Behind the Mask – Attribution of antagonists in cyberspace and its implications on international conflicts and security issues

Urban Bilstrup, Frida Stranne, and Lena Ewertsson

Halmstad University, Sweden

Abstract: *Cyber systems and critical infrastructure are changing the dynamics of international conflicts, security issues, and challenge traditional ways of understanding warfare. Early warning and attribution of who is accountable for a cyber-attack and what is the intention with the attack is crucial information. To be able to efficiently response to a cyber-antagonist the measure of response must be decided at network speed, which is far beyond what is possible with traditional attribution methods. The ongoing "cyber arm raze" push towards the development and use of autonomous cyber response systems. An autonomous cyber response would most probably use the complexity of attack vector as a tool for attribution, not considering the identity of the antagonist for deciding the measure of response. This will challenge traditional ways of understanding conflict, war, and how nation states handle different kinds of aggressions. This leads to a new kind of deterrence increasing the need to theorize cyber conflicts, as well as empirically study how different actors are acting and reacting in relation to this new threat. This paper initiates the discourse on the implications of the use of autonomous cyber response systems for the international system/relations.*

Introduction

The emerging 'digital society' is built around connectivity and a commonly referenced roadmap[1] predicts an increase from today's approx. 15 billion connected devices (CISCO counted 13 billion 2014) to 50 billion connected devices in year 2020[2] and this development creates an estimated 10 trillion U.S. Dollar market for companies and industry worldwide[3]. This super Internet is referred to as Internet of Things (IoT) and is considered as "the first real evolution of Internet – a leap that will lead to revolutionary applications that have the potential to dramatically improve the way people live, learn, work and entertain themselves"[4]. The IoT revolution of transforming Internet into a sensory system (with sensors as: temperature, pressure, vibration, light, moisture, stress etc.) allowing us to become more proactive and less reactive" This evolve Internet from operating as purely a cyber-system to cyber physical systems (CPS), i.e. connects the physical domain to the cyber domain. The pervasiveness of embedded systems, smart objects and IoT is changing everything, providing a never before seen "ability to gather, analyze and distribute data"[5] about everything. The total amount of data is increasingly exponentially and data is not only generated by human activities in cyberspace it is also gathered by: sensors in smart mobile phones, smart TVs, industrial systems, aerial sensory technologies, software logs, CCTV, cameras, microphones, radio-frequency

[1] Ericsson Consumer lab, "PRIVACY, SECURITY and SAFETY ONLINE – Consumer perspective and behaviour", *An Ericsson Consumer Insight Summary Report*, February 2014.

[2] MORE THAN 50 BILLION CONNECTED DEVICES – Taking connected devices to mass market, *ERICSSON WHITE PAPER*, February 2010.

[3] J. Bradely, J. Barbier and D. Handler, "Embracing the Internet of Everything To capture Your Share of $14.4 Trillion," *CISCO White paper*, CISCO 2013.

[4] Dave Evans, The Internet of Things – How the Next Evolution of the Internet Is Changing Everything, *White paper by Cisco Internet Solutions Group*, April 2010.

[5] Ibid.

identification (RFID) readers used in logistic systems, and wireless sensor networks. The worlds stored digital information per capita has roughly doubled every 40 months since the 1980s; as of 2012, every day 2.5 exabytes (2.5×10^{18} bytes) of data were created[6].

The society is increasingly becoming dependent on digital technologies and builds its entire critical infrastructure on these new technologies. These new technologies provide completely new opportunities for optimizing control, distribution systems and logistic systems. A long ongoing trend has been that control systems of critical infrastructure "that once was inaccessible to persons off-premises"[7] become "theoretically accessible for anyone in the world."[8] A direct consequence of this development is that "[t]he more cyberspace is critical to a nation's economy and defense, the more attractive to enemies is the prospect of crippling either or both via attacks on or through it"[9]. To be able to respond to these new cyber-threats it requires as well defensive as offensive cyber response capabilities and the ability to conduct these responses at network speed. Handling this spatially and with high temporal speed as well as independent of the traffic dynamics of an advanced cyber-attack goes far beyond what a human operator can handle. It is clear that if response action should be taken in real-time an automated detection and response system is required, which clearly introduce challenging technical, ethical and responsibility questions.

A response of a cyber-attack follows the so called incident response cycle (detect, analyze, respond, resolve). In order to be able to perform any response the defender first need to detect that he is under attack, then attribute the perpetrator and analyze the purpose of the cyber-attack, i.e. answering the questions: who, what, where, when, why, and how. Early warning and the attribution of *who* is accountable for a cyber-attack and *what* is the intention with the attack are considered the very important information, since any kind of retaliation requires knowing who the attacker is and what the purpose of the attack is. This is consequence of that cyberspace is a multi-jurisdictional domain, the organization or agency that is responsible to handle the retaliation of the "perpetrators" of an attack is dependent upon who is behind and what the intent of the attack is. To answer these two questions one first has develop early warning capability in cyberspace, which is a challenge since things happens at network speed, the cyber domain a super-tactical domain where decisions has to be made in less than milliseconds to be able to repel an attack. The paper will after going through the (known) present state of capabilities and challenges with an autonomous cyber response system and then conclude about some possible consequences. As this paper is focused on who is responsible for an attack from technical perspective it will be limited to the two firs steps in the response cycle: detect and analyze.

Intrusion Detection

Intrusion detection and early warning of an attack is of course crucial in order to response to an incoming cyber-attack. It is not obvious that it easy to know or understand when one is exposed for a cyber-attack, it depends on the nature of the attack and what the goal is of the attack. A breach intended to extract information, in "stealth mode", can be nearly impossible to detect while an

---

[6] M. Hilbert, and P. López, "The World's Technological Capacity to Store, Communicate, and Compute Information". *Science* **332** (6025), 2011, pp. 60–65.

