schackelford, 2010) hiding behind a denial of guilt and the sureness of the impossibility of conducting international investigations. Moreover, it could be inferred that such a behaviour risks to exacerbate also the relations between two countries, country A that states that has been attacked by country B (a claim supported by the highest amount of evidence possible), and country B that has no obligation in admitting nor welcoming or performing internal investigation that could easily dismiss any of these accusations. The point is that there exist circumstantial evidences that could pinpoint to a particular state in case of hostile CNOs such in the Estonia case. The cyber forensics made after the DDoS attacks, which traced back to a person linked to Russian security services, and the political motivations that spurred the attacks, that is to say the removal of the Bronze Soldier, symbol of the fallen Russian soldiers during World War II, from the center of Tallinn, pointed to Russia (Traynor, 2007). Furthermore, Russia has been accused of parallel CNOs against Georgia, Czech Republic, Lithuania and Poland, and to have performed waves of CNOs against Estonia during Russian national holidays, and one of these waves coincided with a fierce speech given by Putin directed against Estonia (*Ibidem*). Of course, there is the possibility that the Russian government did not help directly the individuals that were carrying out the attack, but at the same time did nothing to halt the operations or to shed light on them. NATO intervened helping Estonia in the forensics by sending their anti-terrorist unit but at the same time declared that there was no ground to call for article 5 and collective defence against Russia. The issue here was that invoking article 5 could have led to military escalation between NATO and Russia and furthermore Estonia is dependent from Russia for energy supplies, namely natural gas (Maness and Valeriano, 2015), therefore the situation clearly states that Estonia did actually find itself in a situation of disadvantage as far power relations with Russia are concerned. In this case it is pretty clear that Estonia, despite NATO support, did not have the capability to retaliate against Russia and if it had, the escalation event in the cyber realm would have been unwanted. What could be learned from this case? As it was stated before, lowering the threshold and applying the "due diligence" standard could have helped in the process of blaming Russia and finding it guilty for having allowed hostile CNOs against Estonia, but reality shows that circumstantial evidences could have led to worse consequence, such as a deterioration in the relationship between the two countries to say the least, or an escalation in the kinetic realm that no one would have wanted, to say the worst. However, Estonia wasn't a passive actor in this dispute and it actually did strike back. Estonia did not retaliate with cyber weapons but it responded in the form of norms and laws. Indeed, after the attacks, Tallinn became the headquarter of NATO's Cooperative Cyber Defence Centre of Excellence and became the spearhead in the promotion of the regulation of cyberspace also by hosting the "International Conference on Cyber Conflicts" to which many major western powers attend to discuss norms of behaviour and propose laws in order to avoid escalation during cyber conflicts (Maness and Valeriano, 2015). As a victim of a foreign cyber attack without capabilities to retaliate in kind, Estonia chose to become one of the countries leader in the regulation of cyberspace.

State Actor Vs Private Actor: North Korea Vs Sony

Another example to talk about asymmetry in power as an instrument to influence state responsibility is what is known as Sony Hack. It is a very interesting case, as far as state responsibility is concerned, for two reasons: the first one is that a state actor, North Korea, attacked a non state actor, Sony Pictures (henceforth only "Sony"); the second one is how the standard of proof was produced by the United States, that is to say by previously hacking into North Korean systems in order to know for sure about their operations.

During November 2014, Sony Entertainment started being victim of strange cyber attacks on their systems. The first one was changing all the desktop pictures of more than 7,000 employees with a digitally altered photograph of a beheaded Michael Lynton, former CEO of Sony Pictures. Given the proportion of the breach, Sony's decision was to put all the systems offline and its executives were able to communicate only with old blackberries found in the storage (Cieply and Barnes, 2014). In the following days the executors of the previous attack, called Guardians of Peace, performed other attacks that wiped many hard disk clean and stole many confidential data, such as personal information about the employees and unreleased movies, and asked for ransom otherwise they would have published all the stolen information (Lee, 2014). Only at the beginning of December, the Guardians of Peace asked for the non release of the movie “the Interview” which contains a scene where North Korea’s leader, Kim Jong Un, is killed. Following other digital threats - mentioning also 9/11 - and the release of all the data stolen from the company, Sony withdrew the movie from theatres (*Ibidem*). With all signals pointing to North Korea, this act was seen as a giving in to their threats and became a matter of national security. Former president Barack Obama criticised publicly Sony’s actions and provided the company the intelligence aid of the FBI, which found that the tools used for hacking Sony were strikingly similar to other malwares used by North Korean hackers. Furthermore, Obama during the traditional speech at the end of the year, publicly blamed North Korea as the culprit of the operation and stated that the US would have responded to such attacks “proportionally” and “a place and time and manner that we choose.” (Sanger et al, 2014). Said that, the US retaliated against North Korea imposing sanctions and, allegedly, through cyber means, causing a temporary outage to their national web infrastructure (Nakashima, 2015, Strohm, 2015). As stated at the beginning of the paragraph, two things are worth underlining: in this case we have a state actor that attacked a private actor through a CNO and, maybe because Sony is within American borders, the US decided to respond, in an unprecedented way. The second important point is how the US produced evidence to be so sure about attributing the attack to North Korea. NSA’s TAO (Tailored Access Operations) unit was able to infiltrate North Korean system for intelligence purposes, and due to the poor cyber security operating on those systems they were able to find proofs about the operation against Sony (Van Der Walt, 2017). As far as state responsibility is concerned, it must be underlined that the US did not invoke any article of any international treaty, knowing it was useless due to the issues concerning the regulation of cyberspace also described in this chapter. Instead they acted on their own, sure about the attribution. A question that could arise could be “why didn’t the United States charge legally North Korea through the International Court of Justice, given the high degree of the standard of proof?”. The International Court of Justice would not have dismissed evidences even though they are obtained by illegal measures – namely violating state sovereignty, such in the *Corfu channel* case (Roscini, 2014). The main point to answer the question here is that the methods and the ways through which the NSA’s TAO entered the North Korean systems are supposed to be top secret and therefore they could not have been presented in court.

State Actor Vs State Actor: China Vs United States of America

The cyber skirmish between the United States and China does not have a precise starting date, and compared to the Russia versus Estonia case, or to North Korea versus the United States, it is different because it consists in multiple and repeated attacks. In China Computer Network Attacks and Exploitation began to be used as a strategic tool against the US from the late 90s, but they were allegedly carried out by groups of hackers that acted under a patriotic spur who called themselves "honkers" from the sound of "hong" meaning "red" (Osnos, 2012). The issue was two-faced. On the one hand, attacks performed by civilians questioned the control of the Party over the military use of cyberspace. On the other hand, redirecting the responsibility for these attacks to those independent group served as mean of justification when the Chinese government was blamed as the source of those Computer Network Attacks and Exploitations, for example like the Titan Rain espionage campaign in 2004 (Thornburg, 2005) and the incursions against the Pentagon in 2007 (Sevastopluo, 2007). Nonetheless, patriotic hackers served as a large basis of workforce carrying out sequences of Computer Network Attacks and Exploitation under the consent of the government (Lewis, 2013). This ambiguous situation further blurs the already very thin line that distinguishes an act of an independent civilian, willing to carry out a cyber attack that could bring even little benefit for his, or her, country, and a state-sponsored attack. For this reason, the Chinese government has been able to deny every accusation from other countries, mainly coming from the U.S., of systematically using cyber tools to carry out information gathering and reconnaissance attacks (Hjortdal, 2011). Both information operations and computer network operations are seen as powerful tools to bridge the military asymmetry with other countries, and specifically the United States. The evolution of networked systems, enabled China to enter American cyberspace without effort (Mulvenon, 2009). The exploitation of cyberspace for espionage purposes permits China to leapfrog the technological and economical gap with the US. For example, technologically speaking, China uses cyber incursions to illicitly obtain sensible data, such as 50 terabytes stolen from US defence contractors which contained various American military secrets, such as data on the B-2 bomber, F-22 and F-35 (Gertz, 2016). The Lockheed Martin F-35 Lightning II is one of the most sophisticated fighter jet currently existing, and the Chinese were able to acquire radar modules and engine blueprints and allegedly used that same knowledge to build their J-31.

What is particularly interesting, in light of the focus of this chapter is that, in 2014, the Department of Justice of the United States indicted five 3PLA (a special unit under the Third Department of the General Staff Department) hackers for the attacks against six American entities belonging to nuclear power, metals, and solar products industries, for economic espionage (Schmidt and Sanger, 2014). Espionage is usually justified by the United States but it seems that in this case China has crossed the line, as many commercial secrets are stolen from the United States to seek economic advantage. The indictment was a clear signal that China has to put some boundaries, limiting the scope of actions of its PLA units. (Schmidt and Sanger, 2014). Despite being charged, the members of the PLA Unit were not indicted in China, and straightforwardly not even extradite, because that would be unthinkable. Again, China denied the accusations, but with much surprise, after this historical event, another followed. In 2015 there was the first U.S. – China bilateral agreement on cyber issues: the Cyber Agreement stated that neither of the two governments will willingly support cyber espionage for commercial advantage (differentiating it from espionage carried out for national security reasons) (Rollins, 2105). This was a first, important step in history of cybersecurity and Sino-American relationship, but it was as important as it was pointless. Indeed, during 2016, state sponsored espionage operations originating from China successfully targeted U.S. government and companies (Gady, 2016). In this case, the two states are two major powers in the world, politically, militarily and also as far as cyber capabilities are concerned. Here the United States had enough evidence to indict five individuals - and it is necessary to underline that because backtracing an attack to individuals requires forensic skills that are in the hands of very few governmental agencies in the world - but even such a standard of proof that should have been sufficient to held China accountable was pointless in the absence of an international normative framework *and* in front of China, due to the fact that relative power, that for now seems to substitute accountability mechanisms, in this case is even. Furthermore, two great powers could use CNOs - other than means to an end - as signalling, in order not to resort to kinetic expression of power to show off military capabilities and start an escalation in the physical realm.

It is safe to conclude then, that addressing state responsibility in cyberspace is a pressing need that needs to be regulated. Indeed the debate about CNOs and state responsibility splits the opinions of the academics and scholars that study the subject. For example, the Tallinn manual - a non-legally binding set of norms written by a group of experts after the Estonia case, that *should* regulate state behaviour in cyberspace - in this regard dismisses all attacks in cyberspace that do not constitute physical harm (Schmitt, 2013). Another view sees cyberspace as *res communis,* that is to say a global common - like international waters, international airspace and outer space - where states do not have jurisdiction (Buchan, 2014). The proponents of this stance are evidently ignoring state sovereignty. The physical part of cyberspace resides within national borders and, furthermore, CNOs used for states versus states disputes start in one country and provoke harmful, regardless of the degree, in another one. Albeit a very stretched example, even an Intercontinental Ballistic Missile could fly in sub-orbital, unregulated, portions of the world, but it has a starting point and an ending one, and their use is strictly regulated. The main problem is that an ICBM is very visible and attribution simpler than CNOs. Nonetheless, as was stated before, cyberspace is a domain that is different from the physical one, and should be regulated accordingly. This first part of the chapter provided the problem of an unregulated environment, providing also real-world examples of CNOs performed by states and non-state actors under the connivance of governments.

Due diligence is a standard that should be applied when state-sponsored attacks happen, and it could function as a deterrent. If, allegedly, a state-sponsored CNO is said to have origin in a particular state, then this state should help in the investigation process, reversing the burden of proof, that passes from the victim to the potential attacker. Such a method, however, encounters the difficulties of dealing with states that invoke the principle of non-interference, such as China and Russia, to name two. Nonetheless, due diligence could be applied even without the consent of the accused state, rendering the best way to have state and non-state actors complying to international law. Also lowering the standard of the burden of proof and the standard of proof is a method that should be considered by the International Court of Justice in dealing with cyber disputes, taking also from the criminal jurisprudence by following “means, motives, and opportunity”. Of course there must always be a degree of evidence considered sufficient enough to make accusations. The issues that these two methods of attributing state responsibility met could be that CNOs do not travel in a straight line, but could pass from third states and sometimes multiple states, in case of DDoS attacks and that there is always the issue of false-flag attacks.

Given these problems in determining state responsibility, also in cases of state-sponsored CNOs, and the lack of an international set of norms specifically addressed to cyber disputes, the only option available to governments is resorting to self help. Self help, as it has been outlined, is subject to the variable of power, therefore for now is up to the state determining which is the appropriate response to CNOs that come from a state or from private actors aided, even passively, by another state. Estonia, weaker than Russia (and its nationalist hackers) as far cyber capabilities are concerned, resorted to deterrence, with the help of NATO and becoming a champion of cyber security among states, sponsoring also the Tallinn manual on the international law applicable to cyber warfare. The United States and China, being two superpowers in both cyber and physical world, it could be supposed that both states have the capabilities to attack the other and also determine when it is attacked by the other. In this case, the two states resort primarily to self limitation, in order to avoid that the dispute in the cyber domain could escalate in the physical realm, using self help, like the indictment of the PLA members as a method of signalling. However, the most interesting example is what is called Sony Hack where the United States helped a non-state actor victim of a CNO from another state. In this case, the threat to both American citizens working for the company, and as well to national security, targeting an infrastructure - not near to the requirements of being considered "critical" - spurred a state reaction culminating with both economic sanctions as well as retaliation in cyberspace, exploiting the same kind of non-regulation of the domain. The US could have done that only because their power is superior to the North Korean one, thus eliminating the fear of a retaliation. Nonetheless, another, weaker state that has seen an hostile CNO against a private actor on its soil could have publicly denounced the act. Therefore, it is safe to say that a cyber dispute is a matter of state even when one of the two actors involved is a state, and due to the absence of an internationally shared set of norms, due diligence should be applied but, at least for now, the disputes must be resolved autonomously by states.

Why Do States Wage War In Cyberspace?

Having seen the various scenarios of cyber disputes between states, it could be asked the reason why states wage "war" in cyberspace. The answer to this question is multi-faceted. The first part of the motivation is, straightforwardly, the lack of an internationally shared normative framework. In this sense, cyberspace is deeply anarchic and every state, as stated before, resorts to self-help to regulate its own behaviour and to retaliate against the others.

Given this context of anarchy, states are subject to a classic mechanisms of international relations, that is to say the security dilemma. Cyber weapons and cyber operations are developed and carried out in secrecy. This secrecy is sometimes broken when a CNO, in whichever stage it is in, is found, namely when a foreign state presence it is found within the systems of a particular state. These findings do not tell much about the capability of the attacking state, and do not neither confirm that the attacker is a state. Usually, the "state presence" behind an attack is confirmed when the attack has ended in two ways: the first one is a display of a great degree of sophistication, the second one is thanks to backtracing plus intelligence plus forensics analysis. Sometimes the second one does not exclude the first one. A foreign presence in all those systems containing sensible information for a state, for example governmental agencies, military offices, research and development facilities, or managing critical infrastructures, are nonetheless perceived as a threat, and a government cannot and should not rule out a state or a state-sponsored CNO. Such a posture, however, is really dangerous because it doesn't give to the state the possibility of interpreting any action. Indeed, every reconnaissance phase could be a prelude for a hostile act, even an act of war (Buchanan, 2016). This because intelligence actions aimed at collecting information are crucial to develop state strategy both inside as well as outside cyberspace. For this reason, these incursions need to be long lasting, in order to collect the highest amount of information as possible and in order to do so they need to be stealthy. It is straightforward that knowing that today many states have this capability this generates fear, and what a state could do, in return, is developing or amplifying its cyber capabilities. To test and to signal these capabilities a state could perform reconnaissances or incursions in other state systems, creating a circle of fear and eventually a "cyber arms race" (McAllister, 2015)

In the cyber realm, at least for now, the attacking side has the upper hand for two main strategic features of CNOs: the apparent low entry costs for developing cyber weapons and the vulnerability of the IT infrastructures. These two point will be developed in the next chapter, but they are worth mentioning because they are the two main reasons why cyber disputes increased in numbers in the past years and they are also the reason of two specific trends in cyber attacks. The first one is that, generally speaking, the level of these disputes today remains low, due to states imposing self restraints on their CNOs, because it is very hard to control the consequences and the full potential of a cyber attack, given the interconnection of critical infrastructures in the world, and also between civilian and military ones. Low-level cyber disputes are of medium-high probability with low to medium impact. For example, during conflicts or tense political times low-level cyber attacks are used to try to influence public opinion and government decisions, for example the already mentioned Estonia case of 2007. The peculiarity of these kind of low-level cyber attacks is that they have little consequences on the overall conflict but they have a high impact in terms of visibility and are immediate (Cavelty, 2015). On the other hand we have those cyber attacks that are of low probability but with higher impact. These attacks are very specific due to the fact that those states that are developing better cyber capabilities rely on an intense collection of information that stems from multiple cyber incursion performed in a very long period (Buchanan, 2016). For this reason, the second trend is an ongoing shift from horizontal attacks towards vertical attacks, that is to say from attacks that were intend to affect as many machines as possible (viruses, spam mails) to targeted, custom attacks, aimed at attacking a specific machine with specific vulnerabilities, such as Stuxnet (Cavelty, 2015). Indeed, it could be argued that collection of sensitive data and future attacks are tightly linked, or at least the planning for future attacks and this feature worsens the security dilemma in cyberspace.

Another feature of cyberspace that helps worsening the security dilemma is that defending in cyberspace is the most difficult task. General defence in cyberspace is a passive exercise. There is no software that could detect an attack and respond accordingly. Cyber defence relies on firewalls, anti-virus software, and also operational awareness for the employer of a given facility or infrastructure. As a comparison it could be said that cyber defence is like fortifications. Nonetheless reconnaissance incursions and even attacks could be discovered before they become full-blown attacks and the vulnerabilities could be patched, but it does remain a non-powerful defence. Critical infrastructures could also rely on a “resilient posture”, that is to say an holistic defence that reduces the possibility to attack - by patching the systems and developing a high level of cyber hygiene - and fastens the time of recovery after an attack - by collaborating in close contact with a Computer Emergency Response Team (CERT) or even with government agencies, fostering the public private partnership (Bologna et al, 2013). However, the best kind of defence seems to be an active and aggressive one, such as the one performed by the United States (Buchanan, 2017). This defence relies on an intense collection of information - taken from enemy systems - about potential enemies that could be planning attacks against the US. Furthermore, this type of aggressive stance is both a force posture and also a method to attribute CNOs, such in the Sony hack case, which acts as a deterrent as well.

The security dilemma could also be worsened by the time in which the intrusion has been found. It is straightforward but an intrusion found during a geopolitical crisis could lead to a more severe consequences compared to the same intrusion found during peacetime, increasing also the risk of an escalation in the physical realm. Furthermore, the increasing number of states developing or bettering their cyber capabilities contributes in adding complexity to state calculations in attributing CNOs, causing uncertainty and worsening the security dilemma.

To sum up, the characteristics of the anarchic cyber realm contribute in fostering insecurity because the attackers have incentives to launch CNOs in advance, the best option to defenders is to launch intrusions against many other states as possible to collect the highest number of information, and therefore all intrusions are to be treated as a potential threat. It is clear now to state that the risk of escalation and of cyber, or non-cyber, conflict that no one should want is really high.

Escalation In Cyberspace

After the end of the Cold War, when escalation was a central topic for both the academic and the policymaking environment, the importance of this subject met a sort of a downfall. However, new conditions began to affect the same principles and the same issues around which the "classical" concepts of escalation revolved around. The concept of escalation had its highest point during the Cold War, then the discussion faded during the past twenty years. The rise of new actors, new security problems, and new methods to wage war brought the discussion back to the table (Morgan, 2008). Technology is one of the factors influencing the renewed need to study escalation. Computer network operation do indeed affect the modern security environment being disruptive enough to be an opportunity as well as a concern for leaders and policy makers. The opportunity consists in a new, less destructive way for waging limited disputes but for weaker enemies, weaker states this is also an opportunity to develop cyber capabilities in order to being able to retaliate and escalate if they are attacked, also to amplify their chance of deterrence. In a spectrum of potential weapons to use during an escalation, at one end we would surely have nuclear weapons and at the other end would find cyber weapons, that are definitely less apocalyptic but also less costly, and therefore more attractive, given also the high level of dependence from IT system of the majority of countries around the world (Nye, 2011). This dependence is perceived as a weakness but also as an opportunity to attack and also to escalate in a conflict.

The rise of China, and of other new, immature forces - as far as cyber capabilities are concerned -, that are obtaining cyber power to compete against, and attack the United States and other countries, refreshes the problem of being able to manage a new kind of escalation

New powers, such as Iran, Syria, India, Pakistan, North Korea, are resolving to cyber power in order to shift regional, and international, disputes into cyberspace (Geers, 2014). These new opponents are different from the nuclear Soviet Union of the Cold War, and for this reason they could pose a new challenge regarding how to predict their intentions. Not only, these actors have more options for escalation compared to the USSR, including, but not limited to, tools for conducting cyber operations. Indeed, the increase in the number of cyber-savvy state actors, as we have seen in the previous chapter, worsens the security dilemma and could push state to perform CNOs against other states both as a mean to attack as well as an act of defence. This potential increase in CNOs, in turn, leads to an increase in the risk of escalation.

Escalation in cyberspace could happen for various reasons: as a retaliation to a previous attack, in order to demonstrate a superiority in power and to make the enemy desist; but also as a way of signaling determination and to test the potential reaction of an enemy (Springer, 2015), in this sense, potential escalation in cyberspace is no different from escalation in the physical realm.

Therefore, in cyberspace an escalation process is indeed possible but with different modalities compared to classic escalation, that is to say that it would not follow the so-called escalation ladder proposed by Kahn (Morgan et al, 2008; Libicki, 2012; Springer, 2015). Indeed, it could happen with different speeds and with different intensities, sometimes very apparently and some other times defying the eye of an observer that would fail to recognise it as escalation. Nonetheless, escalation within cyberspace follows a tit-for-tat fashion (Springer, 2015) following what Thomas Schelling already stated in Arms and Influence, that in the escalation process enemies tend to “keep things in the same currency”, that is to say to speak in the same language, responding in the same manner, with the same means to the first attack (Schelling, 1966). The same reasoning could be found also in Axelrod (1984) where he found that the best repeated iteration to promote cooperation in the Prisoner’s dilemma were indeed tit-for-tat strategies. Axelrod concluded that the tit-for-tat move could have been extended to other situations of bargain that were characterized by incomplete information, uncertainty and lack of trust (Axelrod, 1984). It is straightforward that here we cannot really talk about cooperation, because contrary to the simulation performed by Axelrod, during cyber conflicts the first move is all but cooperative. The matter at stake here is more about the willingness not to escalate the conflict by communicating with exchanges on the same level that must be clear, as Axelrod put it (*Ibidem*). However, problems with tit-for-tat exchanges in cyberspace are mainly two. The first one, strictly linked to Axelrod's experiment is that given the non-cooperative nature of the exchange, the two competing states do not attack at the same time, but to an action corresponds a reaction; furthermore, there is the risk, without cooperation, that a tit-for-tat strategy could lead the two state in a grip of coercive bargaining (Potegal and Knutson, 2013). The second problem is that, as it is outlined in the next chapter, the principle of proportionality cannot be applied to cyber weapons, leaving the dangers of a potential escalation in the hands of the perception of the actors involved. Indeed, perception and misinformation are essential elements in the escalation process, and cyberspace is no exception. For example, the level of importance of a target could be high for the attacker, but the opponent could find the attack bearable, and the contrary is also true, and even more dangerous (Libicki, 2012). Or, on the other hand, due to the fact that cyber weapons do not always produce physical consequences, the real effects of an attack could not be fully clear for the attacker, namely a difference in what an attacker thinks he did and what it actually did (Morgan 2008). Indeed, this could also act as an instrument to limit escalation, because also Clausewitz stated that if the consequences of an attack are unsure the escalation is moderated" (Cimbala, 2012). For example, some states could have seen how Stuxnet was planned to stay inside the Natanz nuclear enrichment facility but it spread, allowing many analyst to analyse it. States that do not want inadvertent escalation to the civilian infrastructure or even with third states, or that do not want the characteristics of their secret weapons to be revealed, may be refrained to attack light-heartedly.

Another factor that influences escalation, strictly linked to perception of cyber weapons and targets, is of course the cyber capability of one state, what was called "cyber power" at the beginning of this chapter. The capabilities of a state influence both the sophistication of the cyber weapons it could produce, as well as the types of target it could attack (Springer, 2015). Therefore, it could be stated that the combination of capabilities and perception is what makes cyber exchanges not following Kahn's ladder. Perception is crucial in this case because the decision of employing weapons could be based on "subjective interpretations of the actions of others" as Hammond (1992) put it. Indeed, psychological factors could influence the actions of policy makers, which could act irrationally or without complete information (Jervis, 1976). A state with a low degree of cyber capabilities could respond to a previous high level CNO to one of its national infrastructure with a CNO that *perceives* as medium-high strike to what it *perceives* being an important target for its enemy. Therefore, in the mind of the retaliator, the incremental ladder is still followed, but not in reality, indeed, the outcome could be that the first attacker does not even perceive the escalation.

Classic Escalation and Cyber Escalation

During conventional escalation, the seriousness of an attack is measured in two ways: based on geography and intensity. In the first case the area under attack widens compared to the previous one and/or the number of buildings, critical infrastructures attacked increases, and it is called horizontal escalation (Morgan et al 2008; Sweijs et al, 2016). In the second case, the retaliation is characterised by an increased intensity of violence, whether by using a larger number of the same weapon previously used or using a new, more powerful weapon, compared to the previous one (Morgan et al, 2008; Sweijs et al, 2016). In light of this, what could constitute escalation in the cyber arena? It could be argued that an escalation in the virtual domain should be characterised by the same criteria of conventional escalation. Horizontal escalation should consider the scope of the malware, how it spread through the network, how many infrastructures it was found in or hit, and, also important, the criticality of the infrastructures targeted (Libicki, 2012; Cavaiola et al, 2015). The increase in intensity, namely the vertical escalation factor is straightforward, it is related to the violence of the payload of the malware, whether intended or actual (*Ibidem*). Summing these two criteria should give us a good measurement of the overall attack, hence contributing in creating an escalation ladder. It is very important that the two criteria, geography and intensity, are analysed together and for a simple reason. An attack with minor disruptive power, such as a DDoS, for example, that is easily fixed by restarting the server, could have widespread effects, and block an entire infrastructure for several hours, or days, if the attacks are continuous, like the ones during the cyber dispute between Russia and Estonia in 2007. On the other hand, a very disruptive cyber attack, or the initial attack of a potentially disruptive cyber campaign, could be discovered and blocked before they reach their full potential.

Another criterion that must be taken into account is the nature of the attack, that is to say which kind of cyber weapon the perpetrator uses, and whether it is a CNA or a CNE. However, if we take into consideration intensity, many recent CNE could fall under the label of Advanced Persistent Threats, APTs. They are massive, continuous attack, aimed at stealing sensible information, such as R&D documents, personal data, and so on.[5] Do the

[5] APTs will be analysed more thoroughly in the next chapter.

two different methods of attack require different responses? That is why everything should be taken into consideration when calibrating CNAs to fit an escalation ladder.

Some authors argue that the lack of an internationally-shared idea of escalation ladder in cyberspace will determine miscalculation from both sides, decreasing the power of deterrence (Libicki, 2012). Indeed, brief escalations, or retaliations decreasing the power of the previous attack, or long, prolonged exchanges of cyber attacks seems to be the trend in cyberspace right now . One could argue that states are containing themselves, and this could happen because of the "one-shot" characteristics of cyber weapons and the fear of losing control over the attack, or that the deterrence power is indeed weakened as Libicki stated, and that those attack are enabled by different ideas of escalation ladder, that is to say the perception of both the attacker and the target.

Two important factors help in shaping an escalation process: the mechanisms and the reasons for it. Literature recognises three types of mechanisms: deliberate, inadvertent and accidental that could be applied in cyberspace (Lin, 2012;). When an escalation is considered to be *deliberate* it means that there is the open willingness to escalate the power of the conflict, the reason could be obviously to gain something from the increase in power, and is the typical kind of escalation (Morgan et al, 2008; Sweijs et al, 2016). The impact, apart from being a temporal disruption of systems of the loss of sensitive data, it is also psychological meaning a provocation or a call for attention (Lin, 2012). The best way to respond to, and limit, a deliberate escalation is - of course deterrence - and the next paragraph will look in-depth to the mechanisms of cyber deterrence. An escalation is considered *inadvertent* when the willingness to escalate is not present, but the actions are perceived as escalatory, mostly because they pass a certain threshold set by the enemy of which the attacker is unaware of (Morgan et al, 2008; Sweijs et al, 2016). Therefore, an inadvertent escalation fails to anticipate the potential effects on the escalation stemming from an inaccuracy in understanding the perception of the enemy. This type of escalation is very risky, because it cannot be perceived as such by both the parties in conflict. In this case even though there is the risk of being caught during the attacker persists in trying to find a way inside the enemy's systems. An external presence in the system could be perceived by the victim as an ongoing attack, while it is not by the attacker, spurting escalating retaliation. How to limit the risk of inadvertent escalation? Literature revolves around awareness, about the capabilities and the possible reactions of the enemy, about the existence of inadvertent escalation (in order to take it into consideration when making strategic plans), and also making the enemy aware of the possibility of inadvertent escalation (Morgan et al, 2008).

Therefore, as far as conventional escalation is concerned, there are several self-imposed limits in order to contain the risks of inadvertent escalation, but the question is whether same argument apply to cyber operations? The answer is a tricky one, because I think there are limits to the limits. A similar scenario, but with different awareness from the attacker side, could result in the so called *accidental escalation,* which is characterised by the fact of being unintended (Sweijs et al, 2016). For example, a reconnaissance in the enemy's systems, therefore only an intrusion used as a probe, could be perceived as the beginning of an attack, as we have seen in the previous paragraph. This could lead to a response in kind that is perceived by the first state as a first attack, leading to another response that could become an escalation. Inadvertent and accidental escalation are two very similar concepts, the difference among the two relies in the perception of the actors involved. In the first one the attacker does persist in attacking grounding this action on the fact that in its mind they are not escalatory while for the victim they are perceived as such. In the second one both of the parties are unaware of the escalation risks, therefore the attacker doesn't perceive its actions as a threat - since it is a simple reconnaissance, albeit against state sovereignty - and the victim perceives it as an attack, and responds accordingly. Literature is right to suggest that, together with the inadvertent, accidental escalation has to be managed, because the primary sources are the front lines, not the C&C (Morgan et al, 2008). The management of the cyber operators is difficult to implement, because the minimum tampering with the target systems' data could lead to an accidental escalation, that becomes deliberate for the other party.

Why Do Conflicts Escalate?

The first reason is *instrumentality.* A party escalates for the belief that it will be better off after doing so, whether by ending the conflict or being closer to a victory. Literature suggests that a willing act of escalation with an instrumental motive stems from, for example, a crisis scenario when diplomatic channels and negotiation do not work (Morgan et al, 2008). This could be applied to the rescue of hostages but could also be applied to the Stuxnet case, that was tailored for a specific situation in order to give a signal to the Iranian government as the P5+1 talks were at a stalemate. In this case, a cyber attack is a fait accompli aimed at obtaining a psychological or political effect. The second reason to escalate is signalling to the enemy the potential high costs of a possible escalation and it was anticipated fifty years ago by Thomas Schelling in Arms and Influence (Schelling, 1966). This escalation seeks to deter further attacks, by instilling fear in the enemy that responding to the attacks will bring further and costlier escalation (whether horizontal or vertical or a combination of the two), negating any possibility of gaining something (Morgan et al, 2008). It is important to remember that even an explicit threat could be considered as escalatory if it passes a given threshold. As far as cyberspace is concerned, the threat of considering a cyber attack as an act of war deserving a kinetic response with conventional means, such as the one issued by the Department of Defense of the United States (DOD, 2015), can be considered a suggestive escalation. Another example is a demonstration attack, such as the aforementioned Stuxnet, who was a clear example of the cyber capabilities of the US against Iran. The last example of suggestive escalation is brinkmanship, where one of the two parties willingly creates an "existential shared risk of disaster" where escalating would deprive both sides of any advantage (Morgan et al, 2008; Sweijs et al, 2016). Other motives for escalating are less rational, and consists in escalating for the sake of doing it, for setting an example. In cyberspace it could translate in deliberate attacks in order to prove and to communicate a certain level of cyber power, such as the growing in number and sophistication Chinese CNOs against the US, from Titan Rain in 2004, to APT1 in 2013.

Sometimes escalation is in the very nature of some conflicts, because specific types of weapons are used or a certain target has been attacked. In this case it is difficult for a conflict to de-escalate. The escalation could proceed as a self-feeding process: as losses rises, victory becomes more and more critical driving both parties towards the use of bolder and dangerous actions. However, if the escalation is based on rhetoric or it is symbolic, it is easier to abate the tones of the conflict. In kinetic conflicts there are normative barriers to escalations that grew in importance during history, such as the chemical and the nuclear one, but also technical barriers, such as countermeasures that nullify the risk of escalation. What about barriers in the cyber arena? Could it be that a warfare method such as the cyber one that is precise and does not foresee civilian harm nor even casualties creates new, and lower, barriers? Or does it push it to the extreme, meaning that there are almost no barriers to escalation in cyber conflicts? Furthermore, if we add a lack of international regulation it could be stated that there are basically no barriers to initiate a cyber dispute and to let it escalate. Literature suggests that bloodshed is not necessary for escalation, also national security and reputation are factors that, when in danger, could lead to escalation, whether intentional or not (Morgan et al, 2008).