[7] M. C. Libicki, *CONQUEST IN CYBERSPACE – National Security and Information Warfare*, Cambridge University Press, 2007.

[8] Ibid.

[9] Ibid.

attack where the goal is to take down some Internet services is hard to miss. Currently, network Intrusion Detection Systems, software security patches, and vulnerability scanners are all forms of signature based defense: defensive systems which act on discrete quanta of human knowledge ("signatures"). Human analysts develop these signatures through a process of reasoning about software. In fully autonomous defense, "a cyber-system capable of reasoning about software will create its own knowledge, autonomously emitting and using knowledge quanta such as vulnerability scanner signatures, intrusion detection signatures, and security patches."[10]

As an example of ongoing work in that direction an artificial cognitive architecture based on brain emotional learning[11] dedicated for anomaly detection is considered for Intrusion detection and early warning of cyber-attacks. Previous proposed artificial cognitive architectures are often based on a cognitive cycle that is defined according to the rational reasoning system. A rational goal-oriented decision is made by an intelligent agent while the emotional cycle represents an emotional reaction-oriented action, the system is based on basic emotions like: fear, anger, happiness etc. and the (learning) conditioning of a stimulus. J. Mitola defined the term cognitive cycle[12] (see Figure 1) to describe the intelligent behavior of an artificial cognitive system through five iterative steps: observe, orient, plan, decide and act. Mitola's cognitive cycle has been developed on the basis of the OODA loop[13] that uses Unified Theories of Cognition (UTC) to model rational decision making in human.

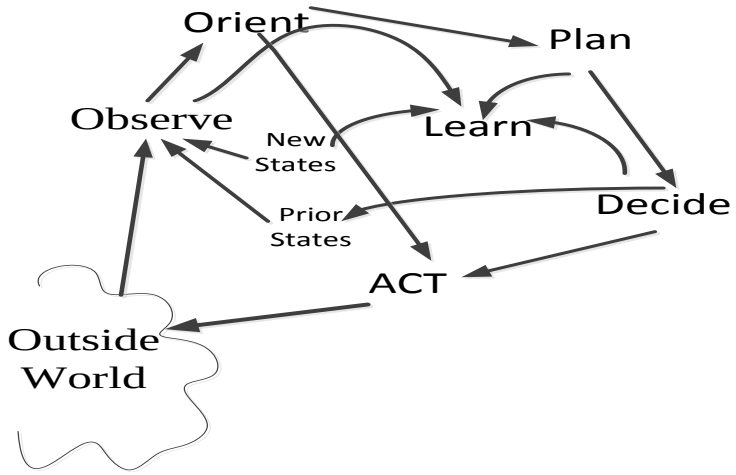

Figure 1 A cognitive cycle based on the rational system

In earlier work an emotional cycle have been developed to represent the functions of the reasoning and learning, assuming environmental changes elicits emotional stimuli that lead to emotional reactions. The emotional cycle imitates the brain's pathway from an emotionally charged stimulus (e.g. fearful stimulus) to an emotionally response (e.g. freezing, run etc.) and it consists of three steps: sensing, learning and acting (see Figure 2). The emotional cycle and Mitola's cognitive cycle differs, especially on how it implements learning, decision making and optimization. Figure 2 depicts the emotional cycle and how it can be mapped on Mitola's cognitive cycle. As Figure 2 indicates

[10] Cyber Grand Challenge – Rules, version 3, DARPA, November 2014.

[11] U. Bilstrup and M. Parsapoor, "A Framework and Architecture for a Cognitive Engine based on a Computational Model of Human Emotional Learning," in the proceedings of *The Wireless Innovation Forum Europe Conference on Communications Technologies and Software Defined Radio, SDR-WInnComm-Europe 2013*, June 11-13, Munich, Germany, 2013.

[12] J. Mitola, Cognitive Radio: An Integrated Agent Architecture for Software Defined Radio, PhD thesis, Royal institute of Technology (KTH), Sweden 2000.

[13] J. Boyd, A discourse on winning and losing: Patterns of conflict, 1986.

emotional sensing corresponds to: observe, orient and act states of the cognitive cycle; *emotional learning* corresponds to plan, learn and decide states of the cognitive cycle; and *emotional reacting* is corresponding to decide and act states of the cognitive cycle.

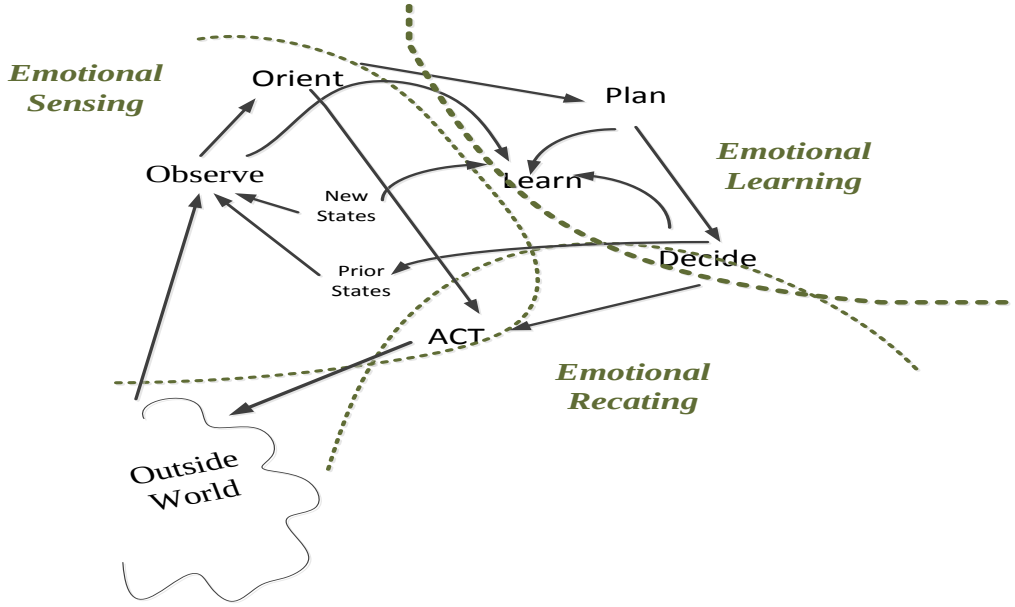

Figure 2 A cognitive cycle based on the emotional system.

A simple structure of an emotion-based computer model, figure 3, has been developed and tested for prediction and classification application[14,15,16].

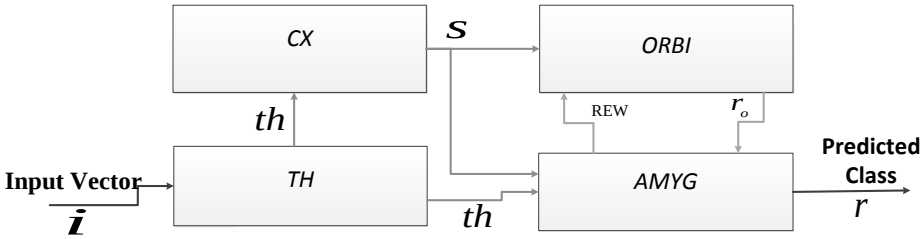

Figure 3. A simple architecture for an emotion-based classifier.

As Figure 3 shows, the architecture consists of four parts: Thalamus (TH), Sensory cortex (CX), Amygdala (AMYG) and Orbitofrontal cortex (ORBI). The CX has a role to select the most informative features and eliminate the redundant features. The AMYG module consists of a classifier and a combiner. The combiner of the AMYG combines the outputs of the AMYG classifier and ORBI classifier to provide the final classification. The combiner strategy depends on the type of classification methods. In the first trials, the weighted k-nearest neighbors (wk-nn) method has been utilized as the classifiers of both AMYG and ORBI. The combiner is also wk-nn and produces the final classification of the input vector, in our case the detection of anomalies in the traffic pattern in a network or system calls in a computer. It should be noted that the classifiers of AMYG and ORBI can be defined on the basis of any supervised classification method, e.g., decision tree, single or

[14] M. Parsapoor, M, U. Bilstrup, "Neuro-fuzzy models, BELRFS and LoLiMoT, for prediction of chaotic time series," in *Proc. of IEEE Int. Conf. INISTA.,* pp. 1-5, 2012.

[15] M. Parsapoor, U. Bilstrup, "Brain Emotional Learning Based Fuzzy Inference System (BELFIS) for Solar Activity Forecasting," in *Proc. IEEE Int. Conf. ICTAI 2012*, 2012.

[16] M. Parsapoor and U. Bilstrup, "An Emotional Learning-inspired Ensemble Classifier (ELiEC)," 8th *International Symposium Advances in Artificial Intelligence and Applications (AAIA'13)*, Kraków, Poland, September 8-11, 2013.

multilayer perceptron, and support vector machine, etc. An ongoing work is to embed these simple classifiers in a network of devices and letting them share information over a peer to peer overlay network, as sketched in figure 4. The hypothesis is that it is possible to set up a distributed anomaly detection networks. The input (input vector ***i*** in figure 3) to the classifier depends on the specific digital device, some inputs are external, the output, ***r***, from other devices' classifier and some are internal parameters for example system call intensity, utilization, communication characteristics etc. As can be seen in figure 4 it can exist feedback loops in the overlay network so that one classifier's output r can be input to another classifier which in turn can be the input to previous classifier.

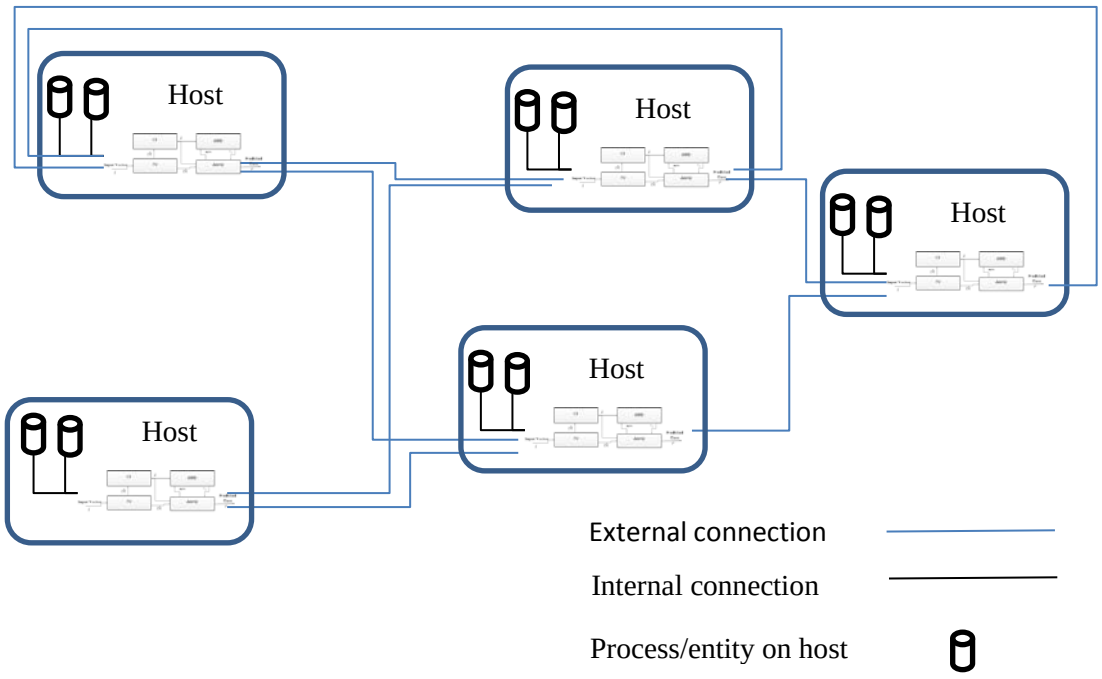

Figure 4. Emotional classifiers interconnected by overlay network.