Another reason why disputes in cyberspace might happen and tend escalate is that the high level of dependence from the IT networked infrastructure is perceived as an opportunity to attack, given the presence of vulnerabilities. Literature describes the windows of opportunity as *"circumstances in which an actor believes it has a significant but temporary ability to attack, escalate, or take some other action and that if it does not do sol, the opportunity will diminish or disappear" (Ibidem*). As far as cyberspace is concerned, the "temporary ability" is the timeframe between the creation of the system and the patching of the vulnerability, or vulnerabilities, that the attacker wants to exploit. With kinetic weapons the expectation of windows of vulnerability closing will make an approach "Use it or lose it" arise. This is also true for cyber weapons that exploit known vulnerabilities, but it is important to say that many newer cyber attacks, such as Stuxnet, use so-called *0day exploits*, meaning malwares that exploit vulnerabilities that are not public yet.[6] If the vulnerabilities are known by a limited number of people, maybe limited to the army of one particular country, the window of opportunity will last indefinitely. In this case, what does arise is not a "use it or lose it" mentality but a "use it then lose it" one, a sort of "first strike vulnerability". If the target discovers the attack, it is able to patch the vulnerability and make the attack useless from that specific point in time onwards. That is another reason contributing to the control of the escalation process that collides with the reliance of countries like the US to the interconnected critical infrastructures.

Vulnerabilities, and their known presence, obviously shape policy makers and leaders' decisions and states strategic postures. But given the absence of a legal international framework, the responses and the doctrines rely only on the national dimension. This could give potential enemies incentives to escalate without the worry of an international condemnation or response. However, this doesn't mean that the response is absent, for example after the Sony attack, the US posed bilateral sanctions on North Korea and attacked through cyber means. Nevertheless, the modern situation in the cyber arena is one characterised by instability, between cyber savvy actors, and cyber rogue states that are exploiting new vulnerabilities to escalate and prove that they are powerful, without worrying about an indictment from the International Court of Justice.

Despite the incentives to attack and to escalate, cyber disputes seems to be constrained somehow. Constraints to conventional escalation are basically costs. These could be

[6] 0day vulnerabilities and exploits will be analysed more thoroughly in the next chapter.

monetary, human and reputational, for example. Cyber disputes seem to be untouched by these constraints unlike, for example, chemical weapons. The constraint that touches both kinetic and cyber escalation is that escalatory act can create a bigger escalation process at the end of which the situation is worse for one or both parties compared to the current status quo. The tendency to de-escalate and a self-restraining behaviour seems to be the norm during cyber escalations, that could be more intense, sometimes they are a little bit more than a retaliation in kind. The hypotheses here are multiple, including not knowing the real capabilities of the enemy, not knowing the perception of the enemy to a particular CNO, meaning the fear of the cyber conflict to switch to a kinetic one. Another explanation could be a remainder of the studies of Snyder and Diesing (1977) where, in a situation of distrust and lack of communication coercive exchanges are the only way two states could rely in making joint decisions. The outcomes of a developing conflict are influenced by the relations between the two parties, and are deeply affected by their bargaining power that is influenced by their interests and their willingness to go to war (Snyder and Diesing, 1997). Another plausible explanation is that, given the unregulated environment and the looming shadow of a possible kinetic conflict, it always seems to depend by the variable of power, understood generally. In this case, it seems that a position of escalation dominance is imperative in order to end the retaliations and to function as a deterrent. In reality, it could be both true and not true at the same time. Escalation dominance is "a condition in which a combatant has the ability to escalate a conflict in ways that will be disadvantageous or costly to the adversary while the adversary cannot do the same in return, either because it has no escalation option or because the available options would not improve the adversary's situation." (Morgan et al, 2008). This factor is key for reading cyber disputes that, again, depend on the variable of power. For example, Russia is far superior to Estonia as far as cyber capabilities are concerned, therefore Estonia could have not retaliated against Russian attack in 2007. With the involvement of NATO, however, Estonia found itself in a position of escalation dominance because if the conflict would have escalated in the physical realm, then Russia would have found itself in a situation of inferiority. However, if we take for example Stuxnet as a starting point of a cyber dispute between US and Iran, we already assume that Iran, at least at that time, did not have cyber capabilities strong enough to respond with a malware with the same power. Allegedly instead, Iran developed Shamoon, an espionage malware found in Saudi Arabian computers to monitor and erase critical files on about 30,000 computers at Saudi Aramco, the world's largest oil company, disabling them (AFP, 2012). Saudi Arabia is one of the major US partners in the Middle East, and that could be considered a test-bed for retaliation. Possible in response, in April 2012 the Iranian government stated that the Kharg Island oil terminal, which exports 80% of the country's daily 2.2m barrels, was hit by a cyber attack, along with terminals on the islands of Gheshm and Kish (Dehgahn, 2012). Again, the US claimed that, in 2013, Iran hacked US Navy computers (Barnes & Gorman, 2013). In May 2014 iSight Partners, that is a security firm based in Dallas, issued a report stating that Iran has been performing a CNE campaign dubbed "Newscaster" for the last three years that targeted military contractors, members of Congress, diplomats, lobbyists and journalists (Perlroth, 2014). Therefore, escalation dominance influences cyber disputes but not as much as conceived theoretically. The exchange of attacks could be continuous but with the increasing power of each retaliatory attack does not follow Kahn's ladder, instead the attack are used to show will and power.

Straightforwardly, the study of deterrence becomes pivotal in order to anticipate potential disasters. This translates into measures to make a dispute not, or less, attractive, and less likely to escalate, given the fact that the menace of a dispute or a limited conflict in cyberspace cannot be completely eliminated, just like kinetic ones.

Deterring Cyber Conflicts

Militarisation of cyberspace seems a process that is ongoing and unstoppable. On the other hand, we have incentives from states to control the risk of a potentially uncontrolled and uncontrollable escalation. These incentives translate in the high number of meetings and conference, both at national as well as at international level. As far as deterring cyber conflicts, three main measures could be outlined: the first two are normative, that is to say is the creation of the so-called TCBMs - Transparency and Confidence-Building Measures - and the attempt to regulate cyber conflicts within the international law framework; the third one is applying deterrence to cyberspace.

TCBMs, a mechanism used mostly during the Cold War, are part of the legal and institutional framework supporting military threat reductions and confidence-building among nations. They have been recognised by the United Nations as mechanisms that offer transparency, assurances and mutual understanding amongst states and they are intended to reduce misunderstandings and tensions. However, when applied to cyberspace TCBMs can address other activities in the virtual realm outside of those performed for by the military or for those performed for national security reasons (*Ibidem*). While TCBMs promote transparency and assurance between states, they do not have the legal force of treaties and states entering into them are bound only by a code of honour to abide by the terms of the instrument. By their nature TCBMs are considered a "top-down" approach to addressing issues. (*Ibidem*). They are not intended to supplant disarmament accords but rather to be a stepping stone to legally enforceable instruments. The OSCE, the Organisation for Security and Cooperation in Europe, published in December 2013 the "Initial set of OSCE CBMs to reduce the risks of conflict stemming from the use of ICTs". (OSCE, 2013). Its focus was the sharing of information on diverse levels and an overall increase in transparency between participants. However, this does not have any legal power so it is mainly an "expression of goodwill" among countries. Nevertheless, such a step forward means that there is common ground for building up further measures. In fact, in 2014 UNIDIR took up the challenge straight away and inside their cyber stability seminar 2014 "Preventing Cyber Conflict" many ways of cooperation and prevention were discussed, among which there are TCBMs and codes of conduct. The important part surely is to extend this type of dialogue to the private sector, without which a complete cybersecurity will be impossible.

Indeed, also international dialogues have been taking place, both between countries and countries' related stakeholders, in order to build confidence and trust among the parties, and most of all, raise awareness and understanding among all the players involved in the cybersecurity process.

Applying the international law to cyberspace has always been a difficult task, and it still is. As it has been outlined at the beginning of the chapter, is very difficult to apply state responsibility standards to cyberspace, and as it will be outlined in the next chapter, regulating cyber weapons still it is an issue untouched by international law. As far as *jus ad bellum* is concerned, the "Tallin Manual on the International Law Applicable to Cyber Warfare" (Schmitt, 2013) and its follow up "Tallin Manual on the International Law Applicable to Cyber Warfare 2.0" (Schmitt, 2017) although they constitute a solid base in-line with existing international law, remain non-enforceable and purely academic and non-binding documents. Therefore, they do not play any role as far as deterrence is concerned, because they do not exercise any kind of power on states.

Deterrence

When one wants to discuss about deterrence and how to apply to a new domain such as cyberspace there is the – natural – tendency to go back to prior studies on the Cold War. All the elements of the classic discourse on deterrence however fail to grasp the new realm and the mechanisms within it. For example, there could be no "mutually assured destruction" in cyberspace because cyber weapons cannot cause level of destruction of nuclear ones. Another issue that is not affected by deterrence is strategic stability, that is to say the interest in normalcy and avoiding arms race (Davis, 2014). CNOs during peacetime among States are the actual normalcy in this time and era, and straightforwardly contribute in raising tensions and insecurity, that is to say the failure of deterrence. For example, another key concept that is applicable both in nuclear strategy as well as in cyber deterrence is the relationship between counter-value and counter-force strategies (Libicki, 2009: Davis, 2014; Gartzke and Lindsay, 2017). If during the Cold War attacking civilian infrastructures (counter-value) was considered abhorrent as well as pointless, and attacking military installations (counter-force) was deemed essential to win a war, in cyberspace this paradigm becomes more blurry. For example, Glaser argues that counter-value attacks are easily deterred because, through context and investigation the culprit is more likely to be found, circumventing the attribution problem (Glaser, 2011). Davis adds that counter-force attacks are difficult to deter because they are part of military operations of conventional warfare (Davis, 2104). Therefore, a State must deter and persuade a potential enemy to avoid attacking through other means. "To prevent from action by fear of consequences" is what Thomas Schelling defines as deterrence in his Arms and Influence (Schelling, 1960 p.71). It is a very broad definition that could be subject to various interpretations, however it could be considered the core of the concept of deterrence, namely reducing the potential of a threat to zero by making an enemy believe that its costs to pursue this threat far exceed its benefits. When facing a threat, deterrence is, essentially a bigger threat. However, as Schelling states in Strategy of Conflict, the threat must be credible in order to sort the desired effect (Schelling, 1960). The second element of deterrence consists in the ability to pursue that threat. Indeed, capability, credibility and communication are the necessary elements for a successful deterrence (NATO, 2016). The "capabilities" mentioned by above are the ones used to retaliate after an attack - therefore constituting a punishment - and also the capabilities employed to defend critical infrastructures or even to discourage an attack, meaning denial. After having obtained capabilities strong enough to perform the aforementioned tasks, it is important that the country that is to suffer from the attack displays its willingness to retaliate, or is indeed capable to defend or deny. Communication is used to establish a certain degree of credibility, for example the aforementioned indiction of the five PLA Chinese officers accused of hacking could be seen under the lens of a declaration or force posture, providing for a threshold that is not to be trespassed. However, it must be bear in mind that, as it has been outlined in the previous paragraph, misinterpretation affects the dynamics of escalation, but it also affects the mechanisms of deterrence (Jervis, 1976). Communication becomes key also during crisis.

The concept of deterrence in cyberspace it is not new, compared to the history of the military exploitation of the domain. The first scholars tried to tackle this problem in the 1990s when it was still called "information warfare". Haynes and Wheatley, in 1996, stated that, as a sovereign state, every attack against the United States - whether criminal or warmongering in nature - is deterred by the same set of policies that deter physical attacks and they also affirm that "information attacks are attacks and therefore subject to international law. Violations of sovereignty and acts of war are no less real because they use the information domain than if they involved more traditional violations." (Haynes & Wheatley, 1996). Of course, this kind of statement ignored all technical problems relative to the characteristics of the domain and also the issue of attribution. However, the research of Harknett, still in 1996, was more profound and very forward-looking. For example he states that deterrence in cyberspace could fail because of incomplete and incorrect information about the attacker, and that the attacker could miscalculate the threats of the deterrer, even if they are sufficient to shift the cost-benefit axis to its disadvantage (Harknett, 1996). He concludes that, influenced by the dynamic of nuclear deterrence, that in cyberspace matter only offence and defence, that is to say an overwhelming attacking power and high degree of defensive capability that make deterrence only a by-product (*Ibidem*).

The first milestone in the field of deterrence in cyberspace it is constituted by "Cyberdeterrence and Cyberwar" of Martin Libicki. Libicki however, focuses only on deterrence by punishment, insisting on the fact that in cyberspace the only retaliation should be in kind (Libicki, 2009). This is generally true because it avoids entirely the problem of escalating in the physical realm, but it could also be considered as wishful thinking because states could retaliate in the physical world, not necessarily with weapons, but with other tools, such as monetary sanctions, or indictments, as it has been outlined in the first part of the chapter. However, deterrence by punishment is not the most useful method of deterrence in cyberspace. Deterrence means to impose costs that exceeds the benefits for the attacker, these costs however should not be seen only as retaliatory costs in kind, but also costs in terms of efforts and even economic costs, therefore another more feasible option is deterrence by denial. Indeed, deterrence by denial means to deny the objective to the attacker, by rendering it very difficult to reach or making it incur in high costs (Snyder, 1959). The problem of information and communication in deterrence remains as crucial as it is in the physical world, but provides a much more challenging task in cyberspace where an attacker is not sure about being threatened or even punished (Libicki, 2009). However, deterrence by denial is better than deterrence by punishment also because it could reduce the problem of misinterpretation increased by the virtual environment. Putting in place well-defended systems and a resilience posture, combined with a convincing deterrence policy, would increase a) the communication of capabilities and intentions and b) the capability and the sophistication needed by the attacker, restricting the roster of the potential culprit to a smaller number of states. In this sense, the risk of wondering if a disruptive CNA stems from a planned operation or a reconnaissance gone out of control (Rid and Buchanan, 2015) is reduced because only to a handful of states. In this scenario, if the deterrer has put in place forms of active defence, then both the risks of misattribution and misperception are reduced.

The problem of attribution remains always one of the main issues in cyberspace, furthermore in those cases where deterrence by punishment is put in place. Forensics measures are available to every state that possess cyber capabilities, of course varying in degree accordingly to their level. Nonetheless, it generally takes a lot of time to be sure about attributing CNOs, even though initial assumptions can be made about attacks and cyber security firms could help in the analysis (Rid & Buchanan, 2016).

Moreover, Libicki is right in focusing on the attribution problem, however it ignored Estonia 2007 - and also Georgia 2008 - where the attribution problem seemed to be an issue that could be overcame resorting to an analysis of the geopolitical situation within witch the attacks happened. In those cases, the pointing to state responsibilities helped in deterring further attacks. Furthermore, as history repeats itself, the same situation rose again during the CNA against the Ukrainian electric grid (Lee et al, 2016). This case is useful for this analysis for two separate reasons. On the one hand, it serves as an example for CNOs used for signalling and deterring measures, on the other hand underlines the major role of geopolitics in the attribution process. In periods of political crises between two states, attribution could be done more clearly like in all those three cases where Russia was involved. Furthermore, in two of those three examples - Georgia and Ukraine - military operations where ongoing, narrowing the scope of attribution and, at the same time, underlining the Russian use of cyberspace as an added weapon during a conflict. And, to conclude, in the Ukraine case the attack was sophisticated and well-prepared enough to point the finger against Russia and to eliminate the possibility of a false flag attack. However, geopolitical analysis alone does not help in overcoming the attribution problem. For example a CNO against JP Morgan in 2012 was attributed to Russia, but then it was discovered that the culprit was a private group (Nye, 2017). Therefore, a clear geopolitical analysis helps deterrence through attributing CNOs only when its combined with good forensics measures, otherwise is communication without credibility. Indeed, credibility is a key factor in attributing a CNO and in fostering deterrence, because it targets two main actors. The first one is the potential attacker, that knows that the capabilities of the deterrer are higher than its own, because it has been found culpable. The second one is the general public: the accused attacker could deny any involvement in the CNO, and the deterrer must convince this audience of its righteousness, as the United States have done in the Sony Hack, and in the PLA officers case, trumping North Koreans and Chinese denial of involvement (Nye, 2017). To conclude, attribution and its capability to influence deterrence is subject to the technological capabilities of states, and as they increase also their deterrence capability increases.

Conclusion

The concept of deterrence goes back to Thucydides and Sun Tzu, who said that the best way to win a war is by not fighting it. Classic deterrence could be difficult to study, because it is like measuring the success of counter-terrorism intelligence, that is to say the absence of an harmful event. Cyber deterrence, on the other hand, does not provide the total absence of cyber conflicts, but a reduction of the potential CNOs against the systems of those states that are capable of acquiring a defence dominant posture with a convincing deterrence policy.

Until an internationally shared code of behaviour regarding what is acceptable and what is not as far as CNOs is established, states will keep on relying on self-help measures that stems from perceptions and misperceptions linked to single countries. Indeed, self-help is for now, the main deterrence strategy but it also is a double-edged sword. To effectively deter CNOs states, like the United States - which are the only country that is known for sure to employ this kind of strategy - infiltrate other countries systems, both enemy that allied, in order to know potential planned and ongoing operations. This kind of deterrence that we could call, at this point, active deterrence relies on an attribution system that is based on potential illegal bases due to the fact that the necessary condition is to violate state sovereignty. The main issue is that national cyberspace is not considered as such by international law, instead it should be. Nonetheless, as stated at the beginning of the chapter, the International Court of Justice has already accepted evidence coming from illegal operations. The problem is that we could face a conundrum: many states would see the active deterrence strategy as attractive, and therefore we would be very likely to see an increase in CNEs. These operations, if performed by low cyber powers, could be victim of misperception, and maybe simple incursions could go rogue if not managed perfectly, worsening the security dilemma and raising the risk of escalation. This risk could be halted by the fact that states able to effectively apply active deterrence would communicate to the enemy the fact that they know about the CNE, maybe with public attribution, sanctions, and/or by signalling through cyber measures, such in the aftermath of the Sony case.

Nonetheless, it is necessary - for those states who are capable of doing so - to adopt a defence dominant posture, that is to say securing the systems with resilient measures. This holistic strategy involves all the national actors responsible of cyber defence, from the government and military apparatus, that should work with those who manage national critical infrastructures (which are most of the time private), that should impose awareness and cyber hygiene measures on their employees. This is crucial because the majority of attacks are enabled by human errors within the target infrastructure (Nye, 2017). Eliminating, or heavily reducing, this entrance mode could, in turn, eliminate or heavily reduce the risk of CNEs. This could be also expanded to other potential national targets - such as in the Sony case - namely private companies whose data could be object of nation states campaigns.

The actors who could put in place such a strategy could effectively state that they could deter the majority of cyber attacks and in case that they are victims of a CNO the roster of culprits is reduced to a handful of potential attackers, due to the fact that such measures are not among the capabilities of many states.

Furthermore, another way to have even more better deterrence is by expanding the deterrence posture to the allies, and this could happen in two ways, one direct or indirect. The direct way is within the deterrence by denial framework, and it is for example what has been done by NATO as a measure against Russia manoeuvres against Estonia (NATO, 2016). Indeed, Estonia shifted the balance of power in its favour by pushing towards the addition of cyber defence measures to the NATO charter and with the development of the CCDOE, the NATO Cooperative Cyber Defence Centre of Excellence. This posture of collective deterrence functioned as deterrent against potential CNOs coming from Russia and other possible threats, due to the fear of an allied response from NATO.

In the physical realm, alliances among states should come in the form of arming and instructing the allies, to deter communal enemies (Mitchell, 2015). However, in cyberspace this is not possible because cyber weapons are not like kinetic weapons, the strongest ones are secret because they are very likely to exploit 0days vulnerabilities. Sharing 0days would nullify completely the benefits they bring, and even sharing the know-how would become a double-edged sword, because potential enemies could penetrate weakest allies' systems in order to possess not only shared strategic plans but also this know how. That is why there is an indirect way that has been employed by the United States, that is to say to preemptively penetrate allied systems to gather intelligence (Edgar, 2017) and to see whether they could be attacked by those communal enemies. This method however goes against the trust among allies and, ones again, state sovereignty. Despite this issue, this method could provide the same deterrence that stems from breaking into enemies' systems, assuring certain attribution. Nonetheless, it could be degrading alliances, therefore the best way could be to provide the allies at least with a better know-how to detect intrusions. Another key factor of deterrence is therefore attribution, and as it has been underlined throughout the chapter, it is not as an insurmountable issue as was previously considered as it is strictly linked to state responsibility, as demonstrated by the Estonia case, the Sony hack and the indictment of the PLA officers. Therefore, to conclude, cyber deterrence has all the means to work, considering both the unique characteristics and mechanisms and cyberspace without forgetting classic geopolitics.

Chapter II: States' New Tools of Warfare, Cyber Weapons

Introduction

As cyberspace constitutes one of the dimensions where states compete for power, a nation that wants to be defined as powerful in cyberspace must acquire the main tool to exert its power in cyberspace, namely cyber weapons. When talking about weapons, the thoughts are always directed to kinetic weaponry. Cyber weapons are the natural result of years of technological advancements in the exploitation of the ICT environment, both military as well as civilian. Given the high degree of interconnection brought by cyberspace, the entire backbone of society is interconnected and, at the same time, dependent on the ICT infrastructure. This infrastructure for its very nature as a man-made environment presents flaws and vulnerabilities that could be exploited to open breaches into systems and networks. These vulnerabilities are essential to cyber weapons to work and indeed shape their peculiar characteristics. Kinetic weapons are means of harm and destruction that, even though are still evolving now, possess the characteristic of being always harmful in time, since their invention. For example, stone-age rudimental axes could still kill a man today, as they did thousands of years ago. Straightforwardly, digital vulnerabilities could be fixed making those cyber weapons based on the exploitation of that given vulnerability useless. Such a characteristic creates weapons that are temporary.

This leads to two possible scenarios. The first one is a "use-it-or-lose-it" approach, where actors could be pushed by the hurry and the fear of the window of opportunity closing and therefore deploying the weapon as soon as possible (Krepinevic, 2012). This kind of approach of course would accelerate the possibility of retaliation and even escalation due to the apparent lack of inhibition in using cyber weapons, that could maybe accelerated by the temporary lack of norms and regulation at international level. The second scenario is a refrain in deploying cyber weapons because, contrary to the previous paradigm, the actual use will automatically close the window of opportunity, so it must be used when really needed (Libicki, 2007).

I deem this scenario as the most realistic one, for one simple reason, that is to say the presence of zero day vulnerabilities. This concept, that would be addressed and expanded later in the chapter, could be summarised with "vulnerabilities unknown to the creator of the piece of software or hardware and therefore temporarily unfixable". If an attacker knows about this vulnerability, a method to exploit it can be created and could be used without any rush. Of course, here the supposition is that the said vulnerability would never be fixed until its exploitation, but there are still chances that it could be fixed, for example during a penetration test, rendering the 0day exploit useless. There are other characteristics of cyber weapons that would be discussed later, for example the unpredictability of the diffusion of the cyber weapons in the systems and networks that brings with it the risk of attacking unintended targets that, of course calls for caution in their use. The apparent refrain in using highly disruptive or even destructive cyber weapons could stem from Cold War reminiscences, that is to say "mutually assured destruction" and "second strike capability". Albeit the first one seems a little farfetched – It could be more the likes of "mutually assured disruption" – the second possibility is very plausible. Without international regulations, and given the "virtual" nature of cyber weapons, statesmen could only infer what the cyber capabilities of other states are. Furthermore, as far as cyber weapons are concerned, there is no direct link between methods for attacking and targets, meaning that it is not necessary a highly sophisticated weapon to cripple an important target. Indeed, a crippling attack towards the electric grid could be responded with a temporary disruption of the banking system with two different kinds of cyber weapon. The one-shot nature and those fears of the unknown are possibly what makes countries refrain in intensely deploying disruptive and destructive cyber weapons.

This refrain and use of cyber weapons mostly for espionage purposes could lead us to think that they are not a subject relevant to study, but such a reasoning would be regarded as wrong. If espionage is a permitted activity – within limits – meaning it is not recognized as an act of war, the increase in the sophistication and intensity of attacks aimed at stealing sensitive data has increased, reaching noteworthy levels. Indeed, the use of cyber weapons for espionage purposes has increased (Leccisotti et al, 2016), including an increase in the employment of a particular kind of cyber weapons called APTs that stands for "Advanced Persistent Threats" against the financial, government and industrial military complexes (Ilascu, 2014; Infosec, 2015; Chandarimani and Monrad, 2016;). What characterises these weapons is the sophistication of the components of the weapons, including the aforementioned 0day exploits, and the persistence with which they attack systems and network in order to achieve their objective. Indeed, these two characteristics – the high degree of sophistication of the weapon and the persistence of the attack – Indicate the likeliness of a state-actor as an origin point. As an example, in 2013 APT1 one of the first APTs discovered, was linked to the PLA Unit 61398 belonging to the Chinese army (Mandiant, 2013). The problem with the "it's only espionage" approach is that it could not be applied in cyberspace. The problems lies in the fact that these weapons could be updated any time the attacker wants; therefore, once inside the systems they could change their nature from a tool meant to extract data to a mean to cause disruption, to say the least. This means that the espionage attack is considered as such only when it constitutes a *fait accompli*. But what about when an attack is discovered before the payload is delivered? Is it right to give a sort of benefit of doubt to the attacker or do we need to tackle this kind of issue in other ways to deter and prevent a further spread of cyber attacks?

In order to deal with these issues it is necessary to know everything possible about cyber weapons available to state actors today and for this reason this chapter is structured as follows: the first part defines the concept of "cyber weapon"; the second part explains the functioning of cyber weapons, namely what permits them to act in both disrupting and non-disrupting ways, what are the possible effects they can and could cause and which objectives they can and could reach; the third part aims at providing an as much as possible complete and holistic taxonomy of the cyber weapons that are available for states today[7]; the last part will focus on the lack of regulation around the use of cyber weapons and why are they so difficult to be framed within a normative corpus, furthermore, this last segment will also provide suggestions for a future regulation of cyber weapons.

Due to the fact that a general description of cyber weapons is not present in the political science literature,[8] this chapter aims at providing an added value to this field of study. In this sense, this chapter would represent an added value to the academic world, that should be a link between the "technical world" which is evolving at a really fast pace and both the analysts one, by providing them with a knowledge that should permit better analysis of cyber conflicts, and policy making one, which is lagging behind in terms of rules and regulations.

[7] Since this dissertation is built around a political science framework, this part does not go into technical details, i.e. programming and code structures of cyber weapons. In those cases where technicalities could not be avoided, the author tried to explain them in the most clear way possible, without diminishing the value of the part of the chapter.

[8] There are, of course, lots of papers and reports on the subject, and various manual on malwares that belong to the cyber crime literature. All these efforts have been used to draft this chapter.

Defining Cyber Weapons

As of today, there still is a lack of commonly accepted definition as of what constitutes a cyber weapon. Single elements such as “virus”, “worm”, “trojan horse” possess different and numerous definitions that try to describe them. The technological evolution in the malware activity poses continuous challenges to the classification of cyber weapons. Indeed, it is becoming more and more difficult to try fit a new malware in rigid definition and to define them accordingly. This research will try to encapsulate those terms in a precise framework, albeit without going in technical (namely code structure) details. This choice is motivated by the fact that what is really needed is to have a common ground to build shared definitions in order to have a basis that allows researchers and practitioners to compare different malwares. Furthermore, defining specifically what are the characteristics of the different cyber weapons will aid policy makers, academics, researchers and strategists to better tackle the issue.

For example, the Tallinn Manual describes cyber weapons as “cyber means of warfare that are by design, use, or intended use capable of causing either (i) injury to, or death of, persons; or (ii) damage to, or destruction of, objects, that is, causing the consequences required for qualification of a cyber operation as an attack” (Schmitt, 2013). It is important to underline two things about this definition. First of all, the term “cyber means of warfare” encapsulates both what could be considered a weapon and all the systems associated with it. In this case, a weapon is “an aspect of the system used to cause damage or destruction to objects or injury or death of persons” and therefore it encompasses also all the devices, instruments, equipment and even software designed, or intended, for cyber attacks (Schmitt, 2013b). Furthermore, the Tallinn Manual specifies that, in order to be considered as a weapon, that specific part of the system must be under the control of an actor. Therefore, the “internet” or other communication infrastructures belonging to the so-called cyberspace are not to be considered as means of warfare (*ibidem*). For the sake of clarity of the chapter, it is important to stress that the Tallin Manual focuses more on attacks rather than on the classification of cyber weapons. Given the definition above, a cyber weapon is the *conditio sine qua non* to have a cyber attack, that straightforwardly here means the generic employment of cyber weapon. However, classifying the usage of cyber weapons as armed attack is useful to have a better understanding and to manage the different kinds of cyber weapons later in the chapter.

Despite this “technical” classification, the problem of defining cyber weapons remains for example in Section 2, rule 13 of the first edition of the Tallinn Manual. This rule regards “Self-defence against armed attack” that starts from the premise that if a State is victim of a of a Computer Network Operation (CNO) whose consequences are on par with the ones of a kinetic attack, then the State may exercise its right for self-defence (Schmitt, 2013c). The question here is straightforward: when does a CNO reach the level of an armed attack?

First of all, an *armed* attack implies, by definition, the utilisation of weapons, described above, and that the scale and the effect of the attack are similar to the ones “that would result from an action otherwise qualifying as a kinetic attack” (*ibidem*). In order to assess the magnitude of a CNO, then, the Tallinn Manual uses two variables that are the *scale* and the *effect.* What it fails to mention are the parameters that must be used in order to assess whether a CNO spurs the consequences of an armed attack. Nevertheless, the Tallinn Manual relies on the clarity of the act. When the CNO damages or destroys property or when it is directly responsible for the death or injuries of people, then it could be classified as a cyber weapon; when they are used for intelligence gathering or for interrupting briefly non-essential cyber service then they could not be considered as an armed attack (*ibidem*).

It is also clear that not every CNO must be at the armed attack-level to be classified as a weapon: there is a wide range of attacks that employ cyber weapons that amount at the level of “use of force”. A CNO belonging within this level of magnitude is a cyber operation equal to a kinetic operation conducted by intelligence forces of a country or by a non-state actor operating under State sponsorship (Schmitt, 2013d). Then, the Tallinn Manual clearly states that the authors, the so-called International Group of Experts, do not know “what actions short of an armed attack constitute a use of force” (Schmitt, 2013e).

The important thing to underline is that the Tallinn Manual aims at providing legal provisions, tracing a line that goes from kinetic acts, whether they could be described as ‘acts of force’ or as ‘armed attacks’, to similar acts in cyberspace. This, of course, is not a simple task. Indeed, when trying to use scope and scale to determine what constitutes an act of force in cyberspace and how is it quantitatively and qualitatively different from an armed attack, the Tallinn Manual renounces in giving an exact threshold. However, it is fair to say that even in the physical world these thresholds are subject to change and could be perceived differently from state to state. Nevertheless, the International Group of Experts mentions an approach in order to try and establish thresholds depending on the level of harm inflicted and qualitative elements. This approach is based on several factors that are probably the ones that the international community takes into consideration when it is the case to consider a particular operation as an use of force or not (*ibidem*). These factors are: severity, immediacy, directness, invasiveness, measurability of effects, military character, state involvement, and presumptive legality (*ibidem*).

Severity depends on scope, duration and intensity, and concerns the consequences of a computer network operation. *Immediacy* regards the timing that the consequences take in order to manifest themselves; immediate consequences are perceived as harbouring stronger damage. *Directedness* is the degree of strength of the link between cause and consequences. *Invasiveness* refers to how much the cyber operation was able to propagate inside a given system; invasiveness depends also on the degree of security of a system or network: the more secure, the more penetration is perceived as a cause of concern. *Measurability of effects* is another very straightforward criteria, as it refers to the measurability of consequences of cyber operations; the more a computer network operation is assessable the better it could be considered as an act of force or not. *Military character* refers to the nature of a specific computer network operation, implying the presence of military or armed forces behind the attack. *State involvement*, fundamental to retaliate or not, is a very polarised factor; in this case one can have “state presence” that includes computer network operations made by the national army or state-sponsored attacks, or the absence of it, for example during cyber operations coming from private actors that have absolutely no involvement with a government. *Presumptive legality* means that acts that are not forbidden by international law are permitted, and so they are less likely to fall under the categorisation of acts of force.

It is clear to see how the Tallinn Manual examines *ex post* effects in order to categorise offensive cyber operations but it is not the only method available in literature. Another model that aims at identifying better what could constitute a cyber weapon is the PrEP model (Herr and Rosenzweig, 2015). Herr and Rosenzweig try to expand the definition given by the Tallinn Manual in order to help policymakers in drafting better cyber control treaties, identifying three main components: a propagation method, an exploit and a payload (hence PrEP). The point here is to take also into consideration the components of the malware, meaning an *ex ante* analysis, to determine whether a cyber weapon constitutes also a military characteristic. They conclude that it is the payload part that helps in distinguishing a military cyber weapon to a non military one; but still, a weapon that does not display a disruptive payload could be a dual-use tool. This means that a particular cyber tool could be used both for military and non-military purposes (*Ibidem*). Indeed, the best definition for a cyber weapon that could be drawn from this research is that a cyber weapon has a propagation method and uses an exploit to deliver a payload. For this research, it does not really matter whether the payload is disruptive or not, it is sufficient to have a measurable payload. The payload is defined as the core of a cyber attack, and it is the mean to reach a desired goal, achieved through a modification of the target system. This modification could range from allowing a simple intrusion to leading to a disruption.