The idea is that when a classifier learns behavior it also connects that knowledge to devices connected through the overlay network, i.e. a distributed multi agent system. The system forms an autonomous system that condition stimulus (different operational patterns of the system) with emotions, e.g. normal (happy) or malicious (fear) system behavior. Results from this work are very preliminary but some promising result indicates that a distributed emotional learning based intrusion detection system could be more robust than previous systems.

The attribution problem

Attribution refers to the process of identifying the agent responsible for an action, answering the question: who did it? Attribution is fundamental to the idea of deterrence, "the idea that one can dissuade attackers from acting through fear of some retaliation."[17] However, attribution on Internet may have different meaning dependent on the specific context but the goal always falls into one of three general categories [18]: a machine, a person, or aggregated identity (organization, nation etc.). In 2010, previous National Security Agency (NSA) Director Mike McConnel stated that "We need to

[17] Ibid.
[18] Ibid.

develop an early warning system to monitor cyberspace, identify intrusions and locate the source of attacks with a trail of evidence that can support diplomatic, military and legal options – and we must be able to do this in milliseconds"[19] A later testimony done by the previous director of the NSA, Army Gen. Keith B. Alexander, in front of the U.S. senate[20] "we feel confident that foreign leaders believe a devastating attack on the critical infrastructure and population of the United States by cyber means would be correctly traced back to its source and elicit a prompt and proportionate response" indicate that U.S. have or develop the capability of attribution of source and offensive means to hit back in network time. Susanne Spaulding stated at a speech[21] that one have gone from the ability to act in the order of weeks down to few hours, and reducing the attribution time of a cyber-attack and trying to act pre-emptive are the main goal of cyber defense. She also pointed out that an obstacle to further reduce the attribution time was the private sectors unwillingness to share information because a fear of revealing secret business information.

From a practical perspective there exist several kinds of attributions on Internet, table 1, and these can be used in different contexts. In some sense one can say that there exist different levels of attribution on internet, identifying an alias that is a pseudonym of a person to the attribution of a specific person.

Table 1. Forms of attribution[22,23].

Attribution	**Comment**
Alias	an anonymous identity used on internet that links to individual without revealing the actual person are, i.e. acting under pseudonym
Mail address	mail addresses is often not worth much since the can be spoofed
IP address	For many IP addresses one can attribute the physical location directly since IP addresses often are allocated in blocks to Internet service providers, organizations, corporations etc. ISP often also has a billing address to a specific customer.
Physical location	Important since it often direct leads to a persons, organizations, or nations.
Person	it may also require deciding on behalf of who the person was acting, himself, organization or government

Actors and deterrence

Cyberspace is a multi-jurisdictional domain and which jurisdictional organization or agency that is responsible to handle a cyber-attack is dependent upon the intent of the attack. Cyber threats may come from a variety of actors, spanning from individuals to nation-states, hostile cyber activities is often classified into four different main groups: cybercrime, cyber spying, cyber terrorism and cyber war dependent on who is accountable for the attack. On the other hand one can say that there exist six main actors: nations, organized crime, cyber terrorists, companies, hacktivists, and Individual perpetrators.

[19] Mike McConnel, Mike McConnel on How to Win the Cyberwar We're losing," *Washington Post*, February 28, 2010.

[20] K. Alexander, Statement before the Senate Committee on Armed Services, Washington D.C. 12 March 2013, pp. 3.

[21] Susan Spaulding, Under Secretary of National Protection and Programs Directorate, Department of Homeland Security, Brookings Institution, Nov 19th "2013, "The Cybersecurity Executive Order and Presidential Policy Directive: What Does Success Look Like?"

[22] J. Hunker, B. Hutchinson, and J. Marglies, Roles and Challenges for Suficient Cyber-Attack Attribution, Institute of Information Infrastructure Protection, January 2008.

[23] David D. Clark and Susan Landau, "Untangling Attribution," *Procedings of workshop on deterring cyberattacks: Informing Strategies and developing Options for U.S. Policy*, xxx

Nations

Many nations today have cyber forces and are preparing them for future cyber conflicts, to understand the magnitude of these activities one should consider the fact that for example United States Air Force (USAF) established a cyber-division (24th air force) that comprises 6000-8000 cyber warriors[24]. United States Air Force (USAF) changed the their operational definition to include cyberspace, "The mission of the United States Air Force is to deliver sovereign options for the defense of the United States of America and its global interests – to fly and fight in Air, Space and Cyberspace"[25]. Another example is North Korea that often is considered as black spot in the context of Internet, but North Korea have several cyber units, specifically Bureau 121 with a large outpost in China, containing roughly 6,000 people[26]. An example of their capabilities is the recent cyber-attack on Sony, an action taken as a retaliation of the release of the comedy movie "The Interview". The data extracted from the breach, conducted by a group called themselves the "Guardians of Peace (GOP)", included personal information about Sony Pictures employees and their families, e-mails between employees, information about executive salaries at the company, copies of (previously) unreleased Sony films, and other information[27]. As consequence of the breach and threats about further retaliations against the company, Sony decided to cancel the release of the movie. President Barack Obama stated that he thought Sony had "made a mistake. We cannot have a society in which some dictator in some place can start imposing censorship in the United States. I wish they'd spoken to me first. I would have told them: do not get into the pattern in which you are intimidated."[28] North Korea temporarily suffered a nationwide Internet outage a couple of days later, and they accuse the U.S. of orchestrating these Internet outages[29]. Many other countries have similar capacity, some examples of strong actors in the cyber domain are: China Russia, Israel etc. but over 100 nations are estimated to have military related cyber capabilities. It is also clear that some countries conduct advanced industry espionage, during the "Titan rain" an offensive cyber-attack from China somewhere between 10 -20 terra byte of data was breached from Pentagon and different companies connected to the U.S. Defense industry[30]. The director of the National Security Agency (NSA), Army Gen. Keith B. Alexander, called cybercrime "the greatest transfer of wealth in history" [31] . Symantec estimate the cost of intellectual property loss via cyber espionage only to the United States companies is in the range of $250 billion a year[32].

[24] Se 24th AIR FORCE hemsida, http://www.24af.af.mil/main/welcome.asp (150102)
[25] http://www.af.mil/news/story.asp?id=123013440 (120102)
[26] R. A. Clarke och R. K. Knake, Cyber war – The next threat to national security and what to do about it, HarperCollins Publisher, 2010.
[27] D. E. Sanger and N. Perlroth, "U.S. Links North Korea to Sony Hacking," *The New York Times*, December 17, 2014.
[28] D. Dwyer and M. Bruce, "Sony Hacking: President Obama Says Company Made 'Mistake' in Canceling 'The Interview'," ABC News December 19, 2014.
[29] P. Helsel, "North Korea Insults Obama, Blames U.S. For Internet Outages," NBC News, December 29, 2014.
[30] R.A. Clarke och R. K. Knake, Cyber war – The next threat to national security and what to do about it, HarperCollins Publisher, 2010.
[31] Speech by NSA director Keith B Alexander at Cybersecurity and American Power, American Enterprise Institute (AEI), July 2012 available at: http://www.aei.org/events/2012/07/09/cybersecurity-and-american-power/
[32] Symantec

Organized crime

The total number of breaches across public and private networks is increasing[33], McAfee estimate that the global financial loss caused by malicious cyber activity is in the range of 300 billion to 1 trillion U.S. dollars per year[34]. Organized crime activities continue to rises in the yearly Data breach Investigation Report from Verizon Business[35] state that approximately 60 percent of their investigated breach cases lead to organized crime. Furthermore, the report states that most of the investigated breaches have their sources in Eastern Europe. In 88 percent of the 761 investigated cases in 2011 was the source geographically identified: 65% Eastern Europe, 19% North America, 6% south East Asia, 3% Eastern Asia, 2 % Western Europe, and 1% from the rest of the world.

Cyber Terrorists

A straightforward of cyber terrorists is that "cyber terrorists are extremists who do not hesitate to make use of extreme means, such as brutal violence towards the innocent or mass destruction of public property, in pursuit of their political goals or ideological agendas." In the low end one should remember that cyberspace has very "low barriers to entry for malicious cyber activity, including the widespread availability of hacking tools, means that an individual or a small group of determined cyber actors can potentially cause significant damage"[36] and can therefore be considered as one of the most dangerous asymmetric threat against the digitalized society. A concern often raised in policy institutions in Washington, were "guerrilla warfare" and the threat from non-state actors (proxy groups) is pointed out as the threat that can and will cost the U.S. most harm by attacking certain domains in cyber space. The perception is that the U.S. will most likely be attacked by smaller groups[37]. In a Joint Subcommittee Hearing in US Congress: "Cyber Incident Response: Bridging the Gap Between Cybersecurity and Emergency Management" it was expressed that the threat from cyber-attack may become as devastating as 9.11 or worse.[38] Internet has evolved from operating as purely a cyber-system to cyber physical systems (CPS), i.e. connects the physical domain to the cyber domain. The pervasiveness of embedded systems, smart objects and the so called Internet of Things revolution is changing everything, providing a never before seen "ability to gather, analyze and distribute data"[39] about everything. The critical infrastructures that are essential for the functioning of a society are to large extent built upon such technology. New features provided by these new technologies give opportunities for optimizing control, distribution systems and logistics i.e. the connectivity provide means for increased efficiency and competitiveness. A consequence of this long ongoing trend is that control systems of critical infrastructure "that once was inaccessible to persons off-premises"[40] become "theoretically accessible for anyone in the world"[41] implies a significant

[33] Symantec Global Internet Security Threats Reports Volume 15, April 2009.
[34] James Lewis, THE ECONOMIC IMPACT OF CYBERCRIME AND CYBERESPIONAGE, Center for Strategic and International Studies, July 2013.
[35] W.H. Baker et.al., 2011 Data Breach Investigation report, Verizon Buisness, 2011.
[36] Department of Defence Strategy for Operating in Cyberspace, department of defence, United States, July 2011.
[37] See for example a discussion at Atlantic Council, November 15th 2013, Cyber Conflict and War: Yesterday, Today and Tomorrow, Dr. Greg Rattray, Richard Bejtlich, Jason Healey
[38] Joint Subcommittee Hearing, Cannon House Office Building, Oct 30, 2013 10:00am
[39] Ibid.
[40] Martin C. Libicki, *CONQUEST IN CYBERSPACE – National Security and Information Warfare*, Cambridge University Press, 2007.
[41] Ibid.

kinetic cyber threat in the near future"[42] A direct consequence is that "[t]he more cyberspace is critical to a nation's economy and defense, the more attractive to enemies is the prospect of crippling either or both via attacks on or through it"[43] and it would probably be ignorant not to prepare for that terrorists sooner or later will utilize these possibilities for reaching their goals even if it haven't happened so far.