Another definition of "cyber weapon" is the one of Thomas Rid and Peter McBurney, who state that a cyber weapon is "a computer code that is used, or designed to be used, with the aim of threatening or causing physical, functional, or mental harm to structures, systems, or living beings." (Rid and McBurney, 2012). Their work is useful for the field of study because it points out two interesting concepts. The first one is a rough classification about the intensity and the sophistication of cyber weapons and the second one is drawing a line about what constitutes a cyber weapon and what does not. The first is that we do not have a unique set of cyber weapons but they span in a wide range, whose poles are referred to as "generic but low-potential tools" and "specific but high-potential weaponry". With the increase in sophistication also the budgetary spending increases in turn (*Ibidem*) which is very agreeable to, and roughly followed by the description of cyber weapons in this chapter. The second one is the most interesting and debatable because it is a recurring argument also present in the Tallinn Manual, that is to say: "the most common and probably the costliest form of cyber-attack aims to spy. But even a highly sophisticated piece of malware that is developed and used for the sole purpose of covertly exfiltration of data from a network or machine is *not a weapon."* (*Ibidem*). This definition concentrates only on the objective of a cyber weapon but it is fair to underline that, as it is outlined later in the chapter, once inside a system cyber weapons could change their nature remotely, thus they have the capability to transform themselves from espionage tools to disruptive tools.

Clay Wilson, instead, does not offer a proper definition of cyber weapon but states that it has to have four general characteristics, namely "combining multiple malicious programs for espionage, data theft, or sabotage; a stealth capability that enables undetected operation within the targeted system over an extended time period; an attacker with apparent intimate knowledge of details for the workings of the targeted system; a special type of computer code to bypass protective cybersecurity technology." (Wilson, 2015). As we can see, these characteristics define more broadly what could be defined as cyber weapons, because it takes into consideration that the need to move away from the link with kinetic weapons, and include also espionage, encompassing all tools that permit carrying out all CNOs.

Another definition is given by Maathius, Pieters and van den Berg that, albeit clarifying that "weapons are not tools for espionage" define cyber weapons as: "a computer program created and/or used to alter or damage (an ICT component of) a system in order to achieve (military) objectives against adversaries inside and/or outside cyberspace" (Maathuis, Pieters, van den Berg, 2016). What is most interesting about this definition is the word "alter". The main point is that files and data are altered in their "confidentiality, integrity and availability" (Committee on National Security Systems, 2015) every time a malware is used.

The lack of a unique definition of "cyber weapon" and the low number of attempts to give one reflect the heterogeneity of the studies on the specific elements of the cyber realm and also the need to study this field and to propose contributions. This research will use an "expanded" definition of cyber weapon with elements derived from the classic definition of the Tallinn Manual but with the decision to include also CNEs that is to say weapons which aim at the exfiltration of data. This is motivated by the definitions of Herr and Rosenzweig and Maathius, Pieters and van den Berg. Herr and Rosenzweig stress the fact that it doesn't matter if the payload aims at disrupting the systems or at siphoning sensitive documents, what matters is the nature of the payload, if it is military it is to be considered a state or state-sponsored cyber weapon. This point is of great usefulness for the thesis and it is fostered also by Clay Wilson, which encourages to move away from concepts bound too much to the classical, kinetic, reality, and include also espionage. The last motivation is grounded in the definition given by Maathius et al, because even though they say that weapons are not used for espionage, altered data could mean both that a document is copied or that a file vital for the functioning of the system is modified. Therefore, since a cvber weapon could modify its very nature CNAs and CNEs should be treated alike.

Functioning of Cyber Weapons

Cyber weapons are constituted by real (meaning physical) and virtual (meaning non-physical) elements, and for this reason they could be considered both as a tool as well as a capability. The functioning of a cyber weapon follows an indirect path (Rid, 2013) and, compared to kinetic weapon the likes of a missile, it causes an indirect damage instead of a direct, and immediate, one. For this reason, it is more likely to see exploitative and disruptive cyber weapons than destructive ones. In the latter case, the destruction will be indirect. Cyber weapons, in the form of both Computer Network Attacks (CNAs) and Computer Network

Exploitations (CNEs) require three essential features to work: target vulnerabilities, the access to them and a payload (Lin, 2010).

Vulnerabilities

How is it possible for cyber weapons to cause this kind of indirect damage? Every offensive attack in cyberspace exploits so-called *vulnerabilities*, that is to say flaws in the construction of a specific system, whether it is software or hardware. These flaws could be exploited to compromise the proper functioning of a machine or to mine the confidentiality, integrity, and availability of information and data. Vulnerabilities exist because they are basically bugs, errors in the code, flaws in the construction of a system, and they are present because they are made by humans, and humans make mistakes. Therefore, we could distinguish among vulnerabilities depending on which part of cyberspace they belong to.

Martin Libicki, one of the first and most prominent experts in cybersecurity, characterises cyberspace as an environment possessing three different layers, respectively defined as the physical, the syntactic and the semantic ones (Libicki, 2009). Straightforwardly, the physical layer is constituted by all tangible components in the form of machinery (servers, wires, computers, routers, HMI interfaces and so on and so forth) that both allow the existence of and the access to what is commonly known as cyberspace. The syntactic layer is composed of all the means that empower a user to communicate with the machines and the ones that enable the processing of information by the machine, that is to say operating systems, software and applications. Information constitutes the third layer, the semantic one. Each of this layer could be separated from the others for analytical purposes, but they work in close connection and cyberspace cannot exist without one of them. The relationship among the three layers is a pyramidal one, in the order in which they were just described: the semantic rests on the syntactic which, in turn, rests on the physical. It is important to underline this relationship because it affects also the consequences of possible offensive operations. An attack aimed at the physical level[9] would have direct consequences on the syntactic and the semantic one.

[9] Attacks aimed at the physical level include both kinetic attacks against the machinery, namely the physical destruction caused by external agents, as well as CNAs. Examples of this are to be found, for example, in the so-called Aurora Generator Test and also in Stuxnet. These attacks proved that attacks through cyberspace could provide destruction to a physical object.

For example, the physical destruction of a server would also result in the disappearance of all the software and information stored in it, impairing all the processes that depend on such machinery. On the other hand, an attack on the information stored on the same server does not directly influence the other two layers. Attacking each of these layers requires different sets of skills and different methodologies that exploit different kinds of vulnerabilities.

Attacks at the physical level could include kinetic attacks against the machinery, cables, namely the physical destruction caused by external agents, but these are not to be considered as "cyber attacks" for the sake of having a proper and, most importantly, useful definition of cyber weapons. It is important to underline that "attacks at the physical level" include both attacks *directed to* the physical layer, and *starting from* this level. Attacks directed to the physical layer could happen by exploiting different vulnerabilities. One instance could be constituted by an attack happening through a hacked piece of hardware that has been unknowingly replaced, which is due to malfunction after a certain amount of time or through a direct order.

One, albeit unconfirmed, example of this attack is the explosion of a gas pipe in Siberia in 1982 allegedly due to a sabotage in the SCADA system of the infrastructure. This sabotage was organised by the United States, in collaboration with a Canadian supplier of Industrial Control Systems, that managed to install a faulty piece of hardware (which had a malicious software embedded in it) that brought the Programmable Logic Controllers to fail and cause an explosion (Reed, 2005). These attacks proved that attacks through cyberspace could provide destruction to a physical object. It is fair to say that the majority of CNAs start from the syntactic layer, where the majority of vulnerabilities reside, since it is the level that concerns software and operating systems.

It must be also underlined that a particular offensive attack does not have to be limited to a specific layer, but it can travel through these different layers. For example, it could begin from the physical one in order to attack the semantic one, in the case of an attack that injects a malware from a USB port capable of deleting or modifying information. Another scenario could involve an attack starting from the semantic layer, modifying particular software that manage machinery in order to cause physical damage. For this reason, to this day an air gapped system – namely a a computer network system that is isolated, thus not connected, to other networks (including the internet) that could be breached - remains the most secure environment, but not immune to attacks. Nonetheless, individuals could try to gain access to an air gapped system by installing some software or even some hardware locally, and this could happen with the help of a willing aide or even an unwilling one. The last vulnerability does not entirely concern the three layers of cyberspace, but adds a new one: the human vulnerability. Indeed, humans could be taken advantage of by means of social engineering.

Nonetheless, some vulnerabilities become useless when they are discovered and then patched, that is to say fixed. The main problem is that sometimes the vulnerabilities are fixed by the software vendor through a patch, sometimes they are not, and it is up to the user of the given system to fix it. This leads to an uneven environment where some systems are exposed to threats due to a vulnerability that in some other system is fixed. This is a scenario that almost the entire world faced in 2017, during the WannaCry epidemic. The diffusion of the infamous ransomware was aided by the fact that some systems were not updated and therefore vulnerable to an exploit that was patched and fixed months before (Sherr, 2017) To avoid wasting efforts and time in reconnaissance phases in the target system to see whether some vulnerabilities are fixed or not, the best way to lower the margin of error or exposure for a cyber weapon is to find, and then exploit, a so-called "Zero Day Vulnerability". This definition, sometimes written as 0day, refers to a particular kind of vulnerabilities that are unknown to the owner and the issuer of the product that has been exploited. Hence, exploiting this kind of vulnerabilities has a success rate of 100% because there is no patch available and the user has zero days to secure itself (Hypponen, 2010). It goes without saying that this kind of vulnerabilities gives the attacker a great advantage. Therefore, zero days vulnerabilities are to be considered very precious elements in building a sophisticated cyber weapon. For their high success rate and for the fact that in order to find those a person or a group must review the entire code of a software, a specific market is born and these vulnerabilities are sold at very high prices.

Regardless the presence of this black market, states tend to balance the diffusion of vulnerabilities and zero days by having special relationships with software manufacturers, for example between the United States of America and Microsoft (Lin, 2010), where the latter gives up the source code to government specialists in order for them to analyse it and discover possible vulnerabilities that went unnoticed during the coding processes. This gives a double advantage to the State: first, it permits to fix the vulnerability thus avoiding being attack through that vector and, second, it allows the State to attack possible targets of interest by exploiting said vulnerability. Furthermore, in order to increase its advantage, a State could convince product vendors to secretly insert so called "backdoors" into softwares that are publicly available (Lin, 2010). States are aware of this possibility, and it should be noted that the main operative systems, as far as commercial purposes are concerned, are produced by two American companies namely Microsoft and Apple.

In light of this, for example, China banned Windows 8 from its computers after Microsoft stopped updating Windows XP that is to be found on between 60%-95% of Chinese personal computers (Kai, 2014). This choice stems from the need for China to try to be more independent in the software market, resulting in huge investments on a national operative system (Kai, 2014) and also in the insertion of backdoors in the first generation of Chinese smartphones, especially Huawei ones (Mathews, 2016). This last example is not connected directly to the Chinese government but it is a reflection of an understanding of the potential advantage that coding commercial softwares brings.

Access to vulnerabilities

Vulnerabilities could be distinguished as local or remote. The difference between the two is the way they could be accessed to: the first ones need a person that has physical access to the system. For example, some hardware vulnerabilities can be exploited only in this way through an entry port, whether USB, firewire, HDMI (Hypponen, 2013). Most likely, Stuxnet was able to propagate inside the system starting from an infected USB key, regardless of the willingness or unwillingness of the user to do so. Not only, the alleged Russian pipeline attack belongs to this kind of access, since the tampered hardware part required a direct installation, therefore extending to all compromising that happens at supply-chain level. Instead, remote vulnerabilities could be accessed even at distance from the target system, usually using the internet as the access path (Lin, 2010), and therefore pose a more serious threat.

The method used to take advantage of vulnerabilities, usually in the form of code, is usually called an exploit. The exploitation of vulnerabilities leads to the propagation of the cyber weapon and/or to the delivery of the payload, depending on the goal of the cyber weapon. Indeed, the attacker has the possibility to exploit vulnerabilities in three different ways: to gain access, to perform an escalation of privileges and to execute code (Herr and Rosenzweig, 2015).

It could be difficult to distinguish between an exploit and the payload and indeed some literature combines the two elements (Bright, 2011). These two features of a cyber weapon have different objectives and occur at different times in the sequence of the attack. We are always talking about code, but the exploit is written in accordance to a specific vulnerability in the target system and it focuses on the structure and function of the software, unlike the payload that is written to achieve a given effect, and therefore it focuses on the output of the code. (Herr and Rosenzweig, 2015). To put in simpler terms, if one thinks about the target system in the form of a house, an attacker could gain access once he has found that the lock of the front door is vulnerable and therefore the lock-picking is the exploit. Exploits could be used and re-used in time, unless they are patched, but they tend to remain in time because they belong to vulnerabilities that are sometimes widespread, whereas a payload is more specific as far as their purpose is concerned, they have to be coded accordingly and they are difficult to re-use once the code they take advantage of has been fixed(Herr and Rosenzweig, 2015).

Of course, also zero day vulnerabilities are exploited, with a so-called zero day exploit. For example, Stuxnet used four of these, and given the aforementioned high prices for the vulnerabilities it was one of the elements that pointed towards a nation-State as the culprit behind the attack. By the presence of a high number of 0days exploit we could also infer the resolution of the attacker of wanting the centrifuge to fail.

Exploits, including zero day ones, can be divided in three different categories: access, escalation of privileges and code execution (Herr and Rosenzweig, 2015). The term "access" indicates an exploitation of a vulnerability aimed at, straightforwardly, gaining access to the target system. This is used in the reconnaissance phase while searching for given files or other vulnerabilities in order to propagate or to extend the reach of the cyber weapon even further. Escalation of privileges means that the attacker exploit vulnerabilities in the privileges management system, trying to gain the highest level of privileges possible, like the one of administrator. This permits the attacker not only to access files, but also to modify them and to run code, feature blocked for more "basic" users. Code execution is indeed the capability, through the exploitation of other vulnerabilities, of being able to run commands in the system. The most dangerous kind of vulnerabilities to exploit are indeed the ones involving code execution and, furthermore, exploited remotely: these are called Remote Code Execution vulnerabilities (Hyponnen, 2013).

Payload

With the term “payload” we refer to the main component of a cyber weapon, that is to say that part whose purpose is to be executed inside a system in order to achieve a particular and pre-defined objective. Indeed, the payload has been described as the *“raison d’etre*” of a cyber weapon (Kirwan, 2011; Herr and Rosenzweig, 2015). Comparing the cyber realm to the kinetic one, if the payload of, for example, a ballistic missile is measured in the number of warheads, it is useful to underline that also cyber weapons are able to deliver more than one payload but these could be very different from one other. Once a malware is able to infect a given system, it could be programmed to perform different actions, from scanning for new vulnerabilities to copy and transmit data and also alter and destroy files (Lin, 2010). Furthermore, the payload can change once it has already been delivered, for example it can be updated or instructed to delete itself. All these action are performed whenever the actor in charge of the malware wants (Lin, 2010), therefore also the timing in the delivery of the payload is very different from the kinetic ones. The mutant capability of a cyber weapon is to be considered a very strong asset since it allows the attacker to deploy the best weapon possible, always depending on the capabilities of the actor. This happens because an attack is very likely to start with a simple reconnaissance of the target system that could mutate through a remote update, to be transformed in the best tool to perform the task intended for a given goal.

Since the malicious goal of the attacker may vary, the payload varies accordingly. For CNAs it is usually a disruptive, or destructive payload, whether for CNEs we should consider a payload aimed at stealing information and for this reason it should cause the least number of visible alterations in the target system in order not to be detected (Lin, 2010). There is a third kind of payload that crosses the borders between CNAs and CNEs, that is to say the capability of hijacking the systems (Furnell, 2010). This payload could be used, for example, to take control of different systems to create a botnet at first (CNEs realm), and in a second time this botnet could also be used to perform a DDOS attack (CNAs realm).

The payload does not change only according to the goal – damaging or stealing – of the attacker, but of course it changes also depends on the target system itself. Hence, this influences the sophistication of the cyber weapon to be deployed. In this sense, sophistication means the degree of complexity of the various elements put together in order to gain a certain capability that will help achieve a given goal. These elements are: the particular type of vulnerability, or vulnerabilities, that are to be exploited, the degree of capability to remain hidden in the system, the advancement of the toolkit employed and the type of target system. The less sophisticated weapons use, for example, vulnerabilities that are already known with already available toolkits, everything done without putting too much effort on avoid being detected. In this sense, the payload could range from a simple defacement[10] to the highly-sophisticated one of Stuxnet (Herr and Rosenzweig, 2015).

Briefly, the purpose of Stuxnet was to cause damage to the centrifuges of the uranium enrichment facility of Natanz, in Iran. In order to do so, the weapon was programmed to make its way into the system of the facility starting from an USB key, if the route that it took was wrong, then the malware deleted itself (Falliere, Murchu, Chien, 2012). Furthermore, Stuxnet is a prime example of a cyber weapon with multiple payloads due to the fact that, for example, one payload was discharged on the SCADA system of the centrifuge, in order to make them spin out of control to stress the metal and make them eventually break, and another one was delivered to the Human-Machine Interface (HMI) with the goal of making the monitor showing a correct functioning of the centrifuges, without displaying any anomaly (Falliere, Murchu, Chien, 2012). This last payload was pivotal in the execution of Stuxnet, since the disruptive payload took months to achieve the desired effect (Herr and Rosenzweig, 2015). Hence, we can see how the sophistication of Stuxnet was due to a complex target system.

After having found the presence of one or more vulnerabilities, and having acquired the means to exploit them, and having deployed its payload, a cyber weapon has some effect on the target system, namely the original purpose of the weapon. CNAs and CNEs have different effects that could be also distinguished between direct and indirect. The goal of the latter is the confidentiality of information that are saved on a system or are passing through a network, whereas the target of CNAs is to cause loss of *integrity*, *authenticity*, *availability* of hardware, software and data.

Loss of integrity refers to the compromise of these three elements such as a system does not operate as intended or expected, but not necessarily without the interruption of the system’s processes. Loss of authenticity refers to the compromise of the source of data, that can be deleted or changed maliciously. Loss of availability refers to the compromise of the functionality that the target network or system should provide normally. To put simply, the system stops functioning, emails remains unsent or do not reach their intended destination, or the speed of the system’s processes become extremely slow. It is straightforward that the consequences of such action could lead to unbearable damage if we refer to systems that manage industrial processes or any physical process (Lin, 2010).

[10] the act of changing the homepage of a website, usually with a message or a slogan.

Taking into consideration CNAs, we could consider direct effects every result of a cyber operation limited to the system that has been targeted, but there are also indirect effects involved. If an actor manages to take control of a critical infrastructure, e.g the system or network that manage an electric grid or the system or network that manage the traffic lights in a particular city, every tampering with these could results in accidents, injuries, and even possible deaths, that even if are not cause directly – such as as the consequence of shooting with a firearm, for example – are nonetheless caused by the employment of a cyber weapon. Indirect effects could also include possible damages caused by a cyber weapon that finds its way out of the borders of the system or network that has been targeted, thus attacking other systems or networks unintentionally.

For example, Stuxnet was found in other systems other than the one of Natanz because, somehow, it managed to "escape" the network of the uranium enrichment facility. In this case the interesting fact is double: on the one hand, regardless of the high degree of sophistication of the malware that was custom made for the Natanz facility it went out of bounds; on the other hand, thanks to the high degree of sophistication the malware did not cause damage to the other systems where it was found, due to the fact that it was made ad hoc for a certain target.

Objectives of Cyber Weapons

When talking about payloads, it was stated that they are deployed to achieve a particular *objective.* Since their nature is different, CNAs and CNEs are employed to reach different goals.

CNEs aim at: determining the structure of a network, so-called *mapping*, and monitor its traffic; exploiting information that is available on the target network such as stealing important information stored in the system or network for espionage purposes (Lin, 2010). The first one is coincidental to the first phase of a CNO, meaning exploring and observing the system that has been targeted to discover it's whole structure, on which software is based on, the nodes that compose it, the users that are joined to the network but also all the machinery connected to the network. Therefore, trough this operation is possible to know which are the most important nodes by examining the quantity and the quality of traffic generated and received by that node, so that one could easily identify possible targets.

The second use of CNEs is to exploit information. This operation is consequential to the first one because, for example, by monitoring the network and analysing the traffic of data, the attacker may filter all the information containing keywords that are deemed as important. By doing a follow-up analysis, important documents such as blueprints and secret plans for operations could be found and extracted, and also other confidential data such as password and other user credentials (Lin, 2010). In this way, other systems could be accessed without even breaking into them but simply impersonating other users. This operation could be done for two separate purposes, the first one being intelligence gathering, the second one being industrial and economic espionage. The line dividing the two could be blurred, even though at first glance the former is usually conducted by a State and the latter by private entities. States like China use industrial espionage to try and close the commercial gap with competitors, such as the United States, and in this case the Chinese national interests combine both military as well as commercial espionage.

The objective of CNAs instead is more oriented towards the detriment of data, disruption and even destruction of the system. Examples are: the alteration, including destruction, of data stored on a system connected to a network or on the network itself, becoming an active node in the network to transmit untrustworthy traffic, the degradation or denial of service on a network (Lin, 2010).

The first case includes an attacker that has access and modifies the military planning of the forces of another State: by doing so, the target will operate on the basis of false premises therefore corrupting the military operations and once the tampering has been found out, the reliability of the network and the database will be undermined; an attacker could destroy or alter important files and data by erasing them and entire databases, causing a malfunction of single or even large numbers of machines. This action does not only erase important documents, but if we take a power plant as an example, it could also destroy or alter files that are managing the infrastructure, thus disabling it. The infrastructure could be also controlled by modifying the files used to run the machinery, such as in the case of Stuxnet.

The second case, the impersonation of a node in the network, is done actively by issuing forged orders or documents, creating a chain of events based on false premises, and it is done by accessing to already existing accounts or creating a new, trusted, one. In environments such as governmental agencies or departments that are very numerous in terms of users and these are used to communicate with people they do not directly know, it could be very easy to run such an operation (Lin, 2010).

The last case is disruptive in its very nature because the objective here is the degradation or denial of service on a network. This is done by a Denial of Service attack (DoS) or a Distributed denial of Service attack (DDoS) that floods a network with numerous but small packages in order to overload the servers and slow down and even crash the machines. This kind of action is useful to cause a temporary interaction of the operations of the target systems but also to impair communications of the target, by rendering its systems useless. As a consequence, the target could then rely to less secure communications systems, thus opening the door to easier and more successful CNEs.

Having described in depth what are the basic elements that constitute cyber weapons, why they are used and what are their effects and objective, the next part will provide some examples of cyber weapons that could be employed.

Examples of Cyber Weapons

Now that it is clear how cyber weapons are defined, how they work and what they could accomplish, this part focuses on their classification. Cyber weapons could roughly be divided into two main families, that is to say "malwares" and "blended threats". This is done mainly for clarity purposes, due to the fact that blended threats are, as the name suggests, multiple malwares that work in sequence or in concert to create a bigger threat compared to single malwares. Providing a general taxonomy of computer viruses is not an easy task. The problem encountered with this part of the chapter derives from the absence of a holistic literature on cyber weapons. There are numerous documents available on malware but it could happen that they are not updated or too technical, while others use the term "virus" as a general term. This is due to the fact that computer viruses are the first form of propagating malware appeared in cyberspace, and the term began to encompass all forms of malwares (a problem also seen in the same word "malware" that will be addressed afterwards). Nonetheless, this part seeks to provide a clear and specific description of which tools are available in the arsenal of states today, trying to follow the main elements characterising cyber weapons described before, namely vulnerabilities-exploit-payload and objectives.

Malwares

According to the Committee on National Security Systems, a malware (crasis of "*mal*icious" and "soft*ware"*) is a "software or firmware intended to perform an unauthorised process that will have adverse impact on the confidentiality, integrity, or availability of an information system" (Committee on National Security Systems, 2015). Malwares are used both for CNAs and CNEs and indeed there are different types of malware that depend on the characteristics that are described above. Nonetheless, they are used to cause the highest level of damage to the target system or network, both in terms of espionage and in terms of loss of confidentiality, integrity or availability of information. They could be roughly divided in: viruses, worms, trojan horses and spywares.

Viruses

Viruses are malicious codes whose peculiarity is to replicate in order to infect the system (hence the name, since they operate very similarly to biological viruses). The principal characteristic of viruses that differentiates them from worms is that the latter alter running code, while the former alter stored code like boot files, executable files and simple files like PDF or JPEG (Subrahmanian et al. 2015).

The infection happens through the insertion of the malicious code into an executable programme or a file, so called "code injection". Viruses exploit flaws in the code of softwares, that could be design flaws, execution flaws (opening an infected executable file attached to an e-mail) or buffer overflow vulnerabilities, that is to say when the buffer memory receives more data than it can handle, this extra portion of data overwrites the software's internal variables. Straightforwardly, this extra portion of data would be scripted for malicious purposes (Filiol, 2006).

The access to the system by the virus usually happens through the simple opening of the infected files triggering also the portion of malware that so has permission to execute code and write to memory. This characteristic makes the propagation of the virus slower compared to a self-replicating malware like a worm. Furthermore, viruses are the most detectable and therefore vulnerable malwares, for at least two main reasons. The first and foremost is the existence of a huge anti-virus industry that is constantly up-to-date, and the second is that viruses run on computer memory and for this reason the system could act abnormally, like slowing down dramatically, crashing or even shutting down, indicating the presence of a virus (Weaver et al, 2003).

Viruses, in turn, could be classified along several criteria. We could distinguish between: file infector, boot sector, master boot record, multipartite and macro viruses. The names are pretty straightforward but it is necessary to describe them separately to offer a general taxonomy. File infector viruses are what is generally conceived as “computer virus” and infect executable program files (to differentiate them from data files, like word or excel documents). Usually the target of this particular kind of virus is code in the form of .com and .exe files, which will have some parts of their legitimate code overwritten by the malicious one. It spreads easily because it can be carried on hard drives, on flash drives, or stored on networks, and furthermore many of these viruses are memory resident, meaning that once the memory is infected, all executable that this memory runs become infected. Mainly the payload is disruptive, due to the nature of the virus to alter data, sometimes important to the functioning of the system, compromising it and rendering it useless. File infector viruses are the hardest to recover from, indeed the best course of action is to format the hard drive and re-install the operative system.

Boot sector viruses infect the boot record on hard disks and usb drives (and they used to infect also floppy disks, today more than obsolete). These viruses start their infection once the system is booted, and they could spread to every hard drive that has been connected to the infected system, since they are memory resident. They could be used to display a particular message on the system’s monitor, like Stoned and its sequent variants, render the machine unusable by freezing it.

Master boot record viruses are almost similar to boot sector ones, with the sole difference that they target the master boot record. For this reason, they are far more dangerous than boot sector, since the MBR is the first thing that a machine runs after the BIOS, meaning that all desktop activities and the whole operating systems are corruptible by the virus and that the machine could also not start. Multipartite viruses combine file infector and boot records viruses, infecting both program files as well as boot records. Macro viruses aim to infect data files, mostly all the software belonging to the Microsoft Office’s suite.

Payloads of the viruses are diverse. The first one, shared with worms, is to have no specific payload at all. As viruses rely on the system’s memory, this could be overused incapacitating the machine and even forcing a reboot. Other payloads of viruses have the aim of tampering with the integrity, authenticity, or availability of information like the overload of CPU or hard disk space; the acquisition of private information (such as personal data like credit card numbers, social security numbers and personal documents); malfunction of e-mail accounts, processing spam messages and therefore spreading itself at the same time; the defacement of a website; the projection of error messages that oblige the user to reboot the system; the damage of files that make the system reboot continuously. Furthermore, they are used to deliver tools for access the system remotely (RAT - Remote Access Tools), pay per install applications, website redirection or are used to display specific messages on the target system's monitor.

Worms

Worms are very similar to viruses, and indeed the distinction between the two is sometimes difficult to make. The main characteristic that distinguishes the two types of replicating malware is that worms are self-replicating, meaning that they do not require user assistance to propagate, even though the sophistication of last-generation viruses enables them to propagate without a user assistance. For example, so-called contagion worms could be described as viruses, because they do not need direct user activation, they are able to propagate also through "otherwise unconnected user action" (Wall, 2009), but for the sake of the definition here are listed as worms. Therefore, we could define worms as *self-replicating malware*.

Worms, as all the other malwares, exploit vulnerabilities and to do this they employ different methods such as: scanning, the use of target lists, and awaiting inadvertent user activation. Scanning is a feature of some worms, such as Code red, that allows them to probe – using a routine – a series of addresses in order to find a host that is vulnerable. The scanning of hosts could be sequential (that is to say it follows an order of addresses) or random (meaning that given a range of addresses it probes them without order). Scanning is the simplest form for finding vulnerable host and for this reason is used by a large number of worms, both active, the fully autonomous ones, as well as passive ones, the ones that require user activation such as the aforementioned contagion worms, and also the one that enter into action after a certain time (Weaver et al, 2003).

Target lists could be of two different types: pre-generated or externally-generated. A pre-generated target lists is embedded in the code of the worm, and it consists in known vulnerabilities that the worm is going to exploit, for this reason pre-generated target lists make the worm infection faster. Known vulnerabilities could be common vulnerabilities of which the attacker is aware that there is the possibility that they could not be patched or could derive from a previous reconnaissance phase and therefore the attacker *knows* about their presence, like in the case of the original Morris Worm (Weaver et al, 2003). Externally generated target lists are a list of vulnerabilities that is inserted in the worm code remotely in order to give the malware brand new targets to attack through the aid of a *metaserver*, that is a service that keeps information about different, dispersed, active servers.

Passive worms, such as contagion worms, do not scan or probe for vulnerabilities but instead await user intervention to infect the system and propagate with copies of itself, like the Gnuman worm (Weaver et al, 2003). These methods have all downsides, for example scanning generates a lot of anomalous traffic and it could be easily detected. Target lists, precisely *ad hoc* and externally generated, are application specific and therefore require a lot more effort to be put in place, but given their topological nature they tend to produce apparently normal traffic. Passive worms do not produce anomalous traffic but since they require user assistance their propagation is slow, like the viruses' one.

The propagation of the worm could occur in different ways, each affecting the speed and the noticeability of the malware. The mechanisms employed are essentially three: the first one is self-carrying, that is to say that the worm transmits itself in a completely autonomous way; in this case the transmission is part of the attack itself and for this reason it is used by *ad hoc* worms, namely topological ones, and self-activating worms, like the scanning ones (Weaver et al, 2003). Another mechanism is called second channel distribution, meaning that the worm exploits a secondary channel of communication to infect the target, like remote procedure calls (Weaver et al, 2003; Subrahmian et al, 2015). The last mechanism is the embedding of the worm to a message or to substitute a normal message in order to be sent along normal communication channels. In this way the spreading looks like a normal pattern of communication without anomalies. Embedded worms are stealthier than others but for this reasons they must be used only in stealth procedures, namely by passive worms (Weaver et al, 2003).

An essential step in the propagation of the worm is its activation, that could be direct, indirect or automatic. Direct activation is the slowest aid to propagation because it relies on a user executing the infected file. In order to speed up this process, sometimes attackers rely on social engineering methods, such as phishing or spear-phishing techniques that lure the victim into opening a corrupted e-mail attachment or to download a file from an apparently trusted source. Indirect activation happens when a user performs a normal task that indirectly triggers the activation of the worm, such as rebooting a machine, thus activating a login script, or inserting an unchecked usb flash drive that automatically executes the worm. Automatic activation includes scheduled activation, for examples by exploiting other scheduled processes on the machines, such as auto updates, and self activation – the fastest way that a worm has to propagate itself, by attaching itself to running code or execute commands automatically, exploiting the permissions associated to the service attacked (Weaver et al, 2003).

Given the high level of efficiency of worms, due to their resourcefulness in infecting a system or network, the payloads they could execute are numerous. The most basic one is, like worms, the absence of payload. As was stated before, since the functioning of some worms could draw a lot of power from CPU and hard drive, they could cause an overload and this effect could be enough to cause problem to vulnerable machines. Worms could lead to the opening of backdoors, subsequently used to execute code arbitrarily, like the Code Red II worm. They could infect machines that then are used as proxy "repeaters" for spam or for phishing activities. Not only, worms are key factors for Denial of Service and Distributed Denial of Service attacks: in this case worms have DoS or DDoS toolkits embedded in their code that are executed remotely without the owner of the machine knowing about it. Worms could be used also for data collection and espionage activities. An early worm like SirCam had the ability to attach files to its mailings, but now they could remotely search for particular keywords. On the other hand, worms could be used to cause damage, meaning both disruption as well as destruction. Early worms like Chernobyl or Klez had data-erasers commands embedded. Furthermore, worms could take control of physical object not connected to the internet, for example by infecting SCADA (Supervisory control and data acquisition) systems[11], like the Stuxnet worm. Worms could also perform DoS against physical machines, by overloading phone lines, as it happened during the cyber attacks against Georgia in 2008 (Markoff, 2008).