Hacktivists

There exists a growing group of cyber activists, which have more or less well defined political goals with their activities. Many of these groups are a loosely connected network of members that ad hoc pile up around different types of political events. The targets of these activists are often somehow connected to events that a specific group objects against. The most known hacktivist group is Anonymous[44], loosely associated international network of activist and hacktivist with a distributed "command structure that operates on ideas rather than directives" members periodically points out different targets in the world, then it is up to the supporter's if he/she want to take part of that specific action. Anonymous hacktivism have targeted[45]: government agencies of the US, Israel, Tunisia, Uganda, and others; child pornography sites; copyright protection agencies; the Westboro Baptist Church; corporations such as PayPal, MasterCard, Visa, and Sony; and lately also Islamic jihadist sites.

Individual perpetrator

The stereotype for a hacker is a social miss adapted young male who have spent his entire life in front of a computer. But in reality there exist a plethora of different kinds of personalities that are involved in cyber activities. Many of them are driven by curiosity but a growing part is also driven by the possibility to earn money[46]. In the book profiling hacker[47] nine typical perpetrator profiles of "cybercriminal" is given: wannabe lamar, script kids, cracker, ethical hacker, quiet paranoid skilled hacker, cyber warrior, industrial spy, government agent, and military hacker. As summary one can state that the span of age for individuals that are active within this area is much larger than the stereotype picture given by media. The last two groups hide behind government's intelligence services or military organizations and questionable if it can be counted as crimes. Connecting specific perpetrator profiles to actors are not easy to perform. A simplified picture is that the first five categories can be counted as individuals or groups with a common interest while "cyber warriors" and "industrial spies" are individuals that work for organized crime syndicates or conduct specific missions against payment. In for example Russia it can be hard to differentiate between individuals that work for the government and crimes syndicates. Those who are really skilled probably work as freelancing consultants and takes on contract from however that affords to pay. In the classification

[42] S. D. Applegate, 5th International Conference on Cyber Conflict, Tallin Estonia 2013.
[43] Martin C. Libicki, *CONQUEST IN CYBERSPACE – National Security and Information Warfare*, Cambridge University Press, 2007.
[44] Xxx, Anonymous, xxx 20xx.
[45] Wikipedia ….
[46] D. Goldberg och L. Larsson, Svenska hackare: en berättelse från nätets skuggsida, Nordstets, 2011.
[47] R. Chiesa, S. Ciappi and S. Ducci, Profiling Hackers – The science of criminal profiling as applied to the world of hacking, Auerbach Publishing Inc., 2008.

given in profiling hackers[48] non-governmental organizations are lacking, for example cyber terrorists, corporations, and patriot militia.

Deterrence

The nation state has traditionally used two basic strategies to maintain the order they need to survive as society. It "maintain internal order by articulating and enforcing a set of proscriptive rules (criminal law enforcement) that discourage the members of the society from preying upon each other in ways that undermine order, such as by killing, robbing, or committing arson. Societies maintain external order by relaying on military force (war) and to an increasing extent, international agreements."[49] This division between internal and external becomes problematic when considering malicious activities on Internet, since it challenging to define borders on Internet. The physical network that carries data traffic is well defined, but for the logical services it is not as easy to state their exact geographical position. Physical computers (servers) have geographical positions but the software that provide services are often divided between many different servers that can be geographically located all over the world. A proposed solution is that retaliation is connected to the purpose of an attack rather than if the act is internal or external to a state. One idea is just to keep the two entities crime and war, but letting the framework for deciding whether a malicious action on Internet is an act of war or a crime be based on the type of actor and the motivation for the malicious action. The problem with this definition is how to draw the line, what is crime and what is considered as war, can for example a non-governmental actor conduct war? Strictly following the idea of asymmetric threats the answer is probably yes. According to the conclusions about cyber warfare and international law in the "Tallinn Manual" the "sovereignty that a state enjoys over territory … gives it the right to control cyber infrastructure and cyber activities within its territory"[50]. This has two consequences: "[f]irst, that cyber infrastructure is subject to legal and regulatory control by the state. Second, the State's territorial sovereignty protects such cyber infrastructure."[51] It has been stated that cyberspace "is not a physical place it defies in any physical dimension or time space continuum"[52], however "practice gives sufficient evidence that cyberspace, or rather components thereof, is not immune from sovereignty and from the exercise of jurisdiction"[53] It is important to notice is that "[i]t does not matter whether it belongs to the government, or to private entities or individuals, nor do the purpose matter"[54] Furthermore, "a state may exercise jurisdiction over a person engage in cyber activities on its territory, over cyber-infrastructure located on its territory and extraterritorially, in accordance with international law."[55] This gives rice to two jurisdictions subjective territorial jurisdiction and objective territorial jurisdiction, the former include "an incident

[48] R. Chiesa, S. Ciappi and S. Ducci, Profiling Hackers – The science of criminal profiling as applied to the world of hacking, Auerbach Publishing Inc., 2008.
[49]
[50] Michael N. Schmitt Gen. Ed., *Tallinn Manual on the international law applicable to cyber warfare*, Cambridge Press, 2013.
[51] Ibid.
[52] Thomas Wingfield, *The Law of Information Conflict: National Security Law in Cyberspace*, Aegis Research Corp, August 2000.
[53] Wolff Heintschel von Heinegg, "Legal Implications of Territorial Sovereignty in Cyberspace", in Proc. of 4th International Conference on Cyber Conflict 2012, June 2012.
[54] Michael N. Schmitt Gen. Ed., *Tallinn Manual on the international law applicable to cyber warfare*, Cambridge Press, 2013.
[55] Ibid.

that is initiated within its territory but completed elsewhere"[56] and the latter is "jurisdiction over individual to the state where the particular incident has effects even though the act was initiated outside its territory"[57]. Another statement in the Tallinn Manual state that "[a] state shall not knowingly allow the cyber infrastructure located in its territory or under its exclusive governmental control to be used for acts that adversely and unlawfully affects other states"[58]. This rule more or less require a state to have means for protecting its cyber infrastructure against illegal use, which implies some form of governmental cyber intelligence organization as well as ability to stop such conduct. The UN general assembly has also "called upon states to [.....] prevent their territories from being used as safe havens [and] cooperate in the investigation and prosecution of International cyber-attacks" as summarized by David Graham[59] and Jason Healey[60]. Following this reasoning is seems that crime and war is still two entities to base retaliation and deterrence upon. However, retaliation is also based on assessing the rationale for the attack in order to apply the appropriate retaliation. "The damage caused is one of the most distinguishing features of a … [cyber-attack]. The damage of a cyber-attack, in contrast to offence that is involved in physical violence, is almost always exceedingly difficult to pin down and to quantify."[61] According to Western Just war, jus in bello action are those that are discriminate and proportionate i.e. the civil casualties should be kept minimum and the strike should be proportionate with what we try to achieve.

Situation awareness

This distribution of data and computation is an increasing challenge for any forensic investigation of unlawful activity in as well cyberspace as used individual. Setting up a time line for the course of events of a cyber-event becomes utterly complex under such premises. To be able to attribute an actor in a specific case without the direct involvement of Internet operators require very good situational awareness in the cyber domain and it requires advanced technical cyber intelligence gathering systems with global coverage and a cyber-intelligence organization. Only the biggest player in cyberspace have the ability to provide such good cyber intelligence that they can attribute a cyberattack from anywhere in the world. The use of sensor fusion and different kinds of distributed and embedded cyber Intelligence entities fundamentally change the prerequisites for a response system to analyze and handle cyber-attacks at network speed[62]. The fast technology evolution has provided new possibilities for massive surveillance and intelligence gathering. Eavesdropping and "all intelligence" gathering, which previously would have required infinite resources is today technically possible, with this opportunity for massive surveillance the risk for threating the integrity of individual citizens have increased. It is also interesting to note that the largest holders of information are private companies like google, facebook, etc. The infrastructure of cyberspace is to a large extent controlled by private owners and operators that not always apply the highest security practice. In June 2013 Edvard Joseph Snowden leaked secrets documents about how NSA use an intelligence

[56] Ibid.
[57] Ibid.
[58] Ibid.
[59] D. Graham, "Cyber Threats and the Law of War," Journal of National Security Law and Policy, 2010.
[60] J. Healey, "The Spectrum of National Responsibility for Cyberattacks," Brown Journal of World Affairs, Volume XVIII, ISSUE I, Fall/Winter 2011.
[61] T Rid and B. Buch, "Attributing Cyber Attacks," Journal of Startegic Studies, 2014.
[62] "Enabling Distributed Security in Cyberspace", this paper was prepared under the direction of Philip Reltinger Deputy under Secretary for the National Protection and program Directorate (NPPD), March 21, 2011

gathering system called “PRISM” to The Guardian and The Washington Post[63]. PRISM provides NSA direct access to information held by nine of the biggest privately held companies that base their business models on “personal information economy” e.g. Microsoft, Google, Facebook, Skype, PalTalk, Youtube, Yahoo, Apple etc. Snowden also revealed another technical system called Boundless Informant that is a tool for dig data analyses and keeping track of NSAs intelligence gathering resources around the globe. The system keep track of all so called SIGINT activity designators (SIGAD), which are IDs for specific communication links that are monitored by a specific SIGINT site[64]. Boundless Informant can answer important SIGINT questions like “which surveillance sites do we have in a specific region” that provide the necessary technical information for setting up a tracking network enabling triangulation of a specific geographical source, IP address, which is necessary for any level of attribution on Internet.

Autonomous cyber defense system

There are many indications that U.S intelligence agencies are in the process of developing autonomous cyber defense capabilities. Previous NSA, Army Gen. Keith B. Alexander state that “[o] ur most recent running of CYBER FLAG introduced new capabilities to enable dynamic and interactive force-on-force maneuvers at net-speed”[65] which indicated autonomous cyber defense capabilities. The NSA whistleblower Edvard Snowden also state[66] that NSA run a research program called Monstermind, which is stated to “be a cyber-defense system that would instantly and autonomously neutralize foreign cyberattacks against the US, and could even be used to launch retaliatory strikes.” It seems that the Monstermind program is a continuation of the Einstein 1 and Einstein 2 programs, which are known after a publicly available Privacy Impact Assessment (PIA) from Department of Homeland Security (DHS)[67]. The PIAs states that the EINSTEIN programs “developed a computer network intrusion detection system (IDS) used to help protect federal executive agency information technology (IT) enterprises. The EINSTEIN 1 program “analyzes network flow information from different ….. government agencies and provides a high-level perspective from which to observe potential malicious activity in computer network traffic”. The follower, EINSTEIN 2, “incorporate network intrusion detection technology capable of alerting the United States Computer Emergency Readiness Team (US-CERT) to the presence of malicious or potentially harmful computer network activity in federal executive agencies’ network traffic. EINSTEIN 2 principally relies on commercially available intrusion detection capabilities to increase the situational awareness of the US-CERT”[68]. Before returning fire of a cyber-attack, “the US would need to know what it is attacking, and what services or systems rely upon it. Otherwise, it could risk taking out critical civilian infrastructure”[69]. "The DARPA program Plan X intend to quantify cyber effects so the military understands how [such