Trojan Horses and Spywares

A trojan horse – or simply trojan – Is, as the name suggest, a file that wants to be perceived as legitimate and harmless but instead hides malicious code in it. The first line that we need to trace is that trojans do not possess the same replication characteristics of viruses and worms, since they rely solely on human activation.

[11] It's interesting to note that many of the payloads hereby listed were described as "not yet seen in the wild" in the all-encompassing and thorough taxonomy provided by Warren et al. The document was published in 2003, and by 2017 all payloads were actually seen in the wild.

A trojan is made by two different parts: a server code, and a client code. The server code is fairly small in size, in order to avoid detection, and it is sent to the target host through social engineering techniques or lured into downloading it. The target opens it, believing it is a legitimate file, such a document, and unwillingly launches the server code that connects to the client code that resides in the attacker machine, which has a console through which it can control the target's machine. Once it has control of the victim's machine, the attacker has complete access to all the files and documents, including the capability of reading, deleting, and modifying them, and also the ability of installing new softwares. Famous trojans like Zeus3, Obad.a and Cryptolocker were used mainly for cybercriminal purposes, hence they were not stealthy, but given the potential of a trojan it is very likely that they could be used for CNEs as well as for CNAs by a state-actor.

Indeed, a trojan codenamed Regin was used for intelligence purposes such as surveillance and espionage activities against offices of several states. It was so sophisticated that the Symantec report stated that only a nation state could have financed and developed it (Symantec, 2014). One of the peculiarities of Regin was the attention to not getting caught, like encryption and custom communication protocols, that allowed the trojan to operate for three years, from 2008 to 2011. Other examples of state-sponsored trojans are found in Titan Rain, which gathered data from defence industry companies in the US between 2003 and 2005 (Kiravuo and Särelä, 2013), and Gauss, a state-sponsored banking trojan which aimed at intercepting data in many Lebanese banks between 2011 and 2012 (Kaspersky Lab, 2013).

As a subcategory of trojans, we could find spywares that, contrary to popular belief, do not constitute a category on their own since they rely on the same methodology used by trojans (Filliol, 2005). A spyware (from "spy" and "malware") is malicious software designed uniquely to collect private information in the form of theft of identity or personal data and in the monitoring of personal activities (Subrahmanian et al, 2015). The payload of spywares consists in searching for specific files, network monitoring, key logging (registering all the keystrokes performed on the keyboard), and access to microphone and camera. All the data collected are sent to the client (Kiravuo and Särelä, 2013) Spywares are static pieces of malware therefore they must be installed on every single computer to siphon information from specific user.

This characteristic has an obvious shortcoming, that is to say that if the same spyware is installed on a large scale – for instance in a government department which has hundreds of machines connected in a Local Area Network (LAN) – the server bandwidth will saturate fast - due to the spyware attempts to connect externally - leading to the discovery of the malware (Filliol, 2005). That is why large espionage campaigns are left to trojan horses, but with the recent technological evolutions the use of spywares has simply adapted to more useful purposes. Indeed, spywares recently began targeting smartphones, which today are very likely to be synchronised with professional e-mail accounts and cloud storage systems. In the last few years examples of mobile spyware have been prolific. On one instance, a Chinese cyber espionage operation (dubbed APT28/Operation pawn storm) targeted hundreds of iPhones (the spyware was custom made for iOS, the operative system of Apple's iPhones) belonging to various financial organisations' representatives (Skhandar, 2015) Furthermore, there was a diffusion of spywares made by private companies and sold to numerous governments for both foreign as well as domestic espionage (for monitoring and identifying dissidents, for example), for example the ones sold by the Italian company Hacking Team and its direct competitor FinFisher GmbH (German) (Marczak et al., 2015).

Therefore, it is interesting to see how, as described in the previous theoretical part, malwares like worms and trojans, but also some types of espionage viruses, actually vary their degree of sophistication depending on the capability of the attacker and also on the objective and target of the CNO. States, straightforwardly, could invest much more compared to the single hacker and this is directly linked to the fact that they have more complex objectives and they want to cause more damage (whether it is a loss of data or a temporary disruption of services) by staying as stealthier as possible. Therefore, we could infer that states have an horizon set much more in the longer term compared to other actors that utilise the "same" methods. On the contrary, cybercriminals or hacktivists think in the short term because they want to reap an immediate gain, whether is to gain money or to send a political message.

Blended Threats

Blended threats, as the name suggests, are advanced cyber weapons that cannot be limited in their definition to the characteristics of malware, as far as functionality, payload and objectives are concerned. Indeed, the characteristics of a blended threat consist in: causing harm, for example launching Denial of Service attacks and extended intelligence gathering; propagation through multiple methods, like scanning for different vulnerabilities, embedding code in HTML, infected USB or e-mail attachments; attacks from multiple sides, like infected executable files, privilege escalation through social engineering methods, compromised network shares, and script code to HTML files; static or self-replicating characteristics, sometimes automatic scanning for vulnerabilities and also self-deleting if the path it followed is erroneous; exploitation of multiple kinds of vulnerabilities at the same time, like known and, most importantly, unknown vulnerabilities (0days) (Wall, 2009). It is important to underline that a blended threat could not have *all* of these characteristics but just some of it, and that's the interesting part of this kind of threat, that is to say the high degree of adaptability combined to a high level of sophistication. This means that a blended threat combines the best methods to exploit different vulnerabilities in a single attack, maximising the velocity of contagion and the resulting damage.

DoS and DDoS - Denial of Service and Distributed Denial of Service

The first type of blended threat to be analysed is also the first one, chronologically, that appeared in cyberspace, that is to say Denial of Service (DoS) and Distributed Denial of Service (DDoS) attacks. Nonetheless, in order to understand DDoS attacks and other types of blended attacks, the concept of "botnet" must be understood. A "botnet" (union of "robot network") is a network of so-called zombie machines that can be controlled remotely and, in order to do this, a high number of machines have to be infected with software that allows remote administration, such as a trojan (Wall, 2009). Therefore, a botnet is essentially a list collecting addresses of computers that are already infected and ready to be controlled; these botnets could also be sold or rented, and for all these reasons they prove a challenge for regulation and attribution mechanisms. Botnets are widely used by cybercriminals but also by state-sponsored actors during cyber disputes, such as the Russia against Estonia in 2007 and Russia versus Georgia in 2009.

DoS and DDoS attacks are two of the most dangerous cyber weapons a state can deploy since they are difficult to foresee and to detect early, due to the fact that the packets they use are part of legitimate traffic, and most importantly in the past decade the availability of free tools to perform this kind of attacks has increased dramatically (Infosec, 2014). Put simply, DoS and DDoS both exploit the communication system with the machines. The payload of DoS and DDoS consists in flooding the request queue of a server with fake requests, resulting in the unavailability of a system or network that results as offline, due to the fact that the server is incapable of handling all these requests (Infosec, 2014).

We could roughly divide DoS and DDoS attacks in three categories: volume based attacks, application layer attacks, and protocol attacks. Volume based attacks are measured in bits per second and aim at the saturation of the bandwidth of the target; application layer attacks which goal is crashing the web server by targeting applications and server resources, and for this reason they are measured in requests per seconds; protocol attacks (such as the SYN flood described below) are measured in packets per second and their objective is to consume server resources and the ones of the intermediate equipment like, for example, firewalls (Infosec, 2014). Nonetheless, it can be easily inferred that the objective of DoS and DDoS attacks is to compromise the availability of information.

Between the two attacks, DDoS attacks are far more dangerous than DoS attacks, and in order to understand that it is necessary to outline the difference between the two cyber weapons. A DoS attack consists in an attack starting from one machine versus one server through a single internet connection. An example of DoS attack is the so-called SYN Flood that exploits a known vulnerability in the TCP protocol, in particular in the its "three-way-handshake". In this type of connection, the communication happens in a three-step fashion (hence the name): first a client sends a TCP SYN packet to a server, this server in response to the SYN packet sends a SYN/ACK packet, and the client re-responds with an ACK packet. To perform a SYN Flood, a client sends a huge number of SYN packets without responding with the ACK one leaving the connections half open, overloading the server and exhausting its resources (Raghavan and Dawson, 2011).

On the other hand, we can have two types of DDoS attacks: automated or human-coordinated. The first one employs the aforementioned botnet: we have a single machine, where the client program is running, that controls thousands, hundred thousands and sometimes also millions of zombie machines (Joeng et al, 2011) that perform a DoS attack against a target. In this case a DDoS could be of two different kinds, that is to say an amplification or a reflection attack. An amplification attack starts from a machine that tells to the botnet it controls to send a packet to a specific target, in this case the original power of the client machine is amplified by a botnet. In a reflection attack, the attacker "spoofs" (simulates) the target IP and sends a communication request to the botnet which in turn starts responding in massive numbers, exhausting the target machine's resources (Raghavan and Dawson, 2011). Human-coordinated DDoS attack use voluntary botnets, such as the peer-to-peer botnet attack used for example in Estonia in 2007 (Raghavan and Dawson, 2011) where a high number of people try to flood a certain target with single DoS actions. Here the issues concerning attribution and IP blocking for prevention prove worthless, since the attacks come from different sources, different IPs in different countries.

APT - Advanced Persistent Threat

Advanced persistent threats or APTs are a relatively new categorisation for a certain kind of cyber weapons and therefore their definition is more debated than the others (Sasanapuri et al, 2016). Nonetheless, we can split the name to identify and describe three different parts. Straightforwardly, "advanced" describes a high degree of sophistication of the cyber weapon, depending on the quantity and quality of the malwares employed in order to exploit different vulnerabilities in the system; "persistent" indicates an attack that lasts over time – obviously in a stealth fashion – therefore indicating a command and control chain between the client machine and the target one; "threat" refers not only to the "menace" posed to the target system, but also to the human orchestration of the attack.

What distinguishes APTs from many other attacks is the fact that their source is an actor that has conspicuous funds to perform a very sophisticated and multi-faceted attack that is long-lasting, and therefore stealthy, against well-defined targets to reach specific objectives.

As far as the source is concerned, it is usually a government and the performance could be direct or indirect. This means that the attack could be carried out by a state military unit specialised in cyber operations, or it could be performed by a private entity sponsored by a government. This kind of actor brings with itself more funds and resources than other actors could allocate, such as military experts, strategists, technical experts, state-of-the art machinery, combined with the capability of buying many zero day exploits (Chen et al, 2011). The sophistication of the attack, the "advanced" part, means that no single individual could perform such attack, but the perpetrators are usually States or private actors with enough funds to sponsor such activity. For the purpose of this research the focus will be only on APTs used by state actors.

Before that, there is a constant repetition of attempts to attack the target system, which draws a line when compared to other cyber weapons that perform widespread attacks at first (Chen et al, 2011). APTs usually exploit 0days and are blended and polymorphic threats, meaning that they use different malwares to penetrate the system and exploit vulnerabilities and that these malwares could be updated and therefore be changed, rendering many anti malware detection systems useless (Sasanapuri et al, 2016). Not only, this polymorphic nature refers to the “persistent” part of the attack, if an hypothetical “Plan A” to penetrate a system or, in a secondary moment, to exploit a particular vulnerability has failed, then the attacker will try to perform the same action with a “Plan B” that is to say with a different method, through a different vulnerability or through a different malware, and in case that even “Plan B” doesn’t work, a “Plan C” will be put in place, and so on and so forth.

Given the funds allocated, the sophistication and the persistence, and in order to reach its objective the APT needs to stay undetected for as long as possible. This cyber weapon has the same “problem” as simple trojan, that is to say it generates and receives traffic. This traffic needs to be concealed as legitimate inside the target’s whole communication traffic, using brief communications, so that it does not arise suspicions. Furthermore, exploiting 0day vulnerabilities avoids the detection of the APT by signature-based security softwares, the concealment is strengthened by employing encryption that reduces the detectability of traffic (Chen et al, 2012).

Lastly, APTs are not broad attacks, but they are “targeted”, meaning that they are custom built for a certain purpose and for a certain target. The targets are usually governments and the military and industrial complex affiliated to it. Indeed, among the objectives of APTs there is strategically important data, like national security documents, intellectual property, secret plans, and trade secrets (Chen et al, 2012).

Since an APT is so complex, we could define it as a sort of “mini campaign” that consists in different stages. Firstly, there is a reconnaissance phase, where the attacker scans for vulnerabilities in the system. It is important to underline that with “vulnerabilities” we include also human vulnerabilities. Indeed, social engineering methods have proved to be very effective in this first phase: for example, it is relatively easy to find a penetration point through phishing or better spear-phishing; by aggregating information publicly available online about a certain person (so called OSINT, Open Source Intelligence); and even through “watering hole attacks”, meaning infecting – for example – websites that are known to be frequently visited by the target (Chen et al, 2012; Sasanapuri et al, 2016). In this case, it is sufficient that a target opens an infected e-mail attachment or visits a certain website believed as legitimate to become an entry point for the attacker, which steals the victim’s credentials and then manages to escalate its privileges.

Once in the system the attacker is able to steal information while establishing persistence, through backdoors, and control of the system by installing a series of custom tools in order to build a communication of command and control (Sasanapuri et al, 2016). The traffic produced by an APT could be one of the shortcomings of APTs, since security companies are advising to spot the so called “callbacks” that is to say the traffic flowing from the infected machines to the command and control ones (FireEye, 2013). Indeed, in order to avoid traffic detection problems, APTs use different methods, such as exploiting blogs and social networks account to send control commands, using the TOR (The Onion Router) network which provides anonymity through hidden services, and the aforementioned remote access tools (RATs) (Chen et al, 2012).

Once the command and control communication system has been established, the attacker will perform the so called “Lateral movement” infecting as many machines in the network in order to steal as much information as possible or to amplify the effects in case of a disruptive attack. In this stage the APT will perform a mapping of the network (a sort of secondary, internal, reconnaissance phase), infecting other systems and escalating privileges further, and identifying where the sensible data is stored. This stage is crucial and indeed takes the longest, given the fact that, in order to avoid detection, the weapon must run slowly and also due to the fact that the attacker wants to reach as many parts of the system as possible in order to exfiltrate the maximum amount of data (Chen et al, 2012). Furthermore, the freedom of communication and movement inside the target’s system allows also the attacker to update the existing tools and even change them according to the needs of the operation. The last stage is the actual payload, that is to say the siphoning of the data or, possibly, the disruption of the target system. Once the location of the critical data has been established, this is compressed, encrypted and channeled to an internal, compromised, server and then sent externally often using secure protocols to amplify the stealthy procedure (Chen et al, 2012).

Knowing about the functioning of cyber weapons is a useful tool that permits better analysis when examining cyber conflicts between states. Nonetheless, this part of the chapter does not pretend to be exhaustive for two main reason. First, other tools to penetrate systems exist, but they are limited to the cyber criminal world, such as ransomware and scareware. Although they could impair the functioning of a systems and are a serious and growing threat to the cyber security environment, they do not classify as a weapon used by militaries. Another “weapon” that is not mentioned above is what is usually called a “logic bomb”. In reality, defining something as a logic bomb does not really classify another kind of cyber weapon, but it only defines the method of delivering the payload, that is to say only predefined conditions could trigger the activation of any kind of malware. Nonetheless, the payload is the one of the malware to which the “timer” is attached to. The second reason is that new cyber weapons could already exist but have not been deployed yet and therefore not seen in the wild.

Without resorting to speculations, it is safe to say that cyber weapons are being developed every day and the trend that could be inferred by the analysis above is that they are continuously evolving. New cyber weapons even resort to alter the fingerprints of a given weapon, thinning the chances to provide a clear attribution. For this reason, full knowledge about the state of the art of cyber weapons is needed both to analyse past and present conflicts, but also to provide regulations on cyber weapons in order to manage future conflicts. What has been done under a normative point of view, and what should be done to catch up to the fast development of cyber weapons, is outlined in the following part.

The problem of regulating Cyber Weapons

As far as 2017, the so-called cyber weapons are not internationally regulated, even though various groups of individuals and intergovernmental organisations tried to develop rules and regulations to manage the military exploitation of cyberspace. The main example of these attempts could probably be found in the Tallinn Manual on the International Law Applicable to Cyber Warfare, an academic study that applies international law to cyber conflicts and cyber warfare, written by an "International Group of Experts". The focus of the Tallinn Manual is the *jus ad bellum*, that manages how States resort to force, and the *jus in bello*, that deals with the way in which warfare is conducted.[12] The first version of the Tallinn Manual was published in 2013 and it was followed by a second version, Tallinn 2.0, published in February 2017. Both versions are somewhat influenced by what could be described as the *zeitgeist* on the perceptions of cyber conflicts and warfare. For example, the first version focused on disruptive and destructive attacks (Schmitt, 2013). It is straightforward that the experiences in Estonia and in Georgia, and the one of Stuxnet heavily influenced how cyber conflicts were perceived at the time, since they posed – and still do – the greatest threat to States.

Once light has been shed on these threat, version 2.0 focuses on the legal framework that should be applied for conducting such CNOs. For example, among others, it takes into consideration human rights, diplomatic law, the responsibility of international organisations, international telecommunications law, peace operations and the so-called peace-time

[12] For further information see ccdcoe.org/research.html

international law (Schmitt, 2017). Unfortunately, despite the fact of being one the most comprehensive document on the governance of cyber conflicts, the Tallinn Manual is a non-binding, unofficial document and cannot be applied – at least legally – in any occasion. Nonetheless, given exactly the fact that it constitutes the first as well as the broadest effort to apply international law to cyberspace, it is widely regarded as the document around which States should shape their cyber strategies and legal framework about cyberspace on.

Other attempts at regulating the use of cyber weapons could be found in the many discussions that took place for extending NATO's article 5 regarding collective defence to cyber attacks as well. The last steps in this direction were made this year after the Warsaw conference. During the conference, NATO members recognised cyberspace as an operational domain on par with air, sea and land and, as a result, cyber defence became part of collective defence (NATO, 2016). The main focus of NATO's strategy in cyberspace it is obviously cyber defence and cooperation among the members of the alliance, but it also mentions the capability to respond to cyber threats (*ibidem*). This is where things tend to be trickier. NATO Secretary General, Jens Stoltenberg, stated that: "a severe cyber attack may be classified as a case for the alliance. Then NATO can and must react […] how, that will depend on the severity of the attack." (Lenoir, 2016).

These postures bring several implications with them. For example, recognising cyberspace as an operational domain to which collective defence could be applied is a strong deterrent factor against the utilisation of cyber weapons. The *how*, despite the deterrence given by the possibility of a collective reaction still poses a problem. Indeed, it is important to underline that the possible reaction would not be necessarily carried out in cyberspace, it could happen in the form of economic sanctions and even the possibility of kinetic intervention must be taken into consideration, which would provide another kind of deterrent.

Nonetheless, the right theoretical steps in regulating the use of cyber weapons are at least clear for NATO policy makers: be certain in attribution process and define precise legal measures for addressing different cyber weapons, whether they are akin to armed conflict or espionage (Fidler et al, 2016). Surely, these are the main points to be regulated, but they are also the hardest to address, due to the absence of proper modifications to the architecture of cyberspace in the past and an international legislative corpus that entered into force prior to the advent of cyber weapons as tools of statecraft.

One example of a first step taken towards the regulation of the cyber realm can be found in the Convention of Cybercrime held in Budapest in 2001 (and entered into force in

2004) that, as the name implies, is limited to criminal activities carried out in cyberspace (European Council, 2001). The Convention was drafted by the Council of Europe with the aid of the United States, Canada, South Africa and Japan and, as of 2016, it was ratified by 53 states, only signed by 4 (European Council, 2016). The Convention asks for enhanced cooperation among the signatories, including the harmonisation of national laws regarding cybercriminal activities and the sharing of intelligence within the members of the Convention. As much desirable and beneficial for international security, the Convention met the opposition of Russia, who refused to sign providing allegations that the Convention constitutes a violation of national sovereignty (Giles, 2012). Nonetheless, the Convention should be an example to look for, both in terms of approach, namely the cooperation among countries that would reduce the uncertainty of attribution if many countries are sharing information about the traffic of data, and also in terms of the problems to expect when proposing an international regulations on cyber weapons, that is to say the reluctancy of countries like Russia to collaborate in this matter.

Another step in regulating the use of cyber weapons for espionage purposes is the 2015 US-China bilateral cyber agreement. The two countries pledged not to perform economic espionage through cyber means one against the other, and the two actors – embodied by former US president Barack Obama and Chinese president Xi Jinping – stated their will to cooperate and share information to reduce the number of cyber incursions (Brown and Yung, 2017a). This agreement surely constitutes a step forward in bilateral regulations for international cyber security, but the facts indicate that the agreement was built around nothing substantial, and was met with suspicion due to the fact that during the meeting that brought to the agreement, the Chinese side never admitted having deployed cyber weapons against the US (Brown and Yung, 2017b).

Furthermore, this absence of substantiality could be linked to the fact that the agreement establishes the *"commitment by each country that it will not be the first to use cyberweapons to cripple the other's critical infrastructure during peacetime"* (Sanger, 2015). The quote implies two important elements that need to be underlined. First, the use of the verb "cripple", which implies an important degree of disruption both in terms of severity as well as duration. Therefore, everything under a "cripple" level should be regarded as accepted, nullifying every effort against weapons for espionage purposes. The second element is "peacetime", that is interesting because it could imply that bilateral agreements could work in place of an absent *jus ad bellum* in cyber disputes involving two countries, but the agreement does not specify the consequences in case of breaching of the accord and an infrastructure is, in fact, crippled. Nonetheless, multilateral measures that are beginning to involve political commitments are beginning to surface, such as the recent G7 declaration on responsible states behaviour in cyberspace, which underlines the possible application of the UN Charter to the employment of cyber weapons given that under some circumstances they fall under the framework of use of force and armed attack, that could possibly lead to exercise the right of self-defence. Furthermore, it pushes towards an increase in cooperation among signatory states that results in information sharing and mutual assistance in case of cyber threats (G7, 2017).

Cyber Weapons and the Law of Armed Conflicts

The government of *jus ad bellum,* the legislative corpus that regulates the legality of the use of force by one state against another, relies on the United Nations Charter, its interpretation and all the international bodies of law derived from the Charter itself. The Charter, with its article 2(4) states that "the threat or use of force against the territorial integrity or political independence of any state, or in any other manner inconsistent with the Purposes of the United Nations" is prohibited for any nations except in those instances where the peace is threatened or, in case that there is a breach of international peace and security, to restore it (United Nations, 1945). The main issue is that the concepts of "threat" and "use" of force and also "armed attack" (mentioned in Article 51, which defines it as the main condition for exercising the right of self-defence) are not specifically defined, but instead rely on common, shared, understandings of what constitutes, and constitutes not, both terms.

As for now, the employment of cyber weapons does not constitute – legally – neither threats nor actual use of force, nor even armed attacks, given the fact that they are not framed as such due to, maybe, the absence of an important precedent, as mentioned in the Tallinn Manual. Furthermore, most of the time, the employment of cyber weapons is aimed at information gathering and since espionage does not constitute an act of war, it is easy to dismiss the attack as such. The problem is that there is the risk of downplaying the actual threat posed by cyber weapons. The issue is that the United Nations Charter was drafted in a sort of "pre-cyber" period, and most importantly, before there was actual evidence that cyber weapons could provoke physical damage and could heavily interfere with the functioning of civil society, as well as governmental and military activities (even without causing casualties and destructive effects), due to the high degree of reliance on cyberspace – as a whole – upon which modern societies rest today (Lin, 2010).

Given this impact, it could be argued that impairing the correct functioning of the whole infrastructure could be intended as an armed attack just as a kinetic attack with the same effects would be regarded as such (Lin, 2010). As far as the "threat" of the use of force is concerned, there are no actual statements from any state against other states to use cyber weapons actively, but there are increasing statements of government officials underlining that many militaries around the globe are developing offensive cyber capabilities (Breene, 2016). These statements could not really be perceived as threats, but they are a method of signalling among countries stating that they possess capabilities, resources and the will to deploy cyber weapons if they are needed. For example, North Korea threatened the US to attack if the movie "the interview" containing shaming of DPRK's leader Kim Jong Un was not withdrawn from theatres, and later attacked Sony through cyber means (Pagilery, 2014).

Nevertheless, threats to the use of force are being displayed as far as hypothetical responses are concerned. For example, in 2011 the US Pentagon compared cyber attacks to "acts of war" to which the response could be through cyber means but not limited to them, stating: *"We reserve the right to use all necessary means – diplomatic, informational, military, and economic – as appropriate and consistent with applicable international law, in order to defend our nation, our allies, our partners and our interests."* (Brookes, 2011). The problem of this statement lies in the "appropriate and consistent with applicable international law" which is, as it is outlined in this chapter, a grey area in international law. Nonetheless, the US held true to this approach, retaliating against North Korean attacks after the Sony hack, by economic sanctions and also, allegedly, cyber attacks (Locatelli, 2015).

As far as *jus in bello* is concerned, cyber weapons have been employed during belligerent actions such in the cases of Estonia in 2007 and Georgia in 2008, but the far more destructive capabilities of kinetic weapons leave cyber weapons in a far less important position (*Ibidem*). Nonetheless, the position of experts in regulating the employment of cyber weapons among belligerent states has proved more complicated than regulating weapons within *jus ad bellum* framework. For example, due to the fact that today commercial and economic activities are mainly carried out on-line, interrupting these activities through cyber weapons targeting the management systems of those enterprises could be compared to a naval blockade in the physical world (Russell, 2014). However, the experts that drafted the second volume of the Tallinn Manual were divided on this analogy and ultimately dismissed the possibility to compare crippling cyber attacks to naval blockades as far as rules are concerned (Schmitt, 2016).

One of the main problems about regulating cyber weapons is that the current international legal framework are concentrated on the effects of those weapons, more precisely on the "scale and effects" as the Tallinn Manual states (Schmitt, 2013). However, the approach should gain different perspectives that would help avoiding getting stuck in this normative impasse. For example, on the one hand concentrating on the effects helps in differentiating between disruptive attacks and espionage attacks; on the other hand it downplays and focuses only on the deployment of the payload. The rationale behind this line of argument is that there is a constant, albeit justified, tendency to try to link cyber weapons to kinetic weapons and try to find an equivalence principle. This is somehow misleading, as one cannot compare a virus to a gun, a DDoS to a machine gun, or saying that a worm is "stronger" than a virus. For this same reason it is very difficult to analyse an exchange of offensive CNOs between two countries in terms of escalation.

Operations involving CNEs are usually dismissed with "espionage is not an act of war" and therefore they do not fall under *jus ad bellum*'s jurisdiction. The certainty of an operation aimed at espionage arrives when a completed attack (meaning that it actually stole data and information) has been found and analysed; but what about when a cyber weapon is found inside a system prior to the delivery of its payload? If there is a RAT inside it, how can an analyst be sure of its real purpose? Could it be dismissed as a simple intrusion?

It seems farfetched, but an attack starts with an intrusion, an illicit one to say the least. If we still take an espionage attack as an example, a spy is as good until it is not caught and if it's caught on enemy's soil it is considered as trespassing and falls under a normative framework that regulates such an activity. This could be done also for cyber operations aimed at espionage. An intrusion should be considered a foreign presence on another country soil, given that the servers reside on national ground. Even when disruptive and destructive effects are not produced, we should not forget that even incursions could be regulated by the international law. In light of these, even if a breach is detected without doing harm it could be linked to a reconnaissance phase and it could be compared to reconnaissance performed by aerial incursions, which constitute a violation of territoriality and also a violation of international law (Oduntan, 2011). Furthermore, if this intrusion is followed by a control gain over the enemy system – that is to say the phase after the reconnaissance one – in order to establish a foothold, it should be viewed as a proper invasion, comparable to a platoon of soldiers taking control of a turret. As described before, cyber weapons could be updated and changed remotely, and for this reason, once inside a network or system, the payload could be changed from espionage-driven to a disruptive one. Treating a foreign presence inside a system in this way would also eliminate this problem of insecurity about the payload without the actual deployment. Therefore the principle of territoriality should be extended to cyberspace during matters of espionage and attack just as it is done to the cybercriminal environment.

Going back to the "scale and effect", the UN Charter and the Tallinn Manual always refer to physical effects when talking about "an act of war", but it could be regarded as pointless to treat physicality so important in a virtual environment. Talking again about the cybercriminal environment, today money is in the form of data, as it could be transferred almost instantly from one account to the other, and it is physical when withdrawn. Nevertheless, malicious hackers could steal money from bank accounts in the form of data, and still it is considered a crime. Something similar should happen in the state vs state cyber disputes, if a cyber attack where actions that produce or could produce effects that interfere with the functioning of infrastructure in the same way kinetic actions would. More clearly, if a foreign state launches a kinetic on an electric grid and this stops its functioning for a given period, it should be considered an act of war by the *jus ad bellum*, due to the fact that is an aggression coming from a foreign military. If the same halt in the operations of the infrastructure would result from a deployment of a cyber weapon it should be regarded in the same way, given the fact that the effect produced is the same and the perpetrator too.

One could wonder why the international regulation of cyber weapons is very late despite their increasing use by state actors. One reason could be that there is the feeling that many of those cyber weapons are employed for espionage, which, according to the UN charter and general belief, does not constitute an armed attack, neither use of force. However, despite being rare, disrupting attacks (with destructive consequences) happened and it is very likely that will happen in the future. Stuxnet, the Ukrainian electrical grid, are only two examples of cyber weapons used for harmful purposes. It is indeed true that certainty of attribution constitutes an obstacle to be able for a state to legally contest to another state to have employed a cyber weapon and then call for the right of self-defence, which could escalate in the non cyber environment.

However, there are a couple of episodes which can constitute a precedent to overcome the impasse posed by the attribution issue, that is to say the so-called "Sony hack" and, to a certain extent, also the APT1 investigation. Without going in the details of the event, it is sufficient to say that a CNE operation was perpetrated against Sony

Entertainment, and the US government were certain about the origin of the attack and who the culprit was, namely North Korea. The certainty was due to the fact that the NSA was already inside North Korean systems, where all the details of the operation were stored (Haggard and Lindsay, 2015). Given this absence of doubt, the US retaliated by means of economic sanctions and, allegedly, CNOs disrupting temporarily the North Korean internet system (Strohm, 2015). This is to say that the current, common uncertainty of attribution, meaning that not every state in the world possess the capability to penetrate other state systems and therefore being able to discover with absolute certainty who the culprit of a offensive cyber operation is, should not constitute an impairment in developing a functioning framework regulating the usage of cyber weapons.

The point is that certain states, albeit few, possess the means to be certain about attribution even though these means fall under covert espionage operations that, according to the UN Charter, do not constitute threat or actual us of force, neither armed attacks. If this could be proven, for example, at UN level then the right of self-defence could be invoked. Of course, the shortcoming of this reasoning is that it seems to advocate an increase in espionage operations among states. That could be argued as true, but the reality is that not every state could be able to penetrate a government or military complex of more secure states. This could provide a sort of asymmetric deterrence that would make some countries refrain from attack fearing that the target might be aware of the ongoing cyber operation.

In the second case, the culprit of the APT1 campaign was found to be PLA Unit 61398 thanks to good old intelligence in the form of GEOINT and OSINT, combining the IP addresses from which the attack seemed to originate from – Pudong, China – and publicly available information about the location of the headquarter of that unit, namely Pudong, China (Sanger, Barboza, Perlroth, 2013). The peculiarity of this investigation is that it was conducted by a private companycalled Mandiant: this to say that intelligence agencies should have in their hands capabilities far superior to the one of a private company, such as SIGINT and HUMINT, providing better means to dissipate the fogs around uncertainty of attribution.

Another issue in regulating cyber weapons could be the one of stockpiling. The end of the Cold War was characterised also by arms control agreements and measurements of the nuclear arsenal of the two superpowers. Also today the presence and the numbers of the various warheads are perfectly known or easily inferred. The same does not hold for cyber weapons, and discussion about stockpiling (Goldsmith, 2015; Denning, 2000; Arimatsu, 2012;) in this domain are pointless. The main reason is that cyber weapons possess certain characteristics that are not shared with other kinetic weapons.

The reason that could be considered the most important is the “temporary” nature of cyber weapons, meaning that once the weapon is used and, after a certain time, discovered, than its no more as useful as it was in the beginning. This happens for different reasons that are somehow overlapping: first of all, once the vulnerability that has been exploited is identified it is patched, therefore it is no longer usable. Of course, the patching of certain vulnerabilities takes time and once it is available and it is not automatic and simultaneous. Some actors could decide – moved by ignorance – not to patch or to delay the patching of the system, extending the window of opportunity of the attacker, that still it is limited once there is a patch. Again, the WannaCry attacks must be mentioned because due to negligence in not patching the risk of having indirect victims was very high because the ransomware made also hospital computers unavailable in the UK which led to emergency transfer of the most serious patients to institutions that were not attacked (Sherr, 2017). Second, some peculiar cyber weapons, namely APTs, are conceived and design to attack certain specific targets, once they are found it is supposed that they have reached their objectives, or part of and furthermore the target will raise its defences against external attacks.

Third, it must not be forgotten how cyber weapons for as carefully designed by experts as they can be they still suffer the shortcoming of uncertainty of control that brings uncertainty in usage, therefore the particular objective that a state seeks to reach by employing a cyber weapon must trump these three reasons. Indeed, all these reasons explain how a country would not have different USB keys in an arsenal dubbed “Stuxnet 1” “Stuxnet 2” and so on and so forth but it will deploy the cyber weapon only once, because the next one it uses will be necessarily different. What could actually be “stockpiled” and be readily available are 0day exploits – that work uniquely with 0day vulnerabilities that were previously found and studied – but it is fairly obvious that, as it was stated before, as they are unknown, there is no method whatsoever to understand how many 0days exploits are in the hands of a state. Therefore, the cyber power of a certain state is to be guessed and inferred through official statements and, of course, to alleged cyber attacks that could be linked to this particular state. Indeed, it could be that a certain country would try to deploy a cyber weapon as a method of signalling, therefore as a mean of deterrence too.

The last, but not least, problem about regulating the use of cyber weapons is the *laissez-faire* approach that is comparable to the one that surrounds espionage. In light of all what was described until now, the fact that:

a) cyber weapons are harmless in terms of physical destruction (considering Stuxnet as a unique and isolated case) and do not provoke casualties;

b) cyber weapons permit a sort of "unofficial signaling" among nations in order to display power stances without resorting to kinetic weapons or official statements;

c) cyber weapons are not comparable to kinetic weapons as far as equivalence is concerned and this characteristic permits countries to attack and retaliate with different methods in the cyber realm and also to balance some – kinetic – asymmetries between countries;

d) even though destructive cyber weapons are not easily deployable for their temporary nature and lack of total control, in absence of international regulations, disruptive cyber weapons constitute a very convenient mean to settle scores internationally without incurring in retaliations and military interventions;

e) cyber weapons generally suffer of the shortcoming of the uncertainty of attribution;

f) the number of countries that are improving or implementing new cyber capabilities is increasing ("everyone is doing it");

all this does not incentivise a real push towards a regulation of cyber weapons comparable to a kinetic threat of actual use of force. But given the fact that there is the possibility that cyber weapons increase in sophistication, as it has been state before, some states have already or are drafting national legislations regarding attack and defence in cyberspace.

The problem is that these are national measures that cannot bring but unilateral consequences, such as sanction (like in the Sony hack case) and nothing much. Linked to this, a possible regulation would face other problems of political nature. As it has been stated before, countries like Russia do not like to cooperate and share information in cyber matters, as many other countries are developing cyber weapons after they've seen the advantages that they bring, especially CNEs there could be a sort of "Kyoto protocol situation" where cyber-developing countries refuse to sign or participate to an international regulation on cyber weapons because they also want to reap the fruits of this legislative gap - such as economic and military espionage – as other state have been doing for the past years. If these countries' possible denial would not be met with other kind of incentives, this lack of cooperation would push a scenario of better international cyber security even more in the future.

If international regulations seem difficult, then also bilateral agreements – as praiseworthy as they can be – have yet to prove their usefulness and dismiss the fear of being only façade accords. Albeit there is only one example up until now, the US-China cyber agreement suffers of the shortcoming of raising the stakes very high, at "Stuxnet level" and implicitly allowing all the activities under this threshold. It surely constitutes a political step forward but not as far as a concrete regulation. China has been using cyber weapons extensively against the United States to extract sensitive and secret data in order to reduce the technological and economical gap between the two countries and there are no current incentives to put an halt on these activities.

Nonetheless, on the political level, an increase in multilateral treaties and agreements even though not all-inclusive would be a huge step forward in three different directions. First and foremost, it would straightforwardly constitute a huge increase as far as international cybersecurity is concerned. Second, it would constitute a a possible and significant decrease in the uncertainty of attribution. The supposed increase in the quantity and the quality of information shared would be hugely beneficial to the the traceability of the path followed by a given cyber weapon, that could be more easily backtracked to the location of origin. This is directly linked to the third direction, that is to say a strong deterrent for the countries that are both inside and outside the signatories of these agreements.

Taking all these problems into consideration, and the unlikeliness of an all-encompassing international regulation on cyber weapons, the only way to tacke a worst case scenario, namely an indirect casualty caused by a cyber attack, it is still to define new norms for the cyber era – like the ones outlined in the Tallinn Manual implemented with the provisions on physicality and territoriality described above by this research. These are capable of regulating the use of cyber weapons in the *jus ad bellum* framework without the actual heavy link with kinetic use of force. This could provide, given all the shortcomings surrounding the use of cyber weapons at least a sort of deterrent factor to their future use.

Chapter III: Case Studies

Introduction

As stated throughout the document, the variable of power seems to be one of the most influencing aspect in cyber disputes. Cyber power is influenced by the classic concept of power which, in turn, is shaped by the budget lines allocated for military expenditures. More resources mean more effective cyber weapons.

It seems to be the case that states resort to CNO to project this power in given geopolitical situations regardless of situations of escalation dominance.

What is power? According to Dahl, power is when “A has power over B to the extent that he can get B to do something that B would not otherwise do” (Dahl, 1957). This means that a state A has enough resources, instruments or means to influence the actions of state B in line with the amount of power that this state A has over state B (*Ibidem*). In Dahl’s definition power is a coercive variable, a zero-sum game from state A in state B’s regards. The power of A over B, however, does not eliminate the possibility of B gaining power over A, therefore power is temporary. For example, if we take the Stuxnet case, state A, the United States of America, exercised power over B, Iran, by attacking the nuclear power enrichment facility of Natanz displaying power in order to reach a political objective. In this sense, we could say the Stuxnet influenced the P5+1 talks favourably towards the US. However, militarily, Iran retaliated through cyber means against the US. Therefore the coercion could be considered political but not military in this sense.

According to Foucault, power is based on a discourse intended as a production of truth (Foucault, 1980). Indeed, following this concept, power is continuously established between two actors, and this ever-evolving discourse helps establishing the mutual perception of the actors involved. For this reason, the Foucault concept of power resembles the interactions of states in the cyber domain, as CNOs are also used to signal a certain amount of capability in order to consolidate a perception of power between two states.

Joseph Nye, while helping in shaping the definition of cyber power, gave also a broader definition of power. In his view, power could be divided in hard power, soft power and smart power, that is the combination of the first kind, namely coercion, and the second one, namely attractiveness (Nye, 2011). To those three, Nye adds cyber power, that is “a set of resources that relate to the creation, control, and communication of electronic and computer-based information" and also "the ability to obtain preferred outcomes through the use of the electronically interconnected information resources of the cyber domain" both within as well as outside cyberspace (Nye, 2011).

Indeed, all three major definitions of power help in analysing one of the main factors shaping offensive exchanges in the cyber domain. It could be coercion, as well signalling and the establishment of power, with actions carried out in the cyber domain but that could have consequence within the domain (for example loss of data, defacing of websites, temporary halt of systems operations) as well as outside the cyber realm, that is to say physical consequences (machinery that stops working, such as centrifuges in uranium enrichment facility or systems that manage electrical grids, but also malware that impedes the access to vital data, such as clinical charts in hospitals).

In this dissertation, the variable of power, in order to assess symmetry or asymmetry, is derived from the 2017 Global Fire Power index 2017 which is a power index assessing world's militaries based on more than 50 factors. Indeed, following the index US, Russia and China could be considered as symmetric, and the US military is definitely asymmetric compared to Iran and South Korea.

Power defines how cyber disputes are carried out, because the collision of two different, or equal, degrees of power shapes retaliation and escalation processes. These could see an increasing amount of intensity in the CNOs, therefore prompting a vertical escalation, but also it could mean that the attacker simply changes the target to attack, even with less intensity, prompting what we could consider an horizontal shift in the chain of events.

The application of power in cyberspace is influenced by four main factors. The first one is the lack of an international shared framework of norms that could regulate the states' behaviour and the usage of cyber weapons. Without a normative corpus together with a system able to impose sanctions, states are likely to feel legitimate to perform CNOs since it is not illegal to do so. This characteristic could lead one to think that a deliberate increase of CNOs should be expected since there are no legal limits to perform them. Such a reasoning it is partly true, it has been already stated in the previous chapters there is an increase in CNOs, however we do not see deliberate escalations or huge campaigns of disruptive attacks. Why? Indeed, it seems the case that the other two factors that influence the application of power in cyberspace are constituted by self-imposed limits applied by states when performing CNOs.

The second factor is what we could call "the fear of the unknown", namely the fear of the possibility that the cyber weapon could surpass the boundaries of the target infrastructure, with a spill-over effect that could a) overcome the limits of the damage intended, creating an event that could spur an harsher retaliation, namely an accidental escalation; b) overcome the limits of the targeted infrastructure, creating a spill-over effect due to the interconnection between national infrastructure. In case this is a first attack, it could be perceived as a wider threat than intended, leading to - again - an accidental escalation, if this is a response to an attack (regarded by the attacker as such), it could be perceived as an horizontal escalation; c) overcome the boundaries of the targeted state, attacking by mistake other neighbouring countries, still due to interconnectivity among infrastructures.

The third factor is the temporary nature of cyber weapons. As was stated in the previous chapter, cyber weapons rely on vulnerabilities in order to propagate and deliver their payload. When a cyber weapon is deployed and found, the vulnerability or the vulnerabilities it exploited could be fixed. The process of rendering a vulnerability fixed in all the systems is very long, but is nevertheless a closing window. The preciousness of a vulnerability is higher when it is constituted by a 0day. Exploiting a 0day vulnerability means having 100% of success that the cyber weapon is going to hit, due to the fact that such vulnerability is unknown. Furthermore there are sophisticated cyber weapons that are custom built for a specific target, for example Stuxnet and Industroyer/CrashOverride that were used respectively to attack the industrial control systems of the nuclear enrichment facility of Natanz and the Ukraine's power grid. Once the secret of these attacks are public, then they become useless. However, if on the one hand IT security operators work on fixing them, on the other hand other malicious attackers that did not have the capability of scripting such a cyber weapon acquire base to work on, and could build variants. Nonetheless, the original cyber weapon loses its purpose after it has been found, whether it is before or after the deployment of its payload. These characteristics therefore define how and when cyber weapons could be deployed.

The fourth factor that influences the exercise of power in cyberspace is given by the geopolitical situation during which it occurs. As it was stated in the first chapter, in period of tension or crisis between states it is to expected an harsher reaction to physical as well as cyber attacks, compared to a situation of peace. In reality, it seems to be the case that during situation of crisis cyber conflicts are used to display power in order to avoid resorting to physical warfare. CNOs, therefore, create a sort of middle layer between "peace" and "war" intended as classic, kinetic, war. This middle layer is used to project power between states due to the fact that an international normative framework is missing and that cyber weapons are generally less costly, both in terms of development that in terms of casualties and physical destruction. For example there are no known examples of major cyber weapons deployed in peacetime, as well as in wartime. The point is that, unless deliberately intended, given all the characteristics of cyber weapons mentioned in the previous chapter, and especially the self-constraining mechanisms mentioned above, it would be pointless to deploy a sophisticated cyber weapon without a political or military objective because it would eventually spur a crisis. Furthermore, during kinetic warfare, major cyber weapons could lose their effectiveness, because a state could cause disruption and destruction through physical means. For example, the cyber attacks performed by Russia during the conflict with Georgia were of secondary importance compared to the physical military actions performed by the two states. Therefore, CNOs acquire greater importance as display of power when they are performed in a situation where both peace and war do not constitute options.

The main problem is to analyse the cyber exchanges between states, because - once again - given the lack of international law applicable to cyberspace and the temporary nature of cyber weapons, the "cyber arsenals" are kept secret as well as the actual usage of cyber weapons. Indeed, what is known both by scholars as well as the general public is a) cyber weapons which payload has already been deployed, namely after the attack and b) cyber operations that were sophisticated enough or attacked targets (both in terms of state as well as physical systems) important enough to deserve media coverage. Of course, this could constitute a problem as far as data gathering is concerned, due to the fact that also media coverage and academic literature on this topic still mainly US centric. Furthermore, it is important to underline that communication is one of the essential parts of the attribution process, and that CNOs take a long time to be analysed by forensic experts.

Studying the retaliation and escalation mechanisms in cyberspace proves a more difficult task compared to kinetic attacks. The reason behind this statement was given in the first chapter, but it is useful to recap the concept. Cyberspace is a domain that is completely different from the physical one as far as the aesthetics are concerned. In the physical world if a country A launches an attack against a country B, it is under everyone's eyes and it is fairly easy to decide which country did strike first. In cyberspace this kind of situation is a bit different, because the risk of inadvertent or accidental escalation is higher. A simple reconnaissance could be judged by country B as an attempt to attack, leading to a reaction whose degree of intensity could vary depending on the perception of the threat. At the same time, country A would not see the reconnaissance it performed as an attack, and it could justify its action by stating that it was a simple gathering of information spurred from the suspicion that country B performed a previous intrusion on its system, since no one would want to be labeled as the first one to attack. Therefore, if we would take all intrusion or suspected intrusions as "attacks" then it would be a dead end, since attempts to intrude state systems are in the order of the hundreds of thousands every day. This could be another reason to justify the case studies that were chosen for this dissertation, namely cyber weapon sophisticated enough to breach state systems and gain international resonance. There is another reason that trumps the difficulties in choosing case studies when the unique origin of information are reports and media articles. That is to say the imperative necessity to push this field of study onwards, without being stuck in the same questions while attacks increase in sophistication and are becoming ever more deadly and more and more states acquire offensive capabilities in cyberspace, worsening the security dilemma. This increase in both quantity and quality calls for academic studies that tailgate every development in the domain, trying to adapt classic concept to this relatively new realm and maybe develop new theories.

Given what has been stated in the previous chapters, retaliation and escalation are expected to be carried out differently compared to kinetic exchanges of the same nature. In the physical realm, an increase of force should be expected, but as underlined in the second chapter, the equivalence principle cannot be applied to the cyber realm, due to the sophistication of the code and all the variables linked to it. In this sense, there is a huge gap between intentions and consequences, and both must be taken into account. In light of the previous chapters, it is expected to have a situation of tension or crisis between two states, the use of cyber weapons as a mean of both self-help and self-constraint that works in two ways, the first as a mean to constrain the use of kinetic weapon, the second one as a mean to constrain the use of the cyber weapon itself, employing cyber weapons useful to acquire the political object but without spilling over in other countries or infrastructures, worsening the situation. Furthermore, the main aim of this chapter is to see whether asymmetry in power plays a role in the development of cyber disputes, where the state that is superior in power generally intended is also superior in the cyber realm but, at the same time, a situation of escalation dominance does not impede the adversary to retaliate in the cyber realm.

The case studies are three, the first two analyse asymmetry, the third symmetry. Every chapter would provide a description of the geopolitical context, the actions carried out, the retaliation and the analysis thereof. Namely, if it was vertical, horizontal, de-escalating, stable on the same level of intensity of if, the state surrendered. The first one describes a powerful state attacking through cyber means a less powerful states, namely the Stuxnet case, involving the US and Iran. The second one concerned the opposite scenario, a low power state against a powerful state, that is to say the Sony Hack, involving North Korea and the US. The third one is a situation of asymmetry and the cases involve the three major powers in the world, US, China and Russia, divided into two dyads US and China, and US and Russia.

United States of America versus Iran – the Stuxnet Case

The first case that is going to be analysed is also one of the most famous cyber events in the past ten years. Namely, it could be considered *the* most famous CNA to ever happen, since it completely changed the use of cyber weapons among state actors: Stuxnet.

The statements above are very bold and must be corroborated by motives. Indeed, Stuxnet could be considered the first malware deployed by a state officially responsible for having caused damage to physical object of another state. The object in question are the centrifuges of the uranium nuclear enrichment facility of Natanz, in Iran. Why and how Stuxnet was deployed, and if and how Iran retaliated are the objects of this paragraph.

To understand why, the analysis must start from its end, that is to say from what spurred all the analyses and let the world know about Stuxnet. However, one machine that was infected by Stuxnet on its road to its main target - the centrifuges, but it was not the ultimate target of the malware - began to continuously reboot (Zetter, 2014). This raised some suspicion and the code extrapolated from the machine was analysed firstly by a small security company named VirusBlokAda, later by two giants of cyber security, namely Kaspersky and Symantec (Collins and McCombie, 2012). Symantec did the most in-depth analysis on and the first thing that it discovered was that the malware was communicating with the outside (Falliere, Murchu, Chien, 2011). Every time it infected a systems it called with encryption two command and control servers one in Denmark and one in Malaysia, masked as football websites (*Ibidem*). This is an interesting part to analyse because it is a method that tries to limit one of the “fears of the unknown” that is to say the risk of a malware going out of control and infecting unwanted machines. Therefore the makers did not have complete control over the malware. Indeed a self propagating malware could not be blocked completely but could be limited to some extent. For example, if Stuxnet encountered a machine that did not possess particular characteristics, for example the fact running a

Siemens WinCC[13] or SIMATIC Step7[14] softwares it shut itself down (*Ibidem*). This could be considered another mean to control the propagation of the malware. Symantec analysts used a common practice during malware forensics, that is to say *sinkholing* (Zetter, 2014). This method, that was mentioned in the previous chapter, consists in placing a DNS server in the middle of the communications system of the malware with the command and control center, in order to receive all malicious traffic to be analysed. What Symantec discovered was that out of 38 000 machines that were analysed, more than 20000 where in Iran (then India, Pakistan and Indonesia) (Chen, 2010).

At this point one must ask itself why these were the countries that were mostly targeted, and in order to understand the reasons behind the attack the whole geopolitical situation must be unfolded.

The Iranian nuclear programme was a source of unrest and tension for the whole Middle East, including also Israel and, consequently, it involved also the US and other western states. Indeed, during 2006 UN voted in favour of sanctions against Iran, because of its nuclear programme (Gootman, 2006) and, furthermore, talks about an airstrike against the plant began to emerge, which the IAEA chief warned against (Jahn, 2007). Furthermore, this picture could be expanded also through India and Pakistan (two of the top three main targets of Stuxnet). Iran, India and Pakistan were planning on building a natural gas pipeline that should have gone from southern Iran through Pakistan into India, which would have also fostered peace and cooperation between the latter two (Sahay and Roshandel, 2010). The US was against this pipeline and put pressure on India that withdrew from the project but joined it again in 2010 (Verma, 2007).

The deterioration of US-Iran relationship begins in 1979, with the overthrowing of the Shah Mohamed Reza Pahlavi substituted by the Ayatollah Khomeini, which shifted the Iran from being a pro-American country to an anti-American one (Peterson, 2001; Bruno, 2010). This deterioration could be seen also by the backing of the Iraqi soldiers by the US during the Iran-Iraq conflict, which, in turn led to the sponsorship of Hezbollah by Iran. However, the Iranian nuclear program began with the Shah, which wanted nuclear power for both

[13] Specific software designed to work as interface for Siemens SCADA (*Supervisory Control and Data Acquisition*) systems, which are software that allow to monitor and control the actions of sophisticated machinery.

[14] Specific software designed to manage PLCs (*Programmable Logic Controller*), which are pieces of hardware that give instructions on how to function to sophisticated machinery, such as rotors, valves, or centrifuges.

civilian as well military uses (Peterson, 2001; Bruno, 2010). At the time, Iran was a US friendly country, and these plans did not constitute a problem for the American government and the Shah was able to strike a deal with Eisenhower under the Atoms for Peace programme's auspices (Sinha and Beachy, 2014). Indeed, Iran signed the Non-Proliferation Treaty in 1968, instituted its atomic energy organisation and US, Germany and France sold Iran components for the two nuclear reactors for the Bushehr facility (Bruno, 2010). After the uprising that led Khomeini to power, the western countries withdrew their support, and Bushehr was bombed during the Iraq-Iran conflict. Khomeini was initially against nuclear power but the deployment of chemical weapons by Saddam Hussein on Iranian people and soldiers plus the rumours that he was going to acquire nuclear power changed his stance (Bruno, 2010; Sinha and Beachy, 2014). The revived Iranian nuclear programme's plans included also an uranium enrichment facility. With the absence of western nuclear power willing to help Iran, between 1985 and 1987 the Ayatollah turned to Pakistan, precisely to Abdul Qadeer Khan, who responsible for acquiring illicitly all the plans and materials for Pakistan's nuclear programme (Bruno, 2010; Squassoni, 2005). Therefore, through Khan, Iran was able to acquire plans and prototypes for building a nuclear enrichment facility and all the instruction to weaponise enriched uranium. The main component of a nuclear enrichment facility are cascades, that are basically groups of centrifuges. These centrifuges are metal tubes that spin at huge speeds, the likes of more than 100 000 spins per minute and in, the form of cascades, separates the two isotopes of uranium hexafluoride gas into U-235, lighter ones, and U-238, heavier ones (Langner, 2013). What it is necessary to have enriched uranium are U-235 isotopes, while the others are discarded. The uranium hexafluoride gas for this secret enrichment programme was sold to Iran by China, but by 1994 Iran managed to have only one centrifuge working at full speed. At the same time, Iran asked Russia for help in re-constructing the facility of Bushehr, symbol of the public nuclear programme (Bruno, 2010; Sinha and Beachy, 2014). At this point, not knowing about the secret enrichment programme, the Clinton administration put pressure on Russia in order to discard the possibility to sell Iran also the know-how and the technology to build an enrichment facility (Einhorn and Samore, 2002). The first successful enrichment experiment was conducted at the end of the 90's, at the secret small facility of Kalaye. After a decade-long effort, the Iranian atomic energy organisation pushed for mass production of enriched uranium, creating the bigger facility of Natanz. Indeed, Bushehr's nuclear reactor would have been Iran's primary reactor, that would have been fuelled by the uranium enriched at the nuclear enrichment facility in Natanz. In 2003 the existence of Natanz went public and the IAEA stormed into it, declaring that the Iranian nuclear programme was more advanced than expected, the likes of two to three years from a nuclear arsenal (Kerr, 2003; Squassoni 2005). Furthermore, the IAEA inspectors found traces of uranium enriched at 70% (Squassoni, 2005), and it is useful to specify that a 90% enriched uranium is classified as weapon-grade uranium. Despite the involvement of the EU3 - United Kingdom, France and Germany - and the IAEA investigations conducted also in other secret facilities working on other stages of the nuclear programme that were discovered in the meantime, Iran pushed forward its plan to enrich uranium and become a nuclear power. In 2005, the IAEA was able to obtain secret documents with the help of the CIA, which showed sketches and plans for missiles and nuclear warheads, plus a footage of a missile test at high altitudes, but Iranian officials accused the IAEA of using forged documents in order to justify an Israel-US aerial bombing on Natanz (Peterson, 2011). Needless to say, the political tension was immense. This tension was aggravated by the fact that in 2005 Mahmoud Ahmadinejad became the president of Iran, who was deeply against the IAEA investigations and halt of a national nuclear programme, which he saw as an indisputable right and a firm point agains the inference of Israel and the US in the Middle East. Only two months after being in power, Ahmadinejad revoked the suspension of activities promised to he EU3, and also the IAEA seals on the facilities, stepping on the gas of the nuclear programme at full speed. The tension in all the countries in the Middle East was very high, but still Iran was having problems with the enrichment process, due to the fact that, starting already in 2003, the CIA was infiltrating tampered pumps and electric power supplies both in front companies as well as in the so-called “Khan network” - referring to the aforementioned Abdul Qadeer Khan - both suppliers of components that Iran was using to build the Natanz facility (Maher, 2012). A batch of defective parts bought from a Turkish supplier made 50 centrifuges explode during a test in Natanz in 2006 (*Ibidem*). However, it wasn’t sufficient to halt the Iranian nuclear programme. Furthermore, the UN voted in favour of toughening the sanctions against Iran in 2007, but even this could not stop Ahmadinejad for pursuing his nuclear dreams fuelled by the fact that the Natanz facility passed from having 1 400 centrifuges at the beginning of 2007 to 3 000 at the end of the same year (Crail, 2008). This amount of centrifuges was enough to enrich weapon grade uranium in less than a year, but the president had plans to double the number of centrifuges by 2008 (*Ibidem*). Iran as a nuclear power was a matter of months, sanctions and sabotage proved pointless and the only option to halt the programme seemed to be destroying the Natanz facility, but this would have meant a certain war.

As was stated at the beginning of the paragraph, the construction of Stuxnet was built around the highest minimisation of risks possible. The following part analyses how Stuxnet was able to reach it target unnoticed, at least until 2010. The start of the infection itself was characterised by a secure method reducing the risk of being caught, for this reason there were no phishing attempts or external attacks (that were still possible, even though the majority of Natanz facility is air gapped), but with an infected USB, or *BadUsb*, insertion. Right away the first 0day comes into play in the form of four infected and hidden .lnk files (Matrosov et al, 2010; Falliere, Murchu, Chien, 2011). A .lnk file essentially acts as a direct link (shortened "lnk") with an executable file which could be found in another location, for this reason opening a .lnk triggers the same consequences as opening the original one. When an external drive is inserted in a Windows-running machine, the Microsoft operating system scans the files stored in the drive. In this specific case, the .lnk thanks to a 0day exploit was modified in such a way that when Windows scanned the .lnk files it triggered the malware that downloaded itself in the machine (Matrosov et al, 2010; Falliere, Murchu, Chien, 2011). Again, the minimisation of the risk was given not only by a direct insertion, but also avoiding the use of triggers exploiting the autorun function of Windows, that now is disabled for most external drives. The sophistication of Stuxnet could be inferred also by the fact that attacked only machines running windows at 32bit and, furthermore, to avoid updated machines, the four .lnk files were in reality four versions of the same 0day exploit that were coded to run on all versions of Windows, starting from Windows 2000 until Windows 7 and Windows Server 2008 r2 (Matrosov et al, 2010; Falliere, Murchu, Chien, 2011). The kind of Windows platform was the only obstacle that the developers of Stuxnet could find in this phase, due to the fact that the exploit was a 0day and therefore no patch was known because the exploit yet to be discovered. Another peculiarity of Stuxnet was that it was able to download and install itself on the machines without triggering any notification about it. Usually, when a software is installed on a Windows-based machine, the operating systems warns the user that a software with no certificate - or an untrusted one - is attempting an installation. Again, Stuxnet was all about the minimisation of risks, and indeed it was signed with a genuine digital certificate. Forged certificates are usually spotted by windows, since they are signed differently and continuously after short periods of time. This means that is very likely that the authors behind Stuxnet obtained a genuine certificate through an intrusion in a legitimate company, in this case RealTek Semiconductor Corps. (Falliere, Murchu, Chien, 2011). Windows recognised Stuxnet as having a genuine signature and therefore allowed the malware driver modules to install themselves effortlessly and without problems. After it installed itself, Stuxnet began escalating privileges - from user to system administrator - on the infected machines exploiting another 0day vulnerability, this time found in a keyboard layout file of windows (*Ibidem*). A deeper analysis showed that Stuxnet searched for specific machines that ran the aforementioned Siemens SIMATIC Step7 software, used for managing Programmable Logic Controllers. Attack on PLCs were unknown at the time because they are very uncommon. The only example of an attack aimed at physical destruction of machinery was a test conducted by Idaho National Laboratories, called *Aurora Generator Test* in order to show how tampering with industrial control systems could result in the physical damage of electric grid components (Wang, Fang, Dai, 2010). PLCs are used to manage industrial control systems, basically they are the combination of hardware and software that regulates the functioning of a specific machinery, such as centrifuges, rotors, turbines and the likes.

Usually, one 0day is enough to make a cyber weapon superior to all the other that do not employ one, due to the increase in the success rate. Two 0days are definitely more than enough to ensure that a cyber weapon reaches its target. The analysts at Symantec discovered two more 0days. Four 0days is very high number for any cyber weapons, and it is very unusual to find a malware with such an amount of 0days. The fact is that a high amount of money is needed to buy 0days, which are sold on the black market for around 100,000 € on average (Greenberg, 2012). The basic motivation behind this overload of 0days that could be inferred was just one: the attacker had access to a huge amount of resources and wanted to have complete certainty that the cyber weapon reached its target. To sum up: the four main 0days that Stuxnet used were:, the .lnk exploit in the USB, an exploit in the keyboard file of windows and in the windows task scheduler, which permitted the escalation of privileges, an exploit in the print-spooler which helped the malware propagate among machines that shared the same printer. Furthermore, Symantec analysts found four additional propagation methods (Falliere, Murchu, Chien, 2011). One of these infected the database that the programmers of the aforementioned Step7 software. By infecting the database, which is shared among every programmer working with Step7, every programmer's machine became infected with Stuxnet, in order to raise the success of finding a machine using a Step7 PLC (*Ibidem*). Furthermore, Stuxnet did communicate with its command and control servers, both to share information about its location as well as to be updated remotely. Not only, stuxnet used a peer-to-peer network in order to make infected machines communicate among each other when connected in the same local area network, so that if one ran an updated version of the malware, all the other machines downloaded the update (*Ibidem*). Exploiting the communication among machines, Stuxnet used also another vulnerability to propagate, one that exploited network shares, that is to say those files shared by machines connected to the same network. It is interesting to note that it was the same vulnerability exploited by the malware Conficker in 2008 (Markoff, 2009), which apparently was not patched although the patched was issued by Microsoft. The sophistication of this cyber weapon is flabbergasting. Natanz is an air gapped facility, it means that it is not connected to the internet. Indeed, none of the exploits used to take advantage of vulnerabilities, propagate in the systems and deliver the payload relied on an internet connection. This is very important because it separates Stuxnet from the majority of all other malwares in history. Albeit how a so-called *BadUsb*, namely a corrupted USB flash drive was inserted in Natanz's facility remains a mystery - although one could infer that intelligence methods were used, such as social engineering or through a spy inside the facility -, it remains the fact that it jumped from machine to machine, infecting those who possessed those characteristics that lead to the Siemens PLC that managed the centrifuges. These "jumps" were not random, but they were perfectly calibrated. However, the two main exploits of Stuxnet were the first one, namely the one exploiting .lnk files, in order to allow Stuxnet to enter Natanz's facility, and the one that infected the Step7 database. The latter was crucial because the operators of PLCs use particular computers that are not connected to the internet but are connected to the shared database. Eventually, a programmer at Natanz inadvertently downloaded Stuxnet in his machine and then infected the Siemens PLC that was managing the centrifuges. As stated at the beginning of this chapter, PLCs are used to manage various type of machinery, ranging from turbines to oil pipes. The interesting thing is that Stuxnet targeted a specific PLC, the Siemens S7-417, with a specific configuration that is to say the one used to manage centrifuges for uranium enrichment purposes (*Ibidem*). So, it is true that Stuxnet infected every machine due to the compromising of the Step 7 database but, at the same time, if the PLC was not configured to operate centrifuge it shut itself down. This made Stuxnet what we could call a precision cyber weapon, the first of its kind. To increase the precision of the cyber weapon, it is worth mentioning that Stuxnet also attacked specific frequency converter with which it tampered in the last stage of the payload delivery. The frequency converter is what makes the rotor of the centrifuge function (Albright et al, 2010). Furthermore, it also infected the *Human Machine Interface,* or HMI, that serves to human operators to see all the process that the machine is performing. By tampering with the HMI, Stuxnet was able to modify the value of the frequencies while displaying that everything was working properly to the operators.

According to the IAEA inspections, during 2009 the enriched gas amount produced by the facility at Natanz dropped rapidly, and the Iranian technicians disconnected eleven out of eighteen cascades that were operating at the facility. Indeed, the last piece of the puzzle, the frequency converters, were the key in understanding Stuxnet. After being inserted into the systems, after all the "road" and all the jumps from machine to machine, after meeting all the requirements embedded into the code of the malware, after infecting the PLC and the HMI, Stuxnet arrived at the frequency converters, that are pieces of hardware that manage the spins of the centrifuges. The normal frequency of a centrifuge is 1,064 Hz, Stuxnet brought this frequency to 1,410 Hz for around fifteen minutes, that is the highest frequency that a centrifuge can tolerate before breaking down (Zetter, 2014). According to the IAEA safeguards report, at least 1000 centrifuges were taken down following the Stuxnet attack (Albright et al, 2010), prompting a halt in all the enrichment process at Natanz and, as a consequence, halting temporarily the Iranian nuclear programme.

The development of Stuxnet began allegedly around 2006. In that period both the US and Israel were confronted with the idea of performing an air strike against the facility at Natanz, that would have taken the Iranian nuclear program back of about three years (Zetter, 2014). However, as it was stated before, the political tension between Iran and the West was so high that an air strike - that was the only possible option among the kinetic weapons, due to the fact that Natanz is built underground - was not an option. The main architects behind the cyber options were essentially two, the US Strategic Commands's general James Cartwright and former NSA director Keith Alexander which brought the idea to president George W. Bush, which approved it in 2007, and was continued by president Barack Obama under the codename "Olympic Games" (Farwell and Rohozinski, 2012). The NSA developed the Stuxnet code at the beginning later combined with Israeli's Defence Force Unit 8200, the NSA counterpart in Israel (Zetter, 2014). Isreal's presence is crucial in the development of Stuxnet because the cyber weapon was tested at the Dimeona facility in the Israeli desert, were the Israeli covert nuclear programme developed its nuclear weapons (Broad, Markoff and Sanger, 2011). This has to be connected to the fact that the Oak Ridge Laboratory in the US was able to obtain centrifuges of P-1 kind, the one which Iran modelled its centrifuges upon (*Ibidem*). In this particular scenario, the choice of a cyber weapon was a logic and a smart one, for a number of reasons. First, the planning of Stuxnet was so thorough that the probability of being discovered was very thin, and therefore the Iranians would have blamed an incident or an unknown malfunction as the source of the problems, avoiding the risk of an international crisis. Furthermore, even if the cyber weapon was discovered, which it was, it would have increased paranoia and fear among the Iranian establishment. Indeed Natanz was shut down after the discovery of Stuxnet in order to analyse and clean all the systems that could harbour Stuxnet (Zetter, 2014). Second, compared to an air strike it is silent and does not expose pilots to potential risks. Third, compared to an air strike it does not provoke casualties, indeed attacking the cascades poses risk for neither nuclear explosions nor uranium poisoning, since the quantities contained in the centrifuges are not lethal. Last, due to the sophisticated architecture of Stuxnet, there was the possibility that it could have spread to other unknown - secret - infrastructures, working on the uranium enrichment programme.