[63] Wiki http://wikipedia.org/wiki/Edward_Snowden (2013-09-30)
[64] SIGINT Activity Designator, http://en.wikipedia.org/wiki/SIGINT_Activity_Designator
[65] Statement of General Keith B. Alexander COMMANDER UNITED STATES CYBER COMMAND before the Senate Committee On Armed Services 12 March 2013.
[66] Edvard Snowden ….
[67] Pursuant to Section 208 of the E-Government Act of 2002 (Public Law 107-347, 44 U.S.C. § 3501, note), the Department of Homeland Security (DHS) must provide this publicly available PIA prior to initiating a new collection of information that uses information technology to collect, maintain or disseminate information that is in an identifiable form or collects identifiable information through the use of information technology.
[68] Privacy assessment for EINSTAEIN 2, U.S Department for Homeland Security, May 2008.
[69] http://www.defense.gov/news/newsarticle.aspx?id=122455 (visited Jan. 2015)

effects] work and what the collateral damage could be,"[70] the DARPA X program manager Frank Pound told American Forces Press Service during a recent interview. Pounds continuous "We want to make sure when we deploy a cyber-effect at an adversary that there's no collateral damage. Right now, that [capability] really doesn't exist, except in small enclaves."[71]

From a technological perspective it seems that the basic techniques for autonomous cyber defense already exist. Technology for detecting a cyber-attack has been available for a long time by for example different kinds of intrusion detection systems (IDS). The problem with these are that so called knowledge based systems only detect known attacks and so called anomaly detection based systems that are able to detect unknown attacks but it suffer from high false alarm rate. However, as presented, new biologically inspired cognitive architectures can possibly provide fine grained distribution of sensors and advanced data driven classification methods show promising results for detecting attacks with good reliability. Global attribution of a cyber-attack at network speed requires extremely good situation awareness and only the biggest players can have this ability. The technical system called Boundless Informant could possibly provide such triangulation abilities in real-time. Together with massive availability of information from the PRISM system it would probably be possible to build a perpetrator profile to identify who is behind (state, organization etc.) an attack in real-time, the problem here is to identify the search questions.

Following that the details of attribution only need to meet the "needs of the policy makers … and few of these needs relay on perfect attribution"[72] the question is what is the needs in sense of correctness. The intelligence that are used in national security situation are very seldom guaranteed to be 100% correct it is rather estimates. For example for a localization attribution based on IP addresses, which today is a commonly used for web service, "[v]arious firms claims that 99-99.9% of IP addresses can be accurately localized within a country, and that 90-96% can be accurately localized to within a state, city or other similar regions."[73] For national security "the occasions when attribution at a level of a person is useful, are very limited"[74] it is rather the question "who is to blame?"[75] that is more important than the question "who did it?"[76]. This statement implies that attribution of IP address is enough if nations are held responsible for major attacks from their national territory or citizens, which seems like a feasible thought, following the Tallin manual. The localizations attribution is probably better using other tools but here will always be an estimate with some value, the big question what level of guarantees that is necessary. Because if the target is wrongly identified there are a big risk for substantial collateral damage.

Discussion

The idea of autonomous cyber defense, that seems to be the next step in the ongoing digital arms race, expose a number of questions regarding accountability and oversight of entity that controls

[70] Ibid.
[71] Ibid
[72] Ibid.
[73] D. D. Clarke S. Landau, "Untangling Attribution," Proceedings of a workshop on Deterring CyberAttacks: Informing Strategies and Developing Options for U.S. policy
[74] Ibid.
[75] J. Healey, "The Spectrum of National Responsibility for Cyberattacks," Brown Journal of World Affairs, Volume XVIII, ISSUE I, Fall/Winter 2011
[76] Ibid.

these intelligence and weapon system. The biggest challenge is that no real-time decisions can be taken by humans (not even hitting the enter button) if a cyber-defense system should be able to repel a cyber-attack, operates in the super-tactical domain with decision times in range of milliseconds. Traditional questions for these kinds of actions, like[77]: who should be accountable (the government, the legislature, or some independent person or body)? For what (expenditure, policy, and operations)? and when (before carrying out operation or after)?, becomes very hard to answer.

Another important issue to consider is whether cyber-defense is only allowed to be conducted by a governmental organization (military, polis, intelligence) or do any organization or individual have the right to protect its assets and in that case with what means, are we as individual allowed to conduct digital self-defense. Even if the required situation awareness limits the type of organizations that have enough cyber intelligence capabilities one should be aware of that most of the infrastructure and content on Internet is owned and held by the private sector. The vast knowledge and capital in this area suggest a strong involvement from private sector in future cyber defense system. There are strong incitements for that we will see privately held cyber forces, similar to contractors as Black Water represent in the physical domain warfare. An example of this trend is that previous chief of Israel's cyberspy program 8200 starts a venture incubator[78], Team8, for security startups together with Google executive chairman Eric Schmidt's Innovation Endeavors, Bessemer Venture Partners, and Marker LLC, along with input from Cisco and Alcatel-Lucent. The goal of Team8 is "To leverage the offensive and defensive skills of veterans of Israel's cyberwar efforts to build new security startups, which the company describes as "disruptive."

[77] Ian Leigh, More Closely Watching the Spies: Three Decades of Experience, In Who's Watching the Spies? Establishing Intelligence Service Accountability, edited by H. Born, L. K. Johnson, and I. Leigh, Potomac books, 2005.

[78] N. Ungerleider, Spymaster launched a new cybersecurity incubator, http://www.fastcompany.com/3042158/googles-eric-schmidt-cisco-israeli-spymaster-behind-new-cybersecurity-incubator-team8-ventur