Iran, at least at that time, did not have cyber capabilities strong enough to respond with a malware with the same power. Allegedly instead, Iran developed Shamoon, an espionage malware found in Saudi Arabian computers to monitor and erase critical files on about 30,000 computers at Saudi Aramco, the world’s largest oil company, disabling them (Leyden, 2012; Bronk and Tikk-Ringas, 2013). Saudi Arabia is one of the major US partners in the Middle East, and that could be considered a test-bed for retaliation. Again, the US officials claimed that in retaliation for Stuxnet, in 2013, Iran hacked US Navy computers (Barnes and Gorman, 2013) as well as performed CNEs against US online banking sites, among which there were JP Morgan, Bank of America, Wells Fargo and PNC Financial Service group (Capaccio, 2013). In May 2014 iSight Partners, that is a security firm based in Dallas, issued a report stating that Iran has been performing a cyber espionage campaign dubbed “Newscaster” for the last three years that targeted military contractors, members of Congress, diplomats, lobbyists and journalists (Perlroth, 2014). What is interesting is that, among the most targeted persons, there was John R. Bolton, an American diplomat. Regarding being a target he stated: “I think the Iranians were after me to get all the secrets that the Obama administration has imparted to me about the Iranian nuclear program [as well as being] the most anti-Iranian regime in Washington” (Sanger, 2014). This has important implications as far as escalation is concerned. It could be argued that the objective of the Iranian cyber attacks that started soon after the discovery of Stuxnet was indeed retaliation. It is important to underline that, according to the Tallinn Manual on the International Law Applicable to Cyber Warfare "acts that kill or injure persons or destroy or damage objects are unambiguously uses of force." (Schmitt, 2013). As mentioned in the second chapter of this dissertation, if the Tallinn Manual was a policymaking document, Stuxnet would have been considered an attack breaking the Geneva Convention, capable of spurring a conventional, armed retaliation from Iran against the US. This factor has the potential to transform the cyber escalation into a classic one.

We could conclude that in this case, the geopolitical context was a fundamental variable. Without an historic hostility between the two country and a situation characterised by a level of tension so high to exclude an air strike that would have spurred a war with absolute certainty, employing a cyber weapon was the best option by the US. The escalation dominant position held by the US did not stop Iran from retaliating, but the retaliation was carried out in a vertical fashion but in a downward movement. That is to say that, due to the fact that Iran does not possess the same cyber capabilities as the US it responded to a disruptive and destructive CNA with Shamoon, which is a tool for CNEs but it doesn't stop at espionage because it also disrupts the master boot records of the system it attacks, rendering the boot process impossible. In this sense, the movement of the retaliation was also minimally vertical as it targeted US banking systems, military contractors and political representatives but also horizontal, because it did not target a critical infrastructure in the US, as Stuxnet did, but a critical infrastructure of one of the main US allies' major assets in the region, namely Saudi Aramco.

North Korea versus the United States of America – the Sony Hack Case

The political relations between North Korea and the United States could be considered one of the remaining legacies of the Cold War. Albeit the conflict ended almost thirty years ago, this already difficult relation got worsened when president George W. Bush during his 2002 State of the Union speech mentioned North Korea, together with Iran and Iraq, as part of what he called "the axis of evil" (Cha, 2002). It is simple to understand by these first few sentences the nature of the relation between the two states. It is important to underline that the "sentiment" is mutual, for the US, North Korea is a country that does not respect human rights and poses a potential threat to international security due to its nuclear programme, on the other hand, the North Korean perception of the US is that of a menace to its own existence. This position stems from the heavy bombing that North Korea suffered during the Korean War. It was during this war that the relations between the two countries was shaped. Indeed, after World War II, the Korean Peninsula was divided along the infamous 38th parallel north in two separate spheres of influence: the north was occupied by the USSR, while the south by the United States of America. After three years, given the impossibility to reconcile the two halves, the southern part founded the Republic of Korea, or ROK, while the north, just a month later, founded the Democratic People's Republic of Korea, or DPRK. The only common ground that these the country shared was the hostility against each other, both wanting the unification of the peninsula only under their respective sovereignty (Wertz and Gannon, 2015). Two years later, in 1950, a DPRK invasion against the south triggered a war, which spurred the intervention of UN forces aiding the ROK, and the Chinese's People Volunteer Army helping DPRK's troops. In the 1953 an armistice was signed which fostered a US presence on ROK's soil, by deploying an armed contingent and tactic nuclear weapons. Furthermore, the Korean War pushed the US into imposing an embargo on DPRK exports, in this way the economic capability of the northern part of the peninsula became severely limited. From half of the Sixties until half of the Seventies it took place what it was called the "Second Korean War" a serious of skirmishes and provocation started by DPRK against US aircrafts, ships, and officers, and the Eighties and the Nineties were characterised by the fear of a North Korean nuclear programme (Armstrong, 2004, Wertz and Gannon, 2015).Pressured by the USSR, in 1985 DPRK signed the Non Proliferation Treaty and it did also sign a Joint Declaration for the Denuclearisation of the (Armstrong, 2004, Wertz and Gannon, 2015). However, the signing of the Declaration implied also that the IAEA would have conducted inspection in both Koreas, and found inconsistencies in North Korean reports, suspecting that DPRK had already enriched enough plutonium to build a nuclear weapon and asked for special inspections (Armstrong, 2004). As an answer, in 1993 DPRK wanted out from the NPT, but in order to halt the North Korean nuclear programme while keeping the country a signatory of the NPT, the US signed with DPRK the Agreed Framework, through which the US would have provided North Korea with heavy fuel oil and with help for the construction of two light water nuclear reactors in exchange of a complete halt of DPRK nuclear programme (Kimball and Crail, 2012). The bilateral agreement was stalled many times through the nineties, for example when the North Koreans shot down a US helicopter in 1994. Despite the agreement, DPRK continued its long range missile programme and ballistic missile programme. In 1999 a North Korean covert nuclear programme was discovered, and led to the US proposal of new talks, coordinated by Japan and the Republic of Korea. The interesting thing is that the Secretary of Defence at the time, William Perry suggested that if the talks would have resulted in failure it would have meant the failure of diplomatic measures that would have led to a military response (Moon, Okogi, Reiss, 2000). In order to halt the missile programme, the US agreed on a lifting, albeit partial, of the sanctions. The US accommodating approach that characterised the Nineties during the Clinton administration changed completely with the advent of president George W. Bush. As stated at the beginning, in 2002, during his first State of the Union speech he explicitly mentioned North Korea as part of the "axis of evil". The same year, officials of the Bush administration went to visit Pyongyang, and during the meeting North Korean officials admitted to have a uranium enrichment programme (Armstrong, 2004). Then, the US stopped the supply of heavy fuel oil as North Korea violated the Agreed Framework, and as a result, DPRK nullified the Agreed Framework, withdrew from the NPT and re-started officially its nuclear programme (Kimball and Crail, 2012). This spurred immediately the so-called six party talks, which began in 2003 among North Korea, US, China, Russia, South Korea and Japan, aimed at dismantling the North Korean nuclear programme, which DPRK agreed to in exchange of energy and food assistance by signing a Joint Statement (*Ibidem*). At the same time, the US began an operation aimed at cutting all illegal financial provisions, by freezing assets in a bank in Macau with the aid of the government, and pressuring several international banks to avoid having ties with DPRK. These actions were met with resentment by the North Korean government which, in 2006, performed a first ballistic missile launch, and also a nuclear test. These actions resulted in the UN placing other sanctions upon North Korea and another round of six party talks. Briefly, several six party talks rounds took place in the next ten years, and DPRK persisted in testing its missiles and nuclear weapons, performing two nuclear tests and twenty missile launches in 2016 only (Armstrong, 2017). To sum up, it is clear that the relations between North Korea and the US were never amicable nor based on cooperation. The US stance with the Obama administration was one of "strategic patience" while the UN kept on posing sanctions on the Korean country. At the same time, however, the US did not stop its double-track strategy, where on the one hand we find diplomatic measures and on the other hand covert operation in order to monitor as much as possible North Korean nuclear and cyber activities. And that is what emerged from the Sony Hack case.

The tension between North Korea and the US had a peak in 2014 and did not have anything to do with missiles or nuclear tests. In November 2014, around 7000 employees of Sony Entertainment, the American branch of the famous Japanese conglomerate, found on their desktop photoshopped pictures of a beheaded Micheal Lynton, chief executive at Sony Entertainment at the time. The studio promptly puts offline all the employees' PCs and the local network, with the upper management resorting to old Blackberries to communicate and the employees receiving phone calls, faxes and paper checks in order to be paid (Pagliery, 2014; Haggard and Lindsay, 2015). Basically, Sony Entertainment suddenly found itself in the Middle Ages. However, initially the attack was belittled by Sony Entertainment, which, in only three weeks found itself in the middle of one of the major crises in the history of the company. The main problem surfaced rapidly, that is to say that the defacement of desktops was not the only attack, but terabytes of data were stolen from the company's servers and, furthermore, half of the data on personal computers and servers were erased (Robb, 2014, Haggard and Lindsay, 2015). Not only, before deleting al those data the attackers overwrote the data seven times, rendering all efforts to recover them pointless. In the three weeks after the attack, the attackers, a group of hackers calling themselves the Guardians of Peace (GOP), published the stolen data in several tranches. The content of the data ranged from personal information about the employees, such as social security numbers, private exchanges between employees, as well as movie that were not out in theatres yet. Indeed, the first batch of data that was dumped were three movies that were available on torrenting sites (Cieply and Barnes, 2014).

After this dump, the FBI was involved in the case. This is the major point that has to be underlined and understood in all the Sony Hack case. Furthermore, it justifies the choice of this case studies in order to better understand state behaviour and asymmetries during disputes in cyberspace. Sony Entertainment is a private company and cyber attacks against private companies are not uncommon, especially espionage activities. However, this time there was the suspect that behind this attack there was a state actor, namely North Korea. This suspect was founded on the fact that the Guardians of Peace asked for the withdrawal of a particular movie *The Interview* from the roster of the movies to be published in December. *The Interview* is a movie directed by Evan Goldberg and Seth Rogen, which narrates the tale of a gossip journalist and his producer, where the latter manages to organise an interview with Kim Jong-Un, DPRK's leader. The two are then appointed by the CIA to kill the leader, and they accept and fulfil the mission (Bond, 2014). In June 2014 a North Korean spokesperson, according to the KCNA, Korea Central News Agency, stated that if Sony Entertainment published the movie, then North Korea would have responded with stern and merciless retaliation (Inkster, 2015). Furthermore, North Korea's ambassador at the UN openly criticised the movie and defined it an act of war, and a mean to sponsor terrorism (Beaumont-Thomas, 2014). After every batch of data published, the Guardians of Peace declared that they would stop if Sony Entertainment withdrew the movie from theatres. Sony Entertainment did not give in, and eventually the hackers menaced physical retribution against the people that would have gone to see the movie, citing the events of 9/11 (Haggard and Lindsay, 2015; Inkster, 2015). After such a threat, 80% of theatres refused to screen the movie, and Sony announced the withdrawal of the film.

Such a posture were met with disdain by president Obama, that put in charge the FBI of the Sony Hack case. This was due for the two aforementioned reasons: first, an American company gave in to the threats of an hostile state actor and two, menacing events comparable to 9/11 made this a case of national security. To sum up, a government attack through cyber means a private company, not even a critical infrastructure, but a) the geopolitical implications and b) the threats posed by the attack spurred the reaction of the government within which borders lies the private company. This was a unique event that had other very interesting implications.

The day that the FBI took charge of the investigation, president Barack Obama declared that the approach followed by Sony Entertainment was erroneous, first because surrender to a threat of a foreign government was not "what America is about" (Laughland and Rushe, 2014) and second - being president Obama a supporter of public-private partnership - because Sony Entertainment should have communicated with the government about the problem sooner. Here lies one of the main obstacle in increasing a culture of cyber security and, effectively, increasing the whole cyber security level of a nation. Private companies are reticent in communicating when they are victims of CNOs for the fear that the event could go public, causing consequences, for example, for the reputation of the company, which could lose the trust of the clients and that could translate in monetary losses. Furthermore, Obama accused directly the North Korean government for being the culprit behind the attack. As expected, DPRK's officials replied that the accusation were unfounded and that North Korea had nothing to do with it, instead they added that it was the work of a group of sympathisers that supported the accusation against the US (McCurry, Carroll, 2014). Due to the difficulties in tracing the source and attributing CNOs and the lack of an international framework that allows to pursue attackers, an official statement along the lines of "my country has noting to do with it" it is pretty common after accusations of conducting hostile CNOs. The main point here is that presidente Barack Obama, resting on the evidences provided by American intelligence agencies, was absolutely sure that the source of the attack against Sony Entertainment was DPRK.

The evidence collected by the FBI amounted to the facts that the attackers used intrusion tools previously used by known activities linked to North Korea, in terms of code, activities, encryption and data deletion (Laughland, 2015). Furthermore, the attackers were sloppy in hiding their IP addresses; indeed, the IP addresses found in the piece of malware used to delete data lined up with IP addresses of North Korean infrastructures (Greenberg, 2015). This is one of the most interesting flaws in the attack performed by DPRK. One rule of thumb in cyber security is that the more a country is depending on the IT infrastructure the more is vulnerable. That is true, because defence must encompass more infrastructures and the consequences of an attack able to bypass those defences will be more extended, compared to a nation with little IT penetration. For a country like North Korea another rule could be applied. Given the limited IT penetration in the country, and given the fact that only a national intranet exists, it should be more difficult for North Korean attackers to conceal their attacks, for example with the use of proxies. Namely, if the attacks are not carried out outside the country, the original DPRK IP could be easier to obtain. Furthermore, a counterfactual argument could be applied. Given how closed the North Korean intranet is, the possibility that a third party used a North Korean IP as proxy - in order to perform a false flag attack - is very little. Another piece of evidence brought by the FBI was that the tools used for the attack against Sony Entertainment were used also in a CNO in March 2014 against South Korean banks, and the attack was perpetrated by North Korea (*Ibidem*).

If the evidence brought forward until know could be seen as circumstantial to the most skeptic analysts, another proof about the fact that North Korea was indeed behind the attack against Sony Entertainment was presented. Namely, the US government was so sure about attributing the attack to North Korea because the NSA pre-emptively intruded North Korean systems (Sanger, Fackler, 2015). In this way they were able to collect evidence and have absolute certainty about the DPRK responsibilities.

After this unprecedented sureness in the attribution process, at the beginning of January 2015, Barack Obama signed an Executive Order which places sanctions on North Korea adducing explicitly motivations linked to the Sony Entertainment Hack case (Korte, Jackson, 2015). Furthermore, and this is also very important for this case study, after the public statement indicating North Korea as the culprit of the attack and after a statement by president Obama about "responding proportionally" to the CNO against Sony Entertainment, DPRK suffered a widespread outage concerning its intranet network (Strohm, 2015; Nakashima, 2015), that could be considered a sort of retaliation in-kind by the United States in response to the Sony Hack. Indeed, North Korean government's spokesperson accused publicly - through state media - president Barack Obama of "not knowing shame" and of "running wild like a monkey" because he "disturbed the internet operations of major media of DPRK" (Kim, 2014). Though these means, the US signalled to North Korea their superiority in political and military terms.

As far as vulnerability, propagation and payload of the North Korean attack, the situation is much more simpler compared to the huge sophistication of Stuxnet. The success of the Sony Hack was due to a combination of factors that could be summed up only by underlining the poor security culture of Sony Entertainment. The DPRK CNE allegedly started with a phishing or spear phishing attack against one of the many authorised users at Sony Entertainment, which could be easily found on linked. Once inside the system as an authorised user, the North Korean attackers could have performed a chain of password resets in order to escalate privileges and obtain system administrator privileges. Otherwise, they could have directed their spear phishing attack against a system administrator from the beginning. As system administrators, they had access to the so-called *active directory* that is a service offered by Microsoft Windows Server which constitutes the core of a whole organisation (Serapiglia, 2016). Having total access to the active directory means having access to all the files stored in the IT infrastructure of a company. From here they downloaded all the files and then deleted them.

The conclusions that could be inferred from the analysis of the Sony Hack are very interesting. First of all, a state actor attacked a private company but, perceiving it as a threat for national security, the country that hosted the private company decided to respond at state level. This constitutes a one of a kind example as fare as cyber disputes are concerned because it also constitutes a precedent and, therefore a deterrent. The US communicated clearly that even if an state or state-sponsored attack targets a private institution but the attack is perceived as a national threat it will respond. Of course, it could be also inferred that this type of response, namely public attribution (political), economic sanctions (economic), and the outage of the intranet networks (military) would be more likely if the attacker finds itself in a position of asymmetry with the US, namely possess a lower level of power.

We could see how, even if this was not an act of war, president Obama retaliated to CNE with "adequate means, whether diplomatic, economic or military measures" as if it was an act of war, following 2011 declaration (Forman, Barnes, 2011). The retaliation in this case, due to the asymmetry in power was vertical both as far as the intensity of the attack as well as in the importance of the infrastructure targeted. To an attack carried out through a CNE that caused the leaking of documents against a private infrastructure, the US retaliated targeting the economy of North Korea, through sanctions, namely spilling over in the physical world, as well as a disruptive CNA against one of DPRK's main critical infrastructures.

The Sony Hack, despite being unsophisticated in nature, was important mainly for the response and the discoveries that it brought forward. The most important being the fact that the NSA through the TAO was already inside the systems of North Korea, following the strategy of active defence that is described in the next section. Furthermore, the Sony Hack underlined how important a strong culture of cyber security is and what are the risks of its absence.

United States of America versus People's Republic China and United States of America versus Russian Federation

The last cases to be taken into consideration is when there is symmetry in power between both parties involved in cyber disputes. What could be considered the most interesting case, namely the exchange of offensive CNOs between major powers is the most difficult scenario to analyse. This stems from a problem that was described in the introduction of the chapter, that is to say a literature that today remains US-centric. Furthermore, the problem in analysing cyber attacks between dyads concerning what could be considered the three major cyber powers in the world today, the United States of America, China and Russia, also derives from their respective school of thought and approach to disputes in cyberspace, which are here described.

United States of America

The US could be considered pioneers in exploiting the cyber domain for military purposes. Indeed, they are the country where the domain itself was born, in 1969, at first as a resilient military infrastructure to resist a potential nuclear attack from the USSR during the Cold War, named ARPANET, and then as a civilian commodity during the Nineties that quickly spread across the globe also thanks to the first commercial Internet Service Providers born in the Eighties (Hafner, Lyon, 1998).

Military operations in cyberspace are an essential part of the broader US military strategy. The US military strategy in cyberspace rests on five different pillars. The first one is the recognition of cyberspace as the fifth domain of warfare, where the US could operate freely meaning defending their digital resources, and the physical ones linked to the virtual domain, and be prepared to attack if a particular occasion calls for the need of offensive cyber operations (Garamone, 2010). The recognition of a new domain of warfare requires a doctrine, strategies and also the institution of a Cyber Command. The fifth domain of warfare was enunciated for the first time in 1995, and it was called information domain (Metz et al., 2006). This - now relatively - new domain was of essential importance in the middle of the nineties for two main reasons. The first concerned the technological advance given by the new data centres that were used as a tool to better the communications between the command and control centres with the operational and tactical troops in land, sea and air (*Ibidem*). The second reason revolved around the increasing dependency of the entire country on the IT infrastructure: the entire governmental, financial, transportation, energetic apparatuses, to say a few, and also the civilian communication system were now almost entirely dependent on cyberspace. The beginning of a "modern" approach to an American cyber security started in the early 2000s. In 2003 the first *National Strategy to Secure Cyberspace,* and in 2006 the *Military Strategy for Cyberspace Operations* (Kramer, Starr and Wentz, 2009). The publication of these strategie were timely to say the least, because 2007 was the year of the Russian CNOs against Estonia, which proved - together with the CNEs operation against the US originating in China, that will be mentioned later - that other state actors were actively using the cyber domain as a mean to project their power postures. Furthermore, in 2008 the Pentagon suffered a major cyber breach, whose perpetrator still remain unknown, dubbed Operation Buckshot Yankee. Due to the insertion of an infected USB key in a military laptop, a worm called agent.btz managed to spread in the US Central Command and was able to create backdoors that allowed a foreign government to steal US military secrets (Lynn, 2010; Healey, 2012). This breach produced two consequences: the first one was the ban of all USB flash drives from military operations, the second one was the realisation that cybersecurity wasn't a matter of the IT department anymore, that classified information were at risk and this risk could have heavy consequences on military operations and national security. As a natural follow-up to this second consequence, in 2009 the US instituted the United States Cyber Command created within the National Security Agency, the NSA (Lynn, 2010). The Cyber Command was an important strategic innovation, because it collected for the first time the planning, the management, and the responsibilities for the actions in the cyber arena in one place. Indeed, the Cyber Command has three main tasks: the first one is in charge of daily defence of US networks and supports all the branches of military and counter-terrorism operations; the second one is providing centrality in the chain of command, and clarity and accountability as far as resources are concerned; the third task is a tight collaboration with all the departments of US government and also with private actors, such as critical infrastructures (*Ibidem*). During 2009, the Pentagon revealed the five pillar strategy described throughout this paragraph.

The second pillar of US strategy in cyberspace is a proactive cyber defence, opposed to a passive, or fortress, defence mentality. A proactive posture could be split in two different sets of actions. The first one involves resilience of government, military and critical infrastructures. A resilient approach means a whole-of-government, holistic posture, which is able to anticipate a potential cyber attack, or it is able - if a CNO manages to pass through defences, to stop the threat and recover as swiftly as possible, ensuring the operational continuity of the infrastructure (Bologna, Fasani, Martellini, 2013). The second set of action concerns an active collection of information as wide and precise as possible about potential threats, in order to prepare adequate responses for both cyber as well as kinetic menaces (Colbaugh and Glass, 2011). We could insert the fact that the NSA was already inside North Korean systems, emerged after the Sony Hack case, as one example of proactive cyber defence. Furthermore, the stockpiling of 0days by the NSA (Burkart, P., & McCourt, 2017) and also the fact that, according to Vault7 leaks, the US established a presence in European allies' systems as another example of active defence, albeit against the mutual understanding and shared rules among allies (Shane, Rosenberg and Lehren, 2017).

The third strategic pillar is the defence of critical infrastructures. Being the backbone of a nation, it is of foremost importance that they are protected against cyber attacks. For this reason, in 2013 president Barack Obama issued the *Executive Order on Improving Critical Infrastructure Cybersecurity* and the *Presidential Policy Directive on Critical Infrastructure Security and Resilience* (Department of Homeland Security, 2013; White House, 2013). These two normative measures go in the direction of an holistic approach for security and resiliency already tackled in the description of the second pillar, focusing however only on critical infrastructures. For example, they foster on the promotion of the adoption of Hugh-level cyber security measure, increasing public-private partnership and information sharing both horizontal - among infrastructures - as well as vertical - with the government.

The fourth pillar consists in what is called collective defence. Collective cyber defences could be applied mainly with US allies, and it is fostered also by NATO. Under the provision of collective cyber defence, all allies pledged to improve their national cyber defence capabilities, which they are responsible of, but also to share important information about potential threats, or occurring hostile CNOs (NATO, 2017). Furthermore, under collective defence one could find joint training sessions and exercises, like war games and scenario analyses (*Ibidem*). Such a collective approach should extend the defensive capabilities of the US which could count on expertise and experience from other allied countries.

The fifth pillar involves keeping the technological advantage. Strictly connected to the introductive portion of this paragraph, being the US the country where cyberspace originated, the US has always had, and keeps in having, the technological upper hand. This is due also by the capability of the US to allocate funds for research and development of cyber tools, both for defensive as well as offensive purposes. As far as offensive capabilities are concerned, in 2013, thanks to the leaks of Edward Snowden, the general public became aware of the previously unknown hostile CNOs carried out by the US and also of the so-called "black budget", namely the secret budget that listed how much money was allocated by the US for cyber operations and also (Gellman and Nakashima, 2013). The Tailored Access Unit (TAO) is the unit responsible for conducting this kind of operation, among which it is found GENIE, a $652 million project to place covert implants in carefully chosen machines around the world (*Ibidem*). However, the budget defines these attacks as part of the aforementioned active defence, and are directed mainly towards China, Russia, Iran and North Korea (*Ibidem*). Therefore the US is employing resources, in terms of money and personnel, in order to keep on having the technological upper hand that means also bolstering the situation of escalation dominance in case of hostilities should happen, both in cyber as well as physical disputes.

China

In China Computer Network Attacks and Exploitation began to be used as a strategic tool from the late Nineties, but they were allegedly carried out by groups of hackers that acted under a patriotic spur. The issue was two-faced. On the one hand, attacks performed by civilians questioned the control of the Party over the military use of cyberspace. On the other hand, redirecting the responsibility for these attacks to those independent group served as mean of justification when the Chinese government was blamed as the source of hostile CNOs. Nonetheless, those patriotic hackers served as a large basis of workforce carrying out sequences of CNA and CNEs under the consent of the government (Lewis, 2013). This ambiguous situation further blurs the already very thin line that distinguishes an act of an independent, willing, civilian to carry out a cyber attack that could bring even a little benefit (even a nuisance to an enemy) for his, or her, country, and a state-sponsored attack. For this reason, the Chinese government has been able to deny every accusation from other countries, mainly coming from the U.S., of systematically using cyber tools to carry out information gathering and reconnaissance attacks (Hjortdal, 2011).

Recent events made known that, apart from using nationalists as "cyber pawns" at military level offensive operations in cyberspace are performed by a specific unit of the PLA, the 61398, and this thanks to an admission of Chinese Official last year, after years of denial (Tiezzi, 2015). The sudden declaration could be a show of force posture from the Chinese government toward the rest of the international community.

China does not have a doctrine focused exclusively on cyber warfare, which it doesn't even call "cyber warfare" but *information warfare* instead, defining a conflict in cyberspace at nation-state level that involves both direct military confrontation and also indirect competition through espionage, reconnaissance, disruption and deception (Kramer, Starr & Wentz, 2009). For this very reason both information operations and computer network operations are seen as powerful tools to bridge the military asymmetry with other countries, and more specifically the United States. The evolution of networked systems, enabled China to enter American (cyber)space without effort (Mulvenon, 2009).

Chinese strategy in cyberspace looks in two directions: espionage and deterrence. The first one is also the simplest one to grasp, because attacks aimed at stealing information and constitute the principal activity of the panoply of attacks coming from Chinese territory. The exploitation of cyberspace for espionage purposes permits China to leapfrog the technological and economical gap with the US. For example, technologically speaking, China uses cyber incursions to illicitly obtain sensible data, such as 50 terabytes stolen from US defence contractors which contained various American military secrets, such as data on the B-2 bomber, F-22 and F-35 (Gertz, 2016). The Lockheed Martin F-35 Lightning II is one of the most sophisticated fighter jet currently existing, and the Chinese were able to acquire radar modules and engine blueprints and allegedly used that same knowledge to build their J-31.

Despite not having a clear and known strategy about the warlike use of cyberspace, we could infer that its military exploitation is part of a bigger military strategy. Indeed, cyber power fits into a greater strategic framework that revolves around an asymmetrical idea of warfare, in order to be able to compete on par with the United States. This strategy is called *Shāshǒujiàn*, literally "assassin's mace" and it is a so-called A2/AD, that stands for Anti-

Access/Area Denial. Its aim is to impair the American ability to project its influence in the western pacific that could enable access to political and economic interests (Krepinevich, 2010). The assassin's mace is a double strategy, and it is composed by an anti-access part, that serves to inhibit US forces from perform military movements in a given theatre of operation, and an area-denial part, which blocks enemy freedom of action in area under Chinese control, thus narrowing American strategic options in the Western Pacific area (McCarthy, 2010). It is important to underline that Computer Network Attacks and Exploitation are only a part of this strategy, which involves also ASAT weapons that is to say anti-satellite measures. The peculiarity of this strategy is that does not involve the use of kinetic weapons but rests on networked measures. That is the essence of the anti-access/area denial strategy, that could remind us of two strategists of the past: first of all, the teachings of Sun Tzu of "winning the war without fighting", and Mao Zedong's concept of "protracted war", that focuses on making the enemy blind and deaf, and confusing its means of communication (Krepinevich, 2010).

The use of Computer Network Operations by the Chinese government is focused specifically on Computer Network Exploitation. These are seen as a mean to balance an imbalance that is mainly military and technological. In line with the Anti-Access/Area denial activities, Computer Network Attacks are to be used preemptively, not as a force multiplier, before a possible kinetic conflict, in order to gain a strategic advantage jamming enemy communications, thus impairing the enemy to collect intelligence and to communicate internally securely (Lewis, 2013). Furthermore, China is known for its extensive use of espionage operations. What makes them attractive is the fact are as cheap as they are useful to acquire military and economic advantage, and for this reason they have increased in numbers in the last ten years (Lewis, 2013). These incursions could be directed for example, to foreign defence contractors, research and development institutes, government agencies, such as the recent attack against the U.S. Personnel Management Office in order to steal personal data of the government's employees.

As stated before, the history of publicly known Chinese offensive cyber operations begins mid-1990, and indeed the beginning of the A2/AD starts in 1993, when Chinese government began exploring this new technology in order to block an American expansion in the Western Pacific. Indeed, at military level, China uses cyber power in order to cast its shadow over Taiwan, and does this in two ways: addressing Taiwan with cyber attacks to coerce its reunification with China, and at the same time excluding the United States from intervening, by means of deterrence (Mulvenon, 2009). This constitutes a great part of

Chinese strategy, namely deterrence by denial against the United States. Knowing where American forces are displaced and being able to perform crippling attacks to the command and control networks is a significant strategic capability, also in light of the fact that the US Command and Control Chain, and the time-phased force and deployment data are heavily dependent on networks, and this translates in the fact that the costs the US incurs for escaping such a measure are too high, and therefore the military power projection is limited.

When analysing the Chinese use of Computer Network Attacks, one should always remember the larger strategic context in which the Assassin's Mace operates. Since the Straits Crisis with Taiwan in 1996, China has been trying to make the costs for the US presence in the Western Pacific prohibitive (Krepinevich, 2010). Computer Network Attacks are a weapon best suited for pursuing a strategy that excludes the use of kinetic weapons and, at the same time the exclusion of the US from projecting influence over the island. Not only, Computer Network Attacks and Exploitation are used, together with other means the likes of military exercises and other means of national power, to directly provoke Taiwan – a heavily interconnected territory – and its critical infrastructures, in order to destabilise the population's will (Mulvenon, 2009). It is straightforward that Taiwan wouldn't be able to militarily respond to a crisis, and a United States intervention in such a crisis is a delicate subject to treat, because it would very likely end into an escalatory process. Therefore, the Chinese government is using these "soft measures" to destabilise the area, to isolate Taiwan from American influence and possibly coerce the islands' government into negotiation.

The added result of all these offensive cyber operation is that China, through gaining military and economic secrets, acquires a sort of security independence that strengthens its position as leader in the region, given the fact that all the other neighbouring countries have tighten ties with the U.S. for security and for acquiring defensive capabilities. Cooperation and security agreements surely aid the protection of a country but limit the development of one's cyber capabilities and bind the defence of a country to the United States, which in turn is able to project power in the region, the main thing that China is trying to avoid.

Russia

If the US concentrates on intelligence and active defence measures (with the exception of Stuxnet as a disruptive attack), and China on espionage for both political and commercial purposes, Russia's strategy sits on a whole different level, namely information warfare and, albeit recently, disruptive attacks. “Information warfare” is how Russian government and military refer to describe a concept very similar to “cyber warfare” (Molander et al, 1996). Russia, together with the US and China was among the first states to recognise the importance and the value of the cyber mean to acquire sensitive information as well as to cause damage and paralyse the enemy.

Russia strategy in cyberspace begins officially in 2000, with president Putin issuing the *Information Security Doctrine of the Russian Federation* (Russian Federation, 2010). The official doctrine constitutes a document unique of its kind, which enlists principles and objectives, and furthermore, how to follow the exact path in order to link these two extremities within the Russian national security policy (*Ibidem*). Compared to countries like the US and its western allies, which focuses on the security of critical infrastructures as one of the main strategic pillars, the doctrine of the Russian federation emphasises the importance of securing federal institutions, in order to have a better and secure leadership for the country (*Ibidem*). Due to the increasing number of hostile CNOs suffered from Russia, in two 2006 the doctrine was updated, underlining the need of establishing security thresholds for the safety of national security and stability, pushing also for cooperation and information sharing among the allies (Gady and Austin, 2010). 2007 was the year of the attacks against the Estonian government, which accused Russia of being the culprit behind the attacks, and 2008 was the years of the attacks against Georgia, which were deployed in parallel with its physical military intervention (Klimburg, 2011; Rid, 2012). Russia was able to signal the importance of information warfare and its power stance to two other countries. Indeed, the Distributed Denial of Service attacks against Estonia could be considered the first instance of state-sponsored CNA against a state actor. In 2010 Russia issued its military doctrine, where, possibly following the experience of the attacks against Georgia, it stressed the usefulness of CNOs in the starting phases of a conflict as a mean to impair the adversary command and control as well as a mean to conduct information campaigns to shift the perception of the international general public in favour of the Russian Federation (Heickero, 2010).

In the following years, Russia followed through both the approaches. Starting from the latter, it could be inferred, for example that the attacks against the last US presidential campaign fall into the framework of information war in order to change perceptions favourably for the Russian Federation (Fidler, 2016). It is important to underline that this influence of information does not necessarily shape a favourable view *of* the Russian Federation but it is more likely to be used to shape a public opinion in a way that is favourable *for* the Russian Federation. As far as the first approach is concerned, namely the impairing of command and control centres, Russia started testing cyber weapons inside wider cyber operations that are able to cause physical damage to infrastructures. For example, in 2015 Russia attacked an electric power grid in Ukraine, that left thousands of Ukrainian people without electricity (Lee, Assante and Conway, 2016). A recent case that due to the lack of extensive analysis (compared to the literature on Stuxnet) was not inserted as a case study as far as asymmetry is concerned, but it is useful to mention briefly.

Ukraine and Russia relations were not amicable since the collapse of the Soviet Union and worsened after Ukraine joined NATO in 2004, which was perceived as an anti-Russia move (Krickovic, 2016). The relations worsened more recently after the Russo-Georgian conflict, where the Kremlin accused Ukraine to support Georgia by selling arms to Georgian troops (Schwartz, 2009). Furthermore, the crisis blew in full scale after the Ukrainian revolution and the subsequent annexation of Crimea by the Russian Federation in 2014 (Gardner, 2016). Given this history of tension, Ukraine has recently become a test-bed for Russian cyber weapons. During December 2015 Russia deployed the BlackEnergy malware, which was able to shut down an electric grid which left half a million Ukrainian in the dark. As was mentioned before, the forensic investigation is still ongoing, but it seems that BlackEnergy exploited vulnerabilities inside macro applications within Microsoft Office, and was able to take control of the management system of the electric grid. This attack was performed jointly with a second attack which targeted the emergency service of the power plant, rendering impossible any telephone communication in order to make the blackout last as long as possible (Kovacs, 2016). Furthermore, to BlackEnergy another cyber weapon was recently discovered, and it was called CrashOverride (Greenberg, 2017). CrashOverride has the same ultimate target as BlackEnergy, namely the hardware of power plants, but has a different payload. If BlackEnergy was able to shut down the electric supply through the intrusion in the electronic management system, CrashOverride targets the hardware itself, in a Stuxnet-like fashion (*Ibidem*). Moreover, it is very likely that the attacks occurred at the end of June 2017, that were masked a ransomware campaign similar to WannaCry (Suiche, 2017) were not criminal in nature, but instead were in reality part of a targeted campaign against Ukrainian infrastructures, with the aim of wiping hard disks from major institutions for example Ukraine's central bank, the state telecommunication service, the metro and Boryspil airport, in Kiev (Brandom, 2017). The fact that it was not a ransomware campaign was that it exploited a vulnerability in the database of MeDoc, the main IT services supplier in Ukraine, to which most of Ukrainian institutions refer to (Suiche, 2017), and that the decryption of data is impossible to perform, and instead, the malware based on Petya (a 2016 ransomware) wipes the hard disk irreversibly. Kenneth Geers NATO ambassador who focuses on cyber security issues, examining the situation stated that “you can’t really find a space in Ukraine where there hasn’t been an attack” (Greenberg, 2017). Petro Poroshenko, Ukrainian president confirmed that only in October and November 2016, Ukrainian was victim of 6,500 cyber attacks against 36 Ukrainian targets and that the investigations pointed against Russia, which attacked Ukraine directly or sponsoring private armies of hackers (*Ibidem*). As far as asymmetry is concerned, is straightforward to see how Ukraine has no ability to retaliate, neither with physical nor kinetic means and it is therefore a victim of the escalation dominant position of the Russian Federation.

After having analysed the posture of the three main cyber powers in the world today, it is useful to analyse also the occasions where there has been a cyber dispute or exchange of hostile CNOs between these countries. As was stated at the beginning, the analysis of cyber disputes between major power is more difficult, due to the fact that both China and Russia tend not to publish news about being victims of hostile CNOs and the US avoids - straightforwardly - telling when it attacked, or attempted to, another country through cyber means. Furthermore, the difficulty in analysing hostile exchanges of CNOs between major power proves difficult because it concerns a level where communication becomes crucial. A misplaced accusation, or worse, a misplaced retaliation poses the risk of a spill-over effect in the escalation process, which could become kinetic or worsening the relations between powerful countries. There have been no major CNAs among these three countries, and one should ask oneself the reason why. The main reason is, in line with the thesis of this dissertation, that the unit taken into consideration are major powers, and also nuclear powers, therefore there is a condition of symmetry. In addition to the self-restraint mechanisms described in the previous chapters, the absence of disruptive CNAs among major powers could be due to the fact that all three are deterred from using powerful cyber weapons in order not to disrupt the already fragile equilibrium in the international arena, that is to say the avoidance of a major crisis that could become a bilateral kinetic conflict. However, the absence of CNAs does not imply the absence of hostile or aggressive cyber activities. Indeed, all three perform campaigns of CNEs in line with their strategic approaches that shape the usage of cyber weapons. Therefore, the lack of an internationally shared normative framework, or a convention in line with the Geneva one, allows these countries to intrude each other territories through cyber means, without the risk of incurring in international sanctions or military intervention. Indeed, all three countries penetrate each other systems to collect information, what differs is the purpose: the US for mainly intelligence purposes as a mean to ensure attribution in case of aggressive CNOs; China collects information in order to bridge the technological gap with the US and to obtain strategic information for their political agenda in the Pacific; Russia intrudes US systems to collect information to be used for its information warfare purposes, for example shaping political perception for the international public.

Before seeing the following cases, is also useful to underline that, generally speaking, cyberspace as a war-fighting domain is a modern conception, which start could be placed after 2007, namely after the Estonia attacks, but definitely after 2010, when it was clear that CNOs could be used to cause damage in the physical realm. For this reason, instances of retaliation are difficult to pinpoint as forensics processes for attribution took longer, and, for example, the application of proactive cyber defence by the United States is a recent measures. Nonetheless it is important to see how these major power behave in cyberspace against one another.

United States of America and China

The first instance recorded in history of a CNO between the US and China, is a Chinese CNE against the US that occurred in 2003 and lasted until 2005. The operation was dubbed *Titan Rain* and was characterised by continuos attempts and actual intrusions into US government and military systems, plus private partners like Lockheed Martin, Sandia National Laboratories, and NASA (Mazanec, 2009). Forensics, headed by Sandia National Laboratories' Shawn Carpenter, traced the origin of the attack to China, more precisely in the Guangdong province (Rogin, 2010). The malware used for this CNE was equipped with a scanner, which searched for vulnerabilities on single machines and, once found, the PLA penetrated the systems and siphoned confidential information (Thornburg, 2005). The data that was stolen was not public but neither classified, as classified information were stored in facilities not connected to internet or intranet, nonetheless revealed projects and logistic information about US military, for example (*Ibidem*).

In 2006, the Naval War College was breached by the Chinese army. The College promptly put offline all of its network and the Pentagon raised the alert status for 5 million personal computers and 12,000 networks. The choice of the target is very likely to stem from the fact that the Naval War College is where the Strategic Group plans technique and conducts war games to practice them for possible cyber conflict, as well as from the fact that the College is the place where military strategy against China is planned (Rogin, 2007). The method used for the penetration was a spear phishing campaign and was able to propagate thanks to the fact that the Naval War College intranet was not updated as far as security measures is concerned (*Ibidem*).

In 2008 a Chinese CNE targeted the US election campaign, trying to find positions on China, as well as obtaining private exchanges between McCain and the recently elected Taiwanese president (Sasso, 2013). Also in this case the vulnerability exploited was the human one, thanks to a phishing attack against the candidates' staff, and once in their system the malware spread into their networks (*Ibidem*).

In 2009 Lockheed Martin, US defence contractor, responsible for the Research and Development of the F-35 programme revealed that it was the target of Chinese hacking (Rogin, 2010). Indeed, the responsible was a state-sponsored hacker named Su Bin that repeatedly broke into Lockheed's systems in order to steal plans for both the F-35 and F-22 fighter jets (O'Hare, 2016). These secret plans were then sent to the Chinese government which was able to use key technologies to build its fifth generation fighter jet, the J-20 (Thornill, 2016). Su Bin was arrested as he was on American soil, and after he admitted his culpability was jailed (O'Hare, 2016). This is one of the best examples to show how China performs intrusions against the US to accomplish political and military goals.

In 2010, Operation Aurora, a series of Advanced Persistent Threats, targeted Google, Symantec, Juniper Networks, Rackspace (all technological, security and defence contractors companies) as well as Yahoo and Adobe (Zetter, 2010). The APT was more sophisticated than previous CNEs, mainly because it exploited a 0day vulnerability found in internet explorer (Kurtz, 2010). The intruding vector was a spear phishing attack aimed at selected individuals who were likely to have access to sensitive information (*Ibidem*). The inadvertent point of entry clicked on a link which exploited the aforementioned 0day and permitted the malware to download itself in the system and open a backdoor, in order to be able to perform reconnaissance and control the infected system, searching for sensitive information and siphon them (*Ibidem*). Operation Aurora demonstrated the evolving capability of Chinese army and state-sponsored attackers to acquire intellectual property. Indeed, Operation Aurora was the joint effort of PLA Unit 69398 - the branch responsible for CNOs -, the Elderwood Gang and the Comment Crew (also called APT1), which employ hundreds of nationalists individuals tied with the government (Sanger, Barboza and Perlroth, 2013).

In 2013 Edward Snowden, former third party security agency contractor for the US government, revealed that the US penetrated Chinese mobile companies to spy on communication and, moreover, hacked into the systems of Tsinghua University. This university, apart from being one of the largest institution on China's soil hosts the China Education and Research Network, one of the main backbone networks of China. An intrusion into these systems means being able to collect huge amount of information such as internet data (Rapoza, 2013). Furthermore, Snowden also revealed that the NSA penetrated and created backdoors into Huawei networks, during an operation codenamed *Shotgiant* (Sanger and Perlroth, 2014). The main objective of the operation was to find a link between Huawei and the PLA, driven by the fear that Huawei - which sells internationally - would use its products, such as smartphones and routers, as beacons to spy on individuals, especially US military officials (*Ibidem*). Moreover, many clients of Huawei are countries that do not have commercial relations with the US nor buy US products. Being in Huawei's network would give the US a strategic advantage, namely monitoring foreign networks of interests, like Iran, Cuba, Afghanistan, Pakistan and Kenya (*Ibidem*).

In 2014, the Department of Justice of the United States indicted five PLA hackers for the attacks against six American entities belonging to nuclear power, metals, and solar products industries, for economic espionage (Schmidt and Sanger, 2014). Espionage is usually justified by the United States but it seems that in this case China has crossed the line, as many commercial secrets are stolen from the United States to seek economic advantage. The indictment was a clear signal that China has to put some boundaries, limiting the scope of actions of its PLA units. (Schmidt and Sanger, 2014). Following this historical event, another followed. In 2015 there was the first U.S. – China bilateral agreement on cyber issues: the Cyber Agreement stated that neither of the two governments will willingly support cyber espionage for commercial advantage (differentiating it from espionage carried out for national security reasons) (Rollins, 2105). This was a first, important step in history of cybersecurity and Sino-American relationship, but it was as important as it was pointless. Indeed, during 2016, state sponsored espionage operations originating from China successfully targeted U.S. government and companies (Gady, 2016). Given that for China commercial benefit overlaps with national security, what remains to be seen is how far will China go in pursuing cyber espionage activities and how much will the United States tolerate this situation, despite the agreement.

United States of America and Russian Federation

The timeline of hostile CNOs between Russia and the US is far more shorter and recent compared to the one between China and the US.

However, the first cyber operation that involved the two countries goes back to 1998. The operation codenamed Moonlight Maze, was a breach into US institutions' systems that originated in the Russian Federation (Rid, 2016). Similarly to early Chinese CNEs, the documents siphoned were unclassified but sensitive nonetheless, as they concerned technologies for military applications (*Ibidem*). The main victims of the breach were the US Army, NASA, the Department of Energy,Los Alamos and Sandia National Laboratories, Air Force Institute of Technology and Army research laboratories indicating that they the attackers had precise targets (*Ibidem*). The case was solved thanks to a so-called honey pot. This very simple technique allows the defenders to infect a file, in this case a document that could be valuable for the attacker, when the attacker steals it and opens it it connects to the defender systems which is able to trace the source. Through this method, US officials discovered that the Russian attackers only used two hop points, one in London, and one in several different universities in parallel, not comparable to a modern proxy, and connected directly to a machine in Moscow (*Ibidem*).

Then Russia laid very low, using the IT infrastructure as a mean of information warfare, building up so-called "troll armies" that is to say state-sponsored internet sockpuppetry or propaganda, namely individuals that publish blog posts and comment in the designated sections in national and international newspapers, posting pro-government arguments (Al-Ketheeb, Agarwal, 2016). Indeed, Russia was the first country to establish such an unofficial institution, in 2003 with the Веб-бригады, literally "Web Brigades".

Information warfare campaigns conducted by Russia reached a peak in 2016. The first step of the campaign was in July, when the an attack coming directly from the Russian government - no state-sponsored proxy used - breached into the Democratic National Committee systems, stealing almost 20,000 personal e-mails and passing them to Wikileaks to be published immediately (Entous, Nakashima and Miller, 2016). The leak had the effect to discredit both Democratic runners, as a great part of the e-mail leaked were aimed at deriding and belittling Bernie Sanders, and showed how the whole DNC fundraising mechanisms, capable of raising millions of dollars thanks to entire dossier on donors, which contained "interests, annoyances and passions" (Confessore and Eder, 2016).

In October 2016, NBC reported that the CIA received an order to prepare a cyber attack to be deployed against the Russian federation in retaliation to the continuous interferences of Russian hackers in the US presidential elections; the order allegedly came directly by President Barack Obama (Arkin, Dilanian, Windrem, 2016). Also Vice President Joe Biden accused directly Putin to be trying to rig the election and that the administration was ready to send him a message, namely retaliation (Sanger, 2016). Both the CNEs were conducted by two separate Russian groups, codenamed *Fancy Bear* (or APT28) and *Cozy Bear* (or APT29) the first operating under the GRU (which stands for *Glavnoe Razvedyvatel'noe Upravlenie*, namely the main Russian military foreign intelligence agency) and the second operating under the SVR (which stands for *Sluzhba vneshney razvedki*, which is the Russian civilian foreign intelligence agency, GRU's counterpart).

Alleged incentives given to Russia by a Trump presidency aside, the infowar it is likely to be retaliation for the fact that former secretary of State and democratic candidate during the 2016 US presidential elections Hillary Clinton spurred civilian and political protests in Moscow against Putin's declaration to run for his third term in 2011 (Elder, 2011). A link could be also threaded between the *Panama Papers* that were published just a a few month before Russian legislative elections of 2016. Inside the Panama papers, documents were tying Vladimir Putin and his entourage to a widespread system of corruption and bribes, which brought analysts and commentators to define Russia as a kleptocracy (Adomeit, 2016). The Kremlin dismissed the allegation of corruption as the enemy's way to destabilise him, and that is why Putin could have purported an information war against the US, as he perceived it as an in-kind retaliation.

Recently, the revelation about the aforementioned CrashOverride malware, troubled US analysts, because, after Estonia 2007, Georgia 2008 and Ukraine 2015 there is the perception that Russia is testing malware aimed at disruption and also physical destruction of power grids which is very likely to cause indirect deaths (Greenberg, 2017). US analysts are worried because such a cyber weapon could easily hit American power grid which is deemed to be vulnerable to such attacks.

We can safely conclude that among countries with the same level of power, that is to say dyads between which there is a situation of symmetry in both power understood generally and, as a consequence, cyber power there is absence of disruptive CNAs. Instead, CNEs are widely used. Whether they are CNEs employed as a mean of intelligence gathering, political, military and commercial espionage or information warfare, intrusions are a safe method of signalling to the enemy a certain level of cyber capability, and broadly speaking: power. CNEs are the best way to project power posture and - at the same time - use a safe, self-restraining approach. The projection of power is due to the fact that a CNE implies the trespassing into another nation by defeating cyber defences in place. Furthermore, at least for now, it seems that responses in kind are silently accepted, until a certain level. Nonetheless, if a red line is trespassed, communication - another form of signalling - takes place. For example the indictment of the PLA officials and the public accusation and threats of retaliation of Joe Biden against Putin are a way of signalling that the US were able to attribute CNEs against its institutions, using public discourse to shape a narrative and to deter further attacks. By employing CNEs, self-restraint and tit-for-tat signalling are the way through which symmetric powers deal each other through cyber means, at least for the time being.

Among symmetric powers, the geopolitical context still matters. China is contending the US the title of global superpower and uses CNEs to bridge the military technological gap by stealing plans and blueprints for technological secrets, and also to better plan and execute its political and military strategy in the Pacific region, by trying to acquire as much information as possible about US strategies and positions in the region. In the case of Russia instead, CNEs aimed at information warfare are used exactly to shape the geopolitical context in its favour. The problem is that Russia is that is increasingly using disruptive offensive capabilities and it is adopting a salami slicing approach. Indeed, Russia is testing both cyber weapon as well as boundaries, to see where is the limit of the freedom of action in cyberspace as far as offensive CNAs are concerned, against targets with which there is a geopolitical context that "justifies" an attack and with which there is a situation of asymmetry and therefore the impossibility of retaliation, namely Estonia, Georgia and Ukraine.

The absence of an internationally shared normative framework that could employ also different degree of punishments allows countries to shift their dispute into cyberspace without risking international political fractures that could become kinetic conflicts. However, even without an international cyber law, normal international law could already be applied. Leveraging on the notion of due diligence, the US could legally prosecute both Russia and China, due to the fact that they never did anything to stop private individuals from attacking to another nation state - in this case the US - which is required by international law.

At first glance, the avoidance of enforcing the international law in this sense could stem from the attribution problem, but in reality this becomes a pointless motivation because - as it was underlined before - the US has the means to overcome this problem. Instead, this could be considered another mean of self- restraint because the lack of a normative framework advantages also the US that trespasses several national jurisdiction in order to employ its active defence.

Conclusions

This dissertation had the objective of answering two main research questions, that where aimed at discovering *when* and *how* cyber disputes take place. The *when* question aimed at researching the main condition that causes states to employ cyber weapons against each other. The *how* question concerned the characteristics of the behaviour or states in cyberspace, and how self-restraining mechanisms and symmetry – or lack thereof – between dyads of states involved in cyber disputes influences how states engage against each other.

This research presents two main original findings. The first finding is that a condition of political tension or hostility, stemming from conflictual strategies and postures, between states – namely, the independent variable – is a condition common to all the analysed cases, namely dyads of states exchanging hostile CNOs, which represents the dependent variable. Counterfactually, there are no recorded cases in literature of allied countries engaging in cyber disputes.

The second finding is that there is a causal mechanism between the states that confront each other in cyberspace and the cyber weapon(s) employed in such disputes. Retaliation between dyads of states constitutes the dependent variable of this research, and indeed this research showed how it is influenced by the two independent variables taken into consideration.

The first one is the symmetry of military power generally understood, which influences the intensity and the scale of CNOs. Asymmetry situations could be divided in two different scenarios, one where a more powerful country attacks a less powerful country, the other one when a less powerful country attacks a more powerful country. The first scenario was represented by the Stuxnet case, where the US attacked Iran. The US attack was a disruptive and destructive CNA against a critical infrastructure, the Iranian retaliation was in-kind, namely through cyber means, but lower in intensity, namely CNEs, and targeted minor critical infrastructures (online banking systems and military contractors), and also horizontal, since also US allies infrastructures were attacked in retaliation. Therefore, we could conclude that escalation dominance does not impede a lower country to retaliate in cyberspace but the lower power influences the intensity and the scope of the retaliation. The second scenario was represented by the Sony Hack case. North Korea attacked with low disruptive CNEs a private company on American soil, the retaliation of the US was vertical in nature, both in-kind as well physical, namely a disruptive CNA against a North Korean major critical infrastructure as well as economic sanctions. Given the power superiority of the US, North Korea did not retaliate against US response.

In situations of symmetry of power there is lack of disruptive CNAs and instead an exchange of intrusions, namely CNEs, in a tit-for-tat fashion is to be found. In situations of asymmetry the movements of attacks and what could be considered retaliations is not vertical, because the intensity always remains the same, that is to say intrusions in enemy's systems to perform CNEs, but only horizontal, namely the variation in the scope of the targets of the CNEs. In both cases, involving dyads of major powers, namely US and China and US and Russia, CNEs were used mainly as instruments to signal strategic postures, and power stances. Penetrating an enemy system without causing physical damage avoids creating conflicts that could spill-over to the physical domain. One interesting results was that, when it seemed that a red line (dependent on the perception of the actors) was crossed, what could be considered a vertical movement in the retaliation process took the shape of communication in the form of public accusations took place, for example through the indictment of PLAs officers and public attribution by Joe Biden against the Russian campaign aimed at interfering in the US presidential election process.

A second variable is constituted by self-restraining mechanisms put in place by states. Self-restraints are based on two pre-existing conditions, that is to say the absence of an internationally shared normative framework and strictly technical problems surrounding cyber weapons. For clarity it is useful to underline that the first on, allows state to perform freely to perform CNOs since it is not illegal to do so, due to the absence of a normative corpus together with a system able to impose sanctions . This characteristic could lead one to think that a deliberate increase of CNOs should be expected since there are no legal limits to perform them. Such a reasoning it is partly true, it has been already stated in the previous chapters there is an increase in CNOs, however we do not see deliberate escalations or huge campaigns of disruptive attacks because, the research concludes, the application of power in cyberspace is characterised by self-imposed limits applied by states when performing CNOs. The second condition involves the fact that states no matter how they are able to code sophisticated malware, do not possess complete control over it, as the Stuxnet case showed clearly. Indeed, self-restraints he fear of the possibility that the cyber weapon could surpass the boundaries of the target infrastructure, with a spill-over effect that could a) overcome the limits of the damage intended, creating an event that could spur an harsher retaliation, namely an accidental escalation; b) overcome the limits of the targeted infrastructure, creating a spill-over effect due to the interconnection between national infrastructure. In case this is a first attack, it could be perceived as a wider threat than intended, leading to – again – an accidental escalation, if this is a response to an attack (regarded by the attacker as such), it could be perceived as an horizontal escalation; c) overcome the boundaries of the targeted state, attacking by mistake other neighbouring countries, still due to interconnectivity among infrastructures. Furthermore, the very technical nature of cyber weapons acts as a restraint on their use. Cyber weapons are temporary in nature, due to the fact that once the vulnerability or vulnerabilities it exploits are discovered, the window of opportunity and the usefulness of cyber weapons decreases.

The empirical analysis ultimately showed that an escalation that follows Khan's ladder does not exists in cyberspace, because in situation of asymmetry of power, the less powerful country does not retaliate, suffering from the escalation dominant position of the counterpart or, if chooses to retaliate, does so by de-escalating, namely targeting a more easily penetrable target with a lower intensity cyber weapon. In cases of symmetry, moreover in cases of symmetry of major power, where one could have expected retaliation and escalation dynamics similar to the one of the Cold War - that is to say following Khan's ladder - we find self-restraint-driven tit-for-tat cyber exchanges of low intensity on a high number of targets, whether aimed at testing boundaries, or at information gathering or at shaping political opinions.

These original as well as important results constitute an important added value in the analysis of the new field of study on cyber disputes. To these main results other added values must be brought forward. The first one involves the first chapter, which is based on a complete and thorough excursus of literature concerning classical concepts of International Relations – namely state responsibility, security dilemma, deterrence and escalation – and how these could be translated into cyberspace. For example, this research constitutes an aid in pushing forward the discussion on attribution and deterrence, which are found to be strictly linked. Attribution always constituted the main obstacle in conducting analyses of cyber disputes, however methods like the active defence of the US, which is based on penetrating enemies' system in order to obtain intelligence provided, in two separate times (Sony Hack case and the Russian information warfare against the US presidential elections) sure attribution. The concept of active defence represents a double-edged sword because it gives incentives to state to penetrate the system of enemies and potential ones, worsening the security dilemma, but at the same time constitutes also an incentive to deterrence, due to the fact that if a state A perceives the fear that a foreign country established a foothold inside than it would be less willing to attack this foreign country or its allies, fearing also public accusations and even stronger retaliations. Furthermore, forensics analysis today is much more precise compared to the past, allowing for better backtracking of attacks that in turns betters the attribution process. Furthermore, as was bought forward by this research, the geopolitical context always matters, and is a crucial help in the attribution process.

Another added value of the research is constituted by the second chapter, which is constituted by an extensive, methodic and updated description of cyber weapons, the likes of which are not found in the political studies literature. The third added value encompasses both the first and the second chapter, throughout which policy suggestions were given tackling mainly state responsibility and the regulation of cyber weapons. These policy suggestions are required because of the lack of an internationally shared normative framework and the lack of applicability of international law to cyber disputes. This dissertation was not concerned with criminal activities, therefore the unit taken into consideration is the state. “The state” includes also state-sponsored attacks, namely CNOs performed by individuals motivated by a patriotic spur or hired directly by the government. This could be considered a mean for a state government to avoid culpability and deny responsibilities in case of public accusations coming from another state, claiming to be attacked. This is pointless for a main reason. The state is the ultimate warrantor for security of the entire nation, and it is responsible for the actions of their citizens when these attack other countries. This responsibility is called due diligence and it is a standard that should be applied when state-sponsored attack happen, and it could function as a deterrent. As duly outlined in the first chapter, a state-sponsored CNO is said to have origin in a particular state, then this state should help in the investigation process, reversing the burden of proof, that passes from the victim to the potential attacker. However, the main obstacle that such method could encounter is constituted by the difficulties of dealing with states that invoke the principle of non-interference, such as China and Russia. Nonetheless, due diligence could be applied even without the consent of the accused state, rendering the best way to have state and non-state actors complying to international law.

Furthermore, to reduce the number and the threat posed by cyber attack and to dissuade new actors to exploit these increasingly dangerous weapon, this dissertation suggests to international policy makers to push towards the extension of the notion of territorial sovereignty to national cyberspace as far international law is concerned, more precisely to the International Law Commission. In this sense, every intrusion should be viewed as a proper invasion, comparable to a platoon of soldiers taking control of a turret. As described in the first and second chapter, cyber weapons could be updated and their nature changed remotely, and for this reason, once inside a network or system, the payload could be changed from espionage-driven to a disruptive one. Treating a foreign presence inside a system in this way would also eliminate this problem of insecurity about the payload without the actual deployment. Therefore, the principle of territoriality should be extended to cyberspace during matters of espionage and attack just as it is done to the cybercriminal environment. This method would also be useful to eliminate completely the distinction between CNAs and CNEs, where the latter could not be considered acts of war for international law but poses an increasing threat nonetheless.

References

Adler, E. (1992). The emergence of cooperation: national epistemic communities and the international evolution of the idea of nuclear arms control. *International organization*, *46*(1), 101-145.

Adomeit, H. (2016). The 'Putin System'. Crime and Corruption as Constituent Building Blocks. *Europe-Asia Studies*, *68*(6), 1067-1073.

Al-Khateeb, S., & Agarwal, N. (2016). Understanding Strategic Information Manoeuvres in Network Media to Advance Cyber Operations: A Case Study Analysing Pro-Russian Separatists' Cyber Information Operations in Crimean Water Crisis. *Journal on Baltic Security*, *2*(1), 6-27.

Alberts, D. S., & Papp, D. S. (1997). *The information age: An anthology on its impact and consequences*. Office of the Assistant Secretary of Defense Washington DC Command and Control Research Program (CCRP).

Albright, D., Brannan, P., & Walrond, C. (2010). *Did Stuxnet take out 1,000 centrifuges at the Natanz enrichment plant?*. Institute for Science and International Security.

Arimatsu, L. (2012). A treaty for governing cyber-weapons: Potential benefits and practical limitations. In *Cyber conflict (CYCON), 2012 4th international conference, IEEE,* 1-19.

Arkin, W.M., Dilanian, K., Windrem, R. (2016). CIA Prepping for Possible Cyber Strike Against Russia. *NBC News.*

Armstrong, C. K. (2004). US-North Korean Relations. *Asian Perspective*, 13-37.

Armstrong, C. K. (2017). North Korea in 2016. *Asian Survey*, *57*(1), 119-127.

Arquilla, J., & Ronfeldt, D. (2001). *Networks and netwars: The future of terror, crime, and militancy*. Rand Corporation.

Arquilla, J., & Ronfeldt, D. (2001). *Networks and netwars: The future of terror, crime, and militancy*. Rand

Barnes, J. E., & Gorman, S. (2013). US says Iran hacked Navy computers. *The Wall Street Journal*, *27*.

Baseley-Walker, B. (2011). Transparency and confidence-building measures in cyberspace: towards norms of behaviour. In *Disarmament Forum: Confronting Cyberconflict* (Vol. 4, pp. 31-40).

Beaumont-Thomas, B. (2014). North Korea complains to UN about Seth Rogen comedy The Interview. *The Guardian, 10.*

Bertrand, N. (2015). Iran is building a non-nuclear threat faster than experts 'would have ever imagined'. *Business Insider.*

Biddle, S. (2010). *Military power: Explaining victory and defeat in modern battle*. Princeton University Press.

Blank, S. (2001). Can information warfare be deterred?. *Defense Analysis*, *17*(2), 121-138.

Bologna, S., Fasani, A., & Martellini, M. (2013). From Fortress to Resilience. In M. Martellini (eds) *Cyber Security*: *Deterrence and IT Protection for Critical Infrastructures* (pp. 53-56). Springer International Publishing.

Bond, P. (2014). Sony Hack: Activists to Drop 'Interview'DVDs over North Korea via Balloon. *The Hollywood Reporter*, *16*.

Boot, M. (2006). *War Made New: Technology, Warfare and the Course of History, 1500 to today,* Gotham Books, New York.

Brandom, R. (2017). A new ransomware attack is infecting airlines, banks, and utilities across Europe. *The Verge.*

Broad, W. J., Markoff, J., & Sanger, D. E. (2011). Israeli test on worm called crucial in Iran nuclear delay. *New York Times*, *15*, 2011.

Bronk, C., Tikk-Ringas, E. (2013). The cyber attack on Saudi Aramco. *Survival*, *55*(2), 81-96.

Brookes, A. (2011). US Pentagon to treat cyber-attacks as 'acts of war. *BBC News*, *1*.

Brown G., Yung C. D. (2017). Evaluating the US-China Cybersecurity Agreement, Part 1, *the Diplomat.*

Brown G., Yung C. D. (2017). Evaluating the US-China Cybersecurity Agreement, Part 3, *the Diplomat.*

Bruno, G. (2008). The evolution of cyber warfare. *Council on Foreign Relations*, *27*.

Bruno, G. (2010). Iran’s nuclear program. *Council on Foreign Relations*, *10*.

Buchanan, B. (2016). The Cybersecurity Dilemma. *London: Hurst*.

Bumiller, E., Shanker, T. (2012). Panetta Warns of Dire Threat of Cyber attack on U.S.. *The New York Times.*

Burkart, P., & McCourt, T. (2017). The international political economy of the hack: A closer look at markets for cybersecurity software. *Popular Communication*, *15*(1), 37-54.

Buzan, B., Wæver, O., & De Wilde, J. (1998). *Security: a new framework for analysis*. Lynne Rienner Publishers.

Capaccio, T. (2013). US General: Iranian Cyber attacks Are Retaliation For The Stuxnet Virus. *Business Insider.*

Cavaiola, L. J., Gompert, D. C., & Libicki, M. (2015). Cyber House Rules: On War, Retaliation and Escalation. *Survival*, *57*(1), 81-104.

Cavelty, M. D. (2015). The normalization of cyber-international relations. *Strategic Trends*, 81-98.

Cha, V. D. (2002). Korea's Place in the Axis. *Foreign Affairs*, *81*, 79.

Chandarimani, Y., Monrad, J (2016). Regional Advanced Threat Report: Europe, Middle East and Africa - 2H2015, FireEye Report.

Chen, T. (2010). Stuxnet, the real start of cyber warfare? [Editor's Note]. *IEEE Network*, *24*(6), 2-3.

Chen, T. M., & Abu-Nimeh, S. (2011). Lessons from stuxnet. *Computer*, *44*(4), 91-93.

Chilcoat, R. A. (Ed.). (1998). *1998 Strategic Assessment: Engaging Power for Peace*. National Defense University.

Cieply, M., & Barnes, B. (2014). Sony Cyber attack, first a nuisance, swiftly grew into a firestorm. *The New York Times. December*, *30*.

Cieply, M., & Barnes, B. (2014). Sony cyber attack, first a nuisance, swiftly grew into a firestorm. *The New York Times*, *30*.

Cimbala, S. J. (2012). *Clausewitz and Escalation: Classical Perspective on Nuclear Strategy*. Routledge.

Clarke, R. A., & Knake, R. K. (2014). *Cyber war*. Tantor Media, Incorporated.

Clausewitz, C. Von (1832, r. 1984). *On War*. translated by Michael Howard and Peter Paret. (Eds.) Princeton University Press, Princeton.

Colbaugh, R., & Glass, K. (2011, July). Proactive defense for evolving cyber threats. In *Intelligence and Security Informatics (ISI), 2011 IEEE International Conference on* (pp. 125-130). IEEE.

Collins, S., & McCombie, S. (2012). Stuxnet: the emergence of a new cyber weapon and its implications. *Journal of Policing, Intelligence and Counter Terrorism*, *7*(1), 80-91.

Collins, S., & McCombie, S. (2012). Stuxnet: the emergence of a new cyber weapon and its implications. *Journal of Policing, Intelligence and Counter Terrorism*, *7*(1), 80-91.

Committee on National Security Systems (2015). Committee on National Security Systems (CNSS) Glossary, CNSSI No. 4009. Available at: https://cryptosmith.files.wordpress.com/2015/08/glossary-2015-cnss.pdf

Confessore, N., Edere, S. (2016). In Hacked DNC Emails, a Glimpse of How Big Money Works. *The New York Times.*

Cordesman, A. H., & Cordesman, J. G. (2002). *Cyber-threats, information warfare, and critical infrastructure protection: defending the US homeland*. Greenwood Publishing Group.

Cornish, P., Livingstone, D., Clemente, D., & Yorke, C. (2010). *On Cyber Warfare* (pp. 21-22). London: Chatham House.

Crail, P. (2008). Iran starts new centrifuge installation campaign. *Arms Control Today*, *38*(4), 42.

Creveld, M. Van (1991). *Technology and War: From 2000 B.C. to the Present,* Free Press, New York.

Davis, A. (2015). A History of Hacking. *The Institute, IEEE.* Available at: http://theinstitute.ieee.org/technology-topics/cybersecurity/a-history-of-hacking

Davis, P. K. (2014). Deterrence, influence, cyber attack, and cyberwar. *NYUJ Int'l L. & Pol.*, *47*, 327.

Denning, D. (2000). Reflections on Cyberweapons Control. *Computer Security Journal*, *16*(4), 43-53.

Department of Homeland Security (2013) Executive Order 13636—Improving Critical Infrastructure Cyber Security.

Der Derian, J. (2000). Virtuous war/virtual theory. *International affairs*, *76*(4), 771-788.

Drogin, Bob (2010). In a doomsday cyber attack scenario, answers are unsettling. *Los Angeles Time.*

Eckstein, H. 1975. Case studies and theory in political science. In Greenstein, F., and N. Polsby, eds. Handbook of political science, vol. 7, Reading, MA: Addison-Wesley, 79-138.

Edgar, T. H. (2017). *Beyond Snowden: Privacy, Mass Surveillance, and the Struggle to Reform the NSA*. Brookings Institution Press.

Einhorn, R. J., & Samore, G. (2002). Ending Russian assistance to Iran's nuclear bomb. *Survival*, *44*(2), 51-70.

Elder, M. (2011). Vladimir Putin Accuses Hillary Clinton of Encouraging Russian Protests. *The Guardian*.

Entous, A., Nakashima, E., & Miller, G. (2016). Secret CIA assessment says Russia was trying to help Trump win White House. *The Washington Post*, *9*.

Eriksson, J. (Ed.). (2017). *Threat Politics: New Perspectives on Security, Risk and Crisis Management: New Perspectives on Security, Risk and Crisis Management*. Routledge.

European Council (2001). Convention on Cybercrime, opened for signature Nov. 23, 2001, Europ. T.S. No. 185. Available from: http://conventions.coe.int/Treaty/EN/projets/FinalCybercrime.htm

European Council (2016). Chart of Signatures and Ratifications of Treaty 185. Available from: http://conventions.coe.int/Treaty/EN /CadreListeTraites.htm

Everard, J. (2013). *Virtual States*. Routledge.

Falliere, N., Murchu, L. O., & Chien, E. (2011). W32. stuxnet dossier. *White paper, Symantec Corp., Security Response*, *5*, 6.

Falliere, N., Murchu, L. O., & Chien, E. (2011). W32. stuxnet dossier. *White paper, Symantec Corp., Security Response*, *5*(6).

Farwell, J. P., & Rohozinski, R. (2011). Stuxnet and the future of cyber war. *Survival*, *53*(1), 23-40.

Farwell, J. P., & Rohozinski, R. (2012). The new reality of cyber war. *Survival*, *54*(4), 107-120.

Farwell, J. P., & Rohozinski, R. (2012). The new reality of cyber war. *Survival*, *54*(4), 107-120.

Federation, Russian. (2010). The military doctrine of the Russian Federation. *Approved by Russian Federation by presidential edict, February*, *5*.

Fidler, D. P. (2016). The US Election Hacks, Cybersecurity, and International Law. *American Journal of International Law*, *110*, 337-342.

Fidler, D. P., Pregent, R., & Vandurme, A. (2016). NATO, Cyber Defense, and International Law. *Journal of International and Comparative Law*, *4*(1), 1.

Filiol, E. (2006). *Computer viruses: from theory to applications*. Springer Science & Business Media.

Forman, S., & Barnes, J. E. (2011). Cyber-Combat: Act of War—Pentagon Sets Stage for US to Respond to Computer Sabotage with Military Force. *Wall Street Journal*, *30*.

Furnell, S. (2010). Hackers, viruses and malicious software. *Handbook of Internet crime*, 173-193.

G7 (2017). Declaration on responsible states behaviour in cyberspace. Available at: http://www.esteri.it/mae/resource/doc/2017/04/declaration_on_cyberspace.pdf

Gady, F. S., & Austin, G. (2010). Russia, the United States, and Cyber Diplomacy. *Opening the Doors, EastWestInstitute, New York*.

Garamone, J. (2010). Lynn Explains US Cybersecurity Strategy. *US Department of Defence News*.

Gardner, H. (2016). The Russian annexation of Crimea: regional and global ramifications. *European Politics and Society*, *17*(4), 490-505.

Gartzke, E., & Lindsay, J. R. (2017). Thermonuclear cyberwar. *Journal of Cybersecurity*, *3*(1), 37-48.

Geers, K. (2011). Sun Tzu and cyber war. *Cooperative Cyber Defence Centre of Excellence, February*, *9*.

Geers, K., Kindlund, D., Moran, N., & Rachwald, R. (2014). *World War C: Understanding nation-state motives behind today's advanced cyber attacks*. Technical report, FireEye.

Gellman, B., & Nakashima, E. (2013). US spy agencies mounted 231 offensive cyber-operations in 2011, documents show. *Washington Post*, *31*.

George, A. L. (1979). Case Studies and Theory Development: The Method of Structured, Focused Comparison. In Paul Gordon Lauren, ed., *Diplomacy: New Approaches in History, Theory, and Policy.* New York: Free Press pp. 43-68.

George, A. L., & Bennett, A. (2005). The Method of structured, focused comparison. In *George, Alexander and Bennett, Andrew, ed. Case Studies and Theory Development in the Social Sciences. Cambridge: MIT*, 73-88.

Gertz, B. (2016). China Hacked F-22, F-35 Stealth Jet Secrets. *The Washington Free Beacon.*

Giacomello, G. and Badialetti, G. (2009). *Manuale di Studi Strategici – da Sun Tzu alle 'nuove guerre',* Vita e Pensiero, Milan.

Giles, K. (2012, June). Russia's public stance on cyberspace issues. In *Cyber Conflict (CYCON), International Conference on*. IEEE.

Glaser, C. L. (2011). Deterrence of cyber attacks and US national security. *2011 Developing Cyber Security Synergy*, 47.

Goldsmith, J. (2015). How Cyber Changes the Laws of War. In *Current and Emerging Trends in Cyber Operations* (pp. 51-61). Palgrave Macmillan UK.

Goldstein, J. (1996). International law and domestic institutions: reconciling North American "unfair" trade laws. *International Organization*, *50*(4), 541-564.

Gootman, E. (2006). Security Council Approves Sanctions Against Iran Over Nuclear Program. *The New York Times*, *24*(2006), A1.

Greenberg, A. (2012). Shopping for zero-days: A price list for hackers' secret software exploits. *Forbes*.

Greenberg, A. (2015). FBI Director: Sony's 'Sloppy'North Korean Hackers Revealed Their IP Addresses.". *Wired.*

Greenberg, A. (2017). How An Entire Nation Became Russia's Test Lab For Cyberwar. *Wired.*

Hafner, K., & Lyon, M. (1998). *Where wizards stay up late: The origins of the Internet*. Simon and Schuster.

Hafner, K., & Markoff, J. (1995). *Cyberpunk: outlaws and hackers on the computer frontier, revised*. Simon and Schuster.

Haggard, S., & Lindsay, J. R. (2015). North Korea and the Sony Hack: exporting instability through cyberspace.

Haggard, S., & Lindsay, J. R. (2015). North Korea and the Sony Hack: exporting instability through cyberspace.

Hammond, G. T. (1993). *Plowshares into Swords: Arms Races in International Politics, 1840-1991*. University of South Carolina Press.

Harknett, R. J. (1996). Information warfare and deterrence. *Parameters*, *26*(3), 93.

Hayes, R. E., & Wheatley, G. (1996). *Information warfare and deterrence* (No. 87). National Defense Univ Washington Dc Inst For National Strategic Studies.

Healey, J. (Ed.). (2013). *A Fierce Domain: Conflict in Cyberspace, 1986 to 2012*. Cyber Conflict Studies Association.

Heickerö, R. (2010). *Emerging cyber threats and Russian views on Information warfare and Information operations*. Defence Analysis, Swedish Defence Research Agency (FOI).

Herr, T., & Rosenzweig, P. (2014). Cyber Weapons & Export Control: Incorporating Dual Use with the PrEP Model.

Hjortdal, M. (2011). China's use of cyber warfare: Espionage meets strategic deterrence. *Journal of Strategic Security*, *4*(2), 1;

Hughes, R. B. (2009) NATO and Cyber Defence: Mission Accomplished?. *NATO-OTAN*

Huntington, S. P. (1958). Arms Races-Prerequisites And Results. *Public Policy*, *8*, 41-86.

Ikenberry, G. J. (1996). The future of international leadership. *Political Science Quarterly*, *111*(3), 385-402.

Ilascu, I. (2014). Predictions for APT Attacks Go from Bad to Worse in 2015, Softpedia. Available at: http://news.softpedia.com/news/Predictions-For-APT-Attacks-Go-From-Bad-To-Worse-In-2015-467227.shtml

Infosec (2014). DoS Attacks and Free DoS Attacking Tools. *Infosec Institute Reports*. Available from: http://resources.infosecinstitute.com/dos-attacks-free-dos-attacking-tools

Infosec (2015). Current Trends in the APT World. Available at: http://resources.infosecinstitute.com/current-trends-apt-world/#gref

Inkster, N. (2015). Cyber Attacks in La-La Land. *Survival*, *57*(1), 105-116.

Jahn, G. (2007). IAEA Chief Exhorts Iran's Critics to Avoid Threats of Force. *Associated Press.*

Jeong, O. R., Kim, C., Kim, W., & So, J. (2011). Botnets: threats and responses. *International Journal of Web Information Systems*, *7*(1), 6-17.

Jervis, R. (1976). Perception and misperception in world politics. *Princeton, NJ: Princeton Univer-sity Press. JervisPerception and misperception in world politics1976*.

Junio, T. J. (2013). How Probable is Cyber War? Bringing IR Theory Back In to the Cyber Conflict Debate. *Journal of Strategic Studies*, 36:1.

Junio, T. J. (2013). How probable is cyber war? Bringing IR theory back in to the cyber conflict debate. *Journal of Strategic Studies*, *36*(1), 125-133.

Kai, J. (2014). Why China banned Windows 8. http://thediplomat.com/2014/05/why-china-banned-windows-8/

Kaspersky Lab (2012). Gauss: Abnormal Distribution, Kaspersky Lab Technical Report. Available at: https://kasperskycontenthub.com/wp-content/uploads/sites/43/vlpdfs/kaspersky-lab-gauss.pdf

Kello, L. (2013). The meaning of the cyber Revolution, perils to theory and statecraft. International Security, vol. 38, No. 2, pp.7-40.

Keohane, R. O., & Nye Jr, J. S. (1998). Power and interdependence in the information age. *Foreign affairs*, 81-94.

Kerr, P. (2003). IAEA presses Iran to comply with nuclear safeguards. *Arms Control Today*, *33*(6), 20.

Kesan, J. P., & Hayes, C. M. (2011). Mitigative counterstriking: Self-defense and deterrence in cyberspace. *Harv. JL & Tech.*, *25*, 429.

Kim, J. (2014). North Korea Blames US for Internet Outages, Calls Obama “Monkey”.

Kimball, D., & Crail, P. (2012). Chronology of US-North Korean Nuclear and Missile Diplomacy. *Arms Control Association*.

Kiravuo T. & Särelä M. (2013). The Care and Maintenance of Cyberweapons, in Rantapelkonen, J., & Salminen, M. (eds). The fog of cyber defence. *National Defence University, Publication Series 2, Article Collection no.10.*

Kirwan, G. (Ed.). (2011). *The Psychology of Cyber Crime: Concepts and Principles: Concepts and Principles*. IGI Global.

Klimburg, A. (2011). Mobilising cyber power. *Survival*, *53*(1), 41-60.

Korte, G., & Jackson, D. (2015). Obama sanctions North Korea for movie hacking. *USA Today*.

Kovacs, E. (2016). BlackEnergy Malware Used in Ukraine Power Grid Attacks. *SecurityWeek, www. securityweek. com*, *4*(01).

Kramer, F. D., Starr, S. H., & Wentz, L. K. (2009). *Cyberpower and national security*. Potomac Books, Inc.

Krepinevich, A. F. (1994). Cavalry to computer; the pattern of military revolutions. *The National Interest* 37:30(13).

Krepinevich, A. F. (2010). Why AirSea Battle? *Washington, DC: Center for Strategic and Budgetary Assessments, February*, *19*.

Krepinevich, A. F. (2012). *Cyber Warfare*. Center for Strategic and Budgetary Assessments.

Krickovic, A. (2016). When ties do not bind: the failure of institutional binding in NATO Russia relations. *Contemporary security policy*, *37*(2), 175-199.

Kuehl, D. T. (2009). From cyberspace to cyberpower: Defining the problem. *Cyberpower and national security*, 24-42.

Kurtz, G. (2010). Operation "Aurora" hit Google, others. Retrieved from: *http://siblog. mcafee. com/cto/operation-% E2*, *80*.

Kushner, D. (2013). The real story of stuxnet. *ieee Spectrum*, *50*(3), 48-53.

Langner, R. (2011). Stuxnet: Dissecting a cyberwarfare weapon. *IEEE Security & Privacy*, *9*(3), 49-51.

Langner, R. (2013). To kill a centrifuge: A technical analysis of what stuxnet's creators tried to achieve. Available from: http://www.langner.com/en/wp-content/uploads/2013/11/To-kill-a-centrifuge.pdf.

Laughland, O. (2015). FBI Director Stands by Claim that North Korea Was Source of Sony Cyber-Attack. *The Guardian*.

Laughland, O., & Rushe, D. (2014). Sony pulling The Interview was 'a mistake' says Obama. *The Guardian, 20.*

Leccisotti, F. Z., Chiesa, R., & De Nicolo, D. (2016). Analysis of possible future global scenarios in the field of cyber warfare: National cyber defense and cyber attack capabilities. In *Handbook of research on civil society and national security in the era of cyber warfare*. IGI Global.

Lee, R. M., Assante, M. J., & Conway, T. (2016). Analysis of the cyber attack on the Ukrainian power grid. *SANS Industrial Control Systems*.

Lee, R. M., Assante, M. J., & Conway, T. (2016). Analysis of the cyber attack on the Ukrainian power grid. *SANS Industrial Control Systems*.

Lee, T. B. (2014). The Sony hack: how it happened, who is responsible, and what we've learned. *Vox.*

Leiner, B. M., Cerf, V. G., Clark, D. D., Kahn, R. E., Kleinrock, L., Lynch, D. C., ... & Wolff, S. (2009). A brief history of the Internet. *ACM SIGCOMM Computer Communication Review*, *39*(5), 22-31.

Lenoir F. (2016). NATO may react to future cyber attacks by deploying conventional weapons. *Reuters.* Available from: http://in.reuters.com/article/cyber-nato-idINKCN0Z203R

Levy, J. S. (2008). Case studies: Types, designs, and logics of inference. *Conflict Management and Peace Science*, *25*(1), 1-18.

Lewis, J. (2013). *Hidden Arena: Cyber Competition and Conflict in Indo-Pacific Asia*. Center for Strategic and International Studies.

Leyden, J. (2012). "Hack on Saudi Aramco hit 30,000 workstations", oil firm admits. *The Register*, *29*.

Libicki, M. C. (2007). *Conquest in cyberspace: national security and information warfare*. Cambridge University Press.

Libicki, M. C. (2012). *Crisis and escalation in cyberspace*. Rand Corporation.

Liff, A. (2012) Cyberwar: A New ‘‘Absolute Weapon’? The Proliferation of Cyberwarfare Capabilities and Interstate War. *Journal of Strategic Studies,* 35:3

Liff, A. P. (2012). Cyberwar: a new ‘absolute weapon’? The proliferation of cyberwarfare capabilities and interstate war. *Journal of Strategic Studies*, *35*(3), 401-428.

Lin, H. S. (2010). Offensive cyber operations and the use of force. *J. Nat'l Sec. L. & Pol'y*, *4*, 63.

Lin, H. S. (2012). Escalation dynamics and conflict termination in cyberspace. *Air University*.

Lindsay, J. R. (2013). Stuxnet and the limits of cyber warfare. *Security Studies*, *22*(3), 365-404.

Lipschutz, R. D. (1992). Reconstructing world politics: the emergence of global civil society. *Millennium*, *21*(3), 389-420.

Locatelli, A. (2015). La reazione in legittima difesa di uno Stato a fronte di un attacco cyber. *CeMiSS.*

Lonsdale, D. J. (1999). Information power: strategy, geopolitics, and the fifth dimension. *The Journal of Strategic Studies*, *22*(2-3), 137-157.

Lynn, W. J. (2010). Defending a new domain: the Pentagon's cyberstrategy. *Foreign Affairs*, *89*(5), 97-108.

Lynn, W. J. (2010). Defending a new domain: the Pentagon's cyberstrategy. *Foreign Affairs*, *89*(5), 97-108.

Maher, R. (2012). The Covert War against Iran’s Nuclear Program: An effective counterproliferation strategy?. *European University Institute Working Papers.*

Mandiant, APT (2013). Exposing one of China’s cyber espionage units. Available at: intelreport.mandiant.com/Mandiant_APT1_Report.pdf.

Manzo, V. (2011). *Deterrence and Escalation in Cross-domain Operations: Where Do Space and Cyberspace Fit?* Strategic Forum, National Defense University Publication.

Marczak, B., Scott-Railton, J., Senft, A., Poetranto, I., McKune S. (2015). Pay No Attention to the Server Behind the Proxy: Mapping FinFisher's Continuing Proliferation. *Citizenlab.org.* Available from: https://citizenlab.org/2015/10/mapping-finfishers-continuing-proliferation/

Markoff, J. (2008). Georgia takes a beating in the cyberwar with Russia. *The New York Times*, *11*.

Markoff, J. (2009). Worm infects millions of computers worldwide. *The New York Times*, *23*.

Mathuis, C., Pieters, W., van den Berg, J. (2016). Cyber Weapons: a Profiling Framework. *In Cyber Conflict (CyCon U.S.), International Conference on.* IEEE.

Matrosov, A., Rodionov, E., Harley, D., & Malcho, J. (2010). Stuxnet under the microscope. *ESET LLC (September 2010)*.

Mazanec, B. M. (2009). The art of (cyber) war. *Journal of International Security Affairs*, *16*, 84.

McAllister, N. (2015) Bruce Schneier: "We're in Early Years of a Cyber Arms Race"'. *The Register*.

McCarthy, C. J. (2010). *Chinese Anti-Access/Area Denial: The Evolution of Warfare in the Western Pacific*. Naval war college newport, joint military operations dept.

McCurry, J., & Carroll, R. (2014). North Korea refuses to deny role in Sony cyber-attack. *The Guardian, 2.*

McGraw, G. (2013). Cyber war is inevitable (unless we build security in). *Journal of Strategic Studies*, *36*(1), 109-119.

Metz, T. F., Garrett, M. W., Hutton, J. E., & Bush, T. W. (2006). *Massing effects in the information domain: A case study in aggressive information operations*. Army Training and Doctrine Command.

Michaels, J. (2007) NATO to Study Defense against Cyber attacks. *USA Today*.

Molander, R. C., Riddile, A., Wilson, P. A., & Williamson, S. (1996). *Strategic information warfare: A new face of war*. Rand Corporation.

Morgan, F. E., Mueller, K. P., Medeiros, E. S., Pollpeter, K. L., & Cliff, R. (2008). *Dangerous Thresholds: Managing Escalation in the 21st Century*. Rand Corporation.

Mulvenon, J. (2009). PLA computer network operations: Scenarios, doctrine, organizations, and capability. *Beyond the strait: PLA missions other than Taiwan*, 257-259.

Nakashima, E. (2015). Why the Sony hack drew an unprecedented US response against North Korea. *The Washington Post*.

Nakashima, E. (2015). Why the Sony hack drew an unprecedented US response against North Korea. *The Washington Post*.

NATO (2016). *On Deterrence.* NATO Review Magazine. Available at: http://www.nato.int/docu/Review/2016/Also-in-2016/nato-deterrence-defence-alliance/EN/index.htm

NATO (2017) Cyber Defence. Available from http://www.nato.int/cps/en/natohq/topics_78170.htm

Nye Jr, S. J. (1997). Understanding International Conflicts: An Introduction to Theory and Conflicts.

Nye, J. (2010). Cyber Power. Harvard Kennedy School. *Belfer Center for Science and International Relations*.

Nye, J. S. (2011). *Nuclear lessons for cyber security*. Air University Press.

Nye, J. S. (2017). Deterrence and Dissuasion in Cyberspace. *International Security*, *41*(3), 44-71

O'Hare, R. (2016). China proudly debuts its new stealth jet it built 'by hacking into US computers and stealing plans'. *The Daily Mail.*

Oduntan, G. (2011). *Sovereignty and Jurisdiction in Airspace and Outer Space: Legal Criteria for Spatial Delimitation*. Routledge.

OSCE (2013). Decision No. 1106 Initial Set of OSCE Confidence-Building Measures to Reduce the Risks of Conflicts Stemming from the Use of Inormation and Communication Technologies.

Osnos, E. (2012). Letter from China. *The God of Gamblers: why Las Vegas is Moving to Macau. The New Yorker. April*, *9*.

Oxford English Dictionary Online (2017). Cyberspace definition. Available at: https://en.oxforddictionaries.com/definition/cyberspace

Paganini, P. (2015). $5 Billion in Military Cyber Spending fivefold increase over last year, *Security Affairs.*

Pagilery, J. (2014). 'Sony-pocalypse': Why the Sony Hack is one of the Worst Hacks Ever. *CNN Money*.

Perlroth, N. (2014). Cyberespionage Attacks Tied to Hackers in Iran. *The New York Times, 29.*

Peterson, S. (2001). Imminent Iran Nuclear Threat? A Timeline of Warnings Since 1979. *The Christian Science Monitor*, *8*.

Phillips, A. (2012). The Asymmetric Nature of Cyber Warfare. *US Naval Institute News, October*, *14*.

Potegal, M., & Knutson, J. F. (Eds.). (2013). *The dynamics of aggression: Biological and social processes in dyads and groups*. Psychology Press.

PwC (2017). *Global State of Information Security Survey 2017.* Available at: https://www.pwc.com/gx/en/issues/cyber-security/information-security-survey.html

Raghavan, S. V., & Dawson, E. (Eds.). (2011). *An investigation into the detection and mitigation of denial of service (dos) attacks: critical information infrastructure protection*. Springer Science & Business Media.

Rapoza, K. (2013). US hacked China universities, mobile phones, Snowden tells China Press. *Forbes.*

Reed, T. C. (2005). *At the abyss: an insider's history of the Cold War*. Presidio Press.

Rid, T. (2012). Cyber war will not take place. *Journal of strategic studies*, *35*(1), 5-32.

Rid, T. (2013). *Cyber war will not take place*. Oxford University Press, USA.

Rid, T. (2016). *Rise of the Machines: A Cybernetic History*. WW Norton & Company.

Rid, T., & Buchanan, B. (2015). Attributing cyber attacks. *Journal of Strategic Studies*, *38*(1-2), 4-37.

Rid, T., & McBurney, P. (2012). Cyber-weapons. *the RUSI Journal*, *157*(1), 6-13.

Rid, T., & McBurney, P. (2012). Cyber-weapons. *the RUSI Journal*, *157*(1), 6-13.

Robb, D. (2014). Sony hack: A timeline. *Deadline, last updated December*, *24*.

Rogin, J. (2007). Cyber Officials: Chinese Hackers Attack ‘Anything and Everything,’. *FCW. com, February*, *13*, 97658-1.

Rollins, J. W. (2015). *U.S.–China Cyber Agreement.* Congressional Research Service Insight.

Sahay, A., & Roshandel, J. (2010). The Iran–Pakistan–India natural gas pipeline: implications and challenges for regional security. *Strategic Analysis*, *34*(1), 74-92.

Saltzman, I. (2013). Cyber Posturing and the Offense-Defense Balance. *Contemporary Security Policy*, 34(1), 40-63.

Sanger, D. E. (2014). Iran Hackers Dangle a Familiar Name to Fish for Data. *The New York Times, 30.*

Sanger, D. E. (2015) U.S. and China Seek Arms Deal for Cyberspace, *the New York Times.*

Sanger, D. E., & Fackler, M. (2015). Nsa breached north korean networks before sony attack, officials say. *New York Times*, *8*, 72-76.

Sanger, D. E., & Perlroth, N. (2014). NSA breached Chinese servers seen as security threat. *New York Times*, *22*.

Sanger, D. E., Barboza, D., & Perlroth, N. (2013). Chinese army unit is seen as tied to hacking against US. *The New York Times*.

Sanger, D. E., Schmidt, M. S., & Perlroth, N. (2014). Obama vows a response to cyber attack on Sony. *New York Times*, *19*.

Sanger, D.E. (2016). Biden Hints at U.S. Response to Russia for Cyber attacks. *New York Times.*

Sasanapuri, C., KC, C. H., Ch, S., & Challa, N. (2016). Classification of APT’s and Methodological Approach to Secure Cloud Services. *International Journal of Applied Engineering Research*, *11*(2), 1000-1005.

Sasso, B. (2013). Report: China Hacked Obama, McCain Campaigns in 2008. *The Hill*.

Schelling, T. C. (1960). The strategy of conflict. *Cambridge, Mass*.

Schelling, T. C. (1966). Arms and influence. *New Haven: Yale*.

Schmidt, M. S., & Sanger, D. E. (2014). 5 in China Army Face US Charges of Cyber attacks.

Schmitt, M. N. (2013). *Tallinn manual on the international law applicable to cyber warfare*. Cambridge University Press.

Schmitt, M. N. (Ed.). (2017). *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations*. Cambridge University Press.

Schwirtz, M. (2008). Claims of Secret Arms Sales Rattle Ukraine's Leaders. *The New York Times, 29.*

Serapiglia, A. (2016). The Case for Inclusion of Competitive Teams in Security Education. *Information Systems Education Journal*, *14*(5), 25.

Sevastopulo, D. (2007). Chinese hacked into Pentagon. *FT.com*.

Shackelford, S. J., & Andres, R. B. (2010). State Responsibility for Cyber Attacks: Competing Standards for a Growing Problem. *Geo. J. Int'l L.*, *42*, 971.

Shane, S., Rosenberg, M., & Lehren, A. W. (2017). WikiLeaks Releases Trove of Alleged CIA Hacking Documents. *NYTimes. com*.

Shankdhar, P. (2013). DOS Attacks and Free DOS Attacking Tools. *Infosec Institute*, *29*.

Sherr, I. (2017). WannaCry ransomware: everything you need to know. *Cnet*.

Sinha, S., & Beachy, S. C. (2014). Timeline on Iran's Nuclear Program. *The New York Times. The New York Times*, *19*.

Slaughter, A. M., Tulumello, A. S., & Wood, S. (1998). International law and international relations theory: A new generation of interdisciplinary scholarship. *American Journal of International Law*, *92*(3), 367-397.

Snyder, G. H. (1959). *Deterrence by denial and punishment*. Woodrow Wilson school of Public and International Affairs, Center of International Studies, Princeton University.

Snyder, G. H., & Diesing, P. (1977). *Conflict Among Nations: Bargaing, Decision Making, and Systems Structure in International Crises*. Princeton University Press.

Springer, P. J. (2015). *Cyber Warfare: A Reference Handbook*. ABC-CLIO.

Squassoni, S. (2005, May). Iran's nuclear program: Recent developments. Library Of Congress Washington Dc Congressional Research Service.

Stone, J. (2013). Cyber war will take place!. *Journal of Strategic Studies*, *36*(1), 101-108.

Strohm, C. (2015). North Korea Web Outage Response to Sony Hack, Lawmaker Says. *Cyber-security, Bloomberg*.

Suiche, M. (2017). Petya.2017 is a wiper not a ransomware. Comae Technologies. Retrieved from: https://blog.comae.io/petya-2017-is-a-wiper-not-a-ransomware-9ea1d8961d3b

Sweijs, T., Ursanov, A., Rutten, R., de Spiegeleire, S., Bekkers, F., Ward, S. M., ... & Skinner, C. (2016). *Back to the brink: Escalation and Interstate Crisis*. The Hague Centre for Strategic Studies.

Symantec Corporation (2017). *Internet Security Threat Report (ISTR) Volume 22*. USA: Symantec Corporation. Available at: https://www.symantec.com/security-center/threat-report

Tadjdeh, Y. (2015). NSA Chief: China, Russia Capable of Carrying Out 'Cyber Pearl Harbor" Attack. *National Defense Magazine*

Thornburgh, N. (2005). The invasion of the Chinese cyberspies. *Time.* Available at: http://content.time.com/time/magazine/article/0,9171,1098961,00.html

Thornhill, T. (2016). Chinese hacker, 51, jailed and fined $10,000 for stealing data from US defense contractors and passing it to Beijing. *The Daily Mail.*

Tiezzi, S. (2014). Taiwan Complains of 'Severe' Cyber Attacks From China. *The Diplomat.*

Traynor, I. (2007). Russia accused of unleashing cyberwar to disable Estonia. *The Guardian*, *17*(05).

Trend Micro (2017). *THE NEXT TIER: Trend Micro Security Predictions for 2017.* Available at: https://resources.trendmicro.com/rs/945-CXD-062/images/Report%20-%202017%20Predictions.pdf

U.S. Air Force Doctrine Center (2003) *Strategic Attack*, Air Force document 2-1.2.

U.S. Department of Defense (2015). The DOD cyber strategy. *Department of Defense: Washington, DC.*

U.S. Department of Defense (2017). DOD Dictionary of Military and Associated Terms. Available at: http://www.dtic.mil/doctrine/dod_dictionary/

U.S. Department of Homeland Security (2017). National Initiative for Cybersecurity Career and Studies' Glossary. Available at: https://niccs.us-cert.gov/glossary#C

United Nations (1945). *Charter of the United Nations*, 1 UNTS XVI.

Van Der Walt, C. (2017). The impact of nation-state hacking on commercial cyber-security. *Computer Fraud & Security*, *2017*(4), 5-10.

Verma, S. K. (2007). Energy geopolitics and Iran–Pakistan–India gas pipeline. *Energy Policy*, *35*(6), 3280-3301.

Wang, C., Fang, L., & Dai, Y. (2010). A simulation environment for SCADA security analysis and assessment. In *Measuring Technology and Mechatronics Automation (ICMTMA), 2010 International Conference on* (Vol. 1, pp. 342-347). IEEE.

Wendt, A. (1992). Anarchy is what states make of it: the social construction of power politics. *International organization*, *46*(2), 391-425.

Wertz, D., & Gannon, C. (2015). A History of US-DPRK Relations. *Issue Brief. Np: The National Committee on North Korea. Accessed December*, *12*, 2016.

Wess Mitchell, A. (2015) The Case of Deterrence by Denial. *The American Interest.* Available at: https://www.the-american-interest.com/2015/08/12/the-case-for-deterrence-by-denial/

White House (2013) Presidential Policy Directive - Critical Infrastructure Security and Resilience (PPD-21)

Wilson, C. (2015). Cyber weapons: 4 defining characteristics. *CGN.com.*

Zetter, K. (2010). Google hack attack was ultra sophisticated, new details show. *Wired Magazine*, *14*, 33-36.

Zetter, K. (2014). *Countdown to Zero Day: Stuxnet and the Launch of the World's First Digital Weapon*. Broadway books.