Center for International and Security Studies at Maryland
School of Public Policy, University of Maryland

On Critical Infrastructure Protection and International Agreements

By Nicolas Christin

CISSM Working Paper

March 2011

This paper was prepared as part of the Advanced Methods of Cooperative Security Program, with generous support from the John D. and Catherine T. MacArthur Foundation and the Yamamoto-Scheffelin Endowment for Policy Research.

Center for International and Security Studies at Maryland
4113 Van Munching Hall, School of Public Policy
University of Maryland
College Park, MD 20742
(301) 405-7601

UNIVERSITY OF MARYLAND

On Critical Infrastructure Protection and International Agreements

Nicolas Christin
Carnegie Mellon University CyLab
4720 Forbes Ave., Rm. 2108
Pittsburgh, PA 15213, USA
nicolasc@cmu.edu

March 31, 2011

Abstract

This paper evaluates the prospects for protecting critical social functions from "cyber" attacks carried out over electronic information networks. In particular, it focuses on the feasibility of devising international laws, conventions or agreements to deter and/or punish perpetrators of such attacks. First, it briefly summarizes existing conventions and laws, and explains to which technological issues they can apply. The paper then turns to a technical discussion of the threats faced by critical infrastructure. By distinguishing between the different types of attacks (theft of information, destructive penetration, denial of service, etc.) that can be conducted, and examining the role of collateral damages in information security, the paper identifies the major challenges in devising and implementing international conventions for critical infrastructure protection. It then turns to a practical examination of how these findings apply to specific instances of critical networks (power grids and water systems, financial infrastructure, air traffic control and hospital networks), and draws conclusions about potential remedies. A notable finding is that critical functions should be isolated from non-critical functions in the network to have a chance to implement viable international agreements; and that, given the difficulty in performing attack attribution, other relevant laws should be designed with the objective of reducing negative externalities that facilitate such attacks.

1 Introduction

Critical infrastructure systems such as power plants, hospitals, air traffic control systems, to name a few, nowadays all rely on computing devices. Thus, a large portion of our infrastructure becomes potentially vulnerable to "cyber"-attacks. Rather than trying to directly harm the physical system ultimately targeted, in a cyber-attack, the attacker instead attempts to compromise the underlying computing devices to either indirectly seize control over the system, or acquire the ability to disrupt system operations.

Given that computing systems are increasingly federated in interconnected networks, attackers further acquire the ability to carry out cyber-attacks remotely, without ever physically accessing any of the targeted infrastructure. Examples of computer viruses or worms, such as Slammer [35] or Stuxnet [23], that have significantly disrupted critical functions indicate that the threat of cyber-attacks is not only real but that the consequences of carefully orchestrated cyber-attacks by hostile parties can be dire.

Damages on critical infrastructure can further be caused by collateral effects of cyber-attacks. The Slammer worm, for instance, was not trying to specifically incapacitate critical functions; instead it was built to replicate itself as fast as possible over as many Internet hosts as possible. The very fast rate of propagation of the worm led to a surge in Internet traffic, which in turn caused packet routers to be unable to cope with the demand, and to discard some routing messages. Internet routes, which are essential to ensure network connectivity, then became unstable and parts of the network became unreachable [35]. Critical infrastructure systems that relied on Internet availability to ensure continuity of their operations ended up being impeded by the Slammer worm.

More recently, information security experts have witnessed the emergence of malicious programs seemingly specifically designed to target critical infrastructure. The Stuxnet worm [23], in particular, has received considerable media coverage as it is targeting industrial control systems by altering the code present on programmable logic controllers that govern these industrial control systems. The sophistication in the design of the Stuxnet worm indicates that its designers have had access to considerable research and development capabilities, which in turn suggests very powerful adversaries – possibly with available resources at the level of that of a nation-state.

Numerous technological remedies, ranging from network intrusion detection systems (see [8] for a survey), to formal logic for access control (see, e.g., [2, 3] for an overview of works in the field) have been proposed in the academic and industrial research literature. While a number of these research advances have greatly contributed in making computing systems considerably more secure than they used to be, attackers and defenders are seemingly engaged in a technological arms race. The complexity of the computing systems involved indeed makes it possible for novel attack vectors to be continuously discovered, and defenses have to be found almost immediately. As such, it is important to determine whether technological remedies can be complemented by other, non-technological, advances. For instance, building strong economic disincentives for attackers to carry out their misdeeds [39], or devising and implementing legal prohibitions, may usefully supplement technical protection.

This paper follows on this observation by exploring whether international laws, conventions or agree-

ments, could be drafted to dissuade and/or punish perpetrators of cyber-attacks targeting critical infrastructures.

To that effect, the paper deliberately takes a technology-oriented approach. By distinguishing the various technical factors that make it possible to carry out cyber-attacks, it identifies the major challenges in implementing international agreements to provide critical infrastructure protection. The paper notably argues that segmentation of networks between critical and non-critical functions must be enforced by technology; and that relevant laws and conventions should be primarily devised with the objective of reducing negative externalities that make security attacks easier. This argumentation is illustrated with a few selected case studies of specific instances of critical networks (power grids and water systems, financial market clearing operations, air traffic control, and hospital networks), before drawing conclusions about the viability of different potential remedies.

The rest of this paper is organized as follows. The paper reviews existing conventions and laws in Section 2, before presenting a technical discussion of the cyber-threats to critical infrastructure in Section 3. It then discusses implications on some of the existing critical networks in Section 4 before turning to a principled discussion of the possible legal remedies in Section 5.

2 Existing Conventions and Laws

Before discussing any further whether (and how) additional conventions or laws could be implemented, this paper first considers existing conventions that have been drafted or have already entered into force to combat cyber-attacks.

Historically, a number of international conventions pertaining to computing activities were derived from international laws applying to different, seemingly unrelated, fields. Perhaps the best known example is that of the successive conventions restricting the export of cryptographic materials (e.g., CoCom, [62] and Wassenaar Arrangement [1]). Rather than defining new laws pertaining to digital assets, the parties to these conventions used existing international agreements on merchandise export and applied them to digital data. One key focus of these agreements is dual-use goods, that is, goods that can be used both in a civilian and a military context. Weapons, as well as cryptography, fall under the "dual-use" umbrella, and as a result, cryptographic material was treated as equivalent to weapons and munitions.

More generally, many areas pertaining to information security are already addressed by existing conventions and laws, and do not necessarily need additional legislation regardless of the (digital vs. physical) medium used. For instance, issues related to the circulation of intellectual property fall mostly outside the scope of this paper – although it does consider the issue of breaches of confidentiality. This includes international patent enforcement, as well as copyright agreements. While there have been numerous arguments that some of the existing conventions may be inappropriate to properly regulate digital media (see, e.g., [15] and the references therein), this paper decidedly avoids debating a body of intellectual property law whose scope considerably exceeds that of cyber-attacks; and instead focuses here on laws and conventions directly enacted in response to cyber-crime.

2.1 Convention on Cybercrime

Perhaps the best-known effort, so far, to produce an international agreement to combat cyber-attacks is the Convention on Cybercrime [19]. The Convention on Cybercrime was written by the Council of Europe and was adopted in Budapest in November 2001. By October 2010, 30 states had signed and ratified the treaty, with another 16 states signing it. The United States signed (in 2001) and ratified (in 2006) the Convention on Cybercrime. The Convention eventually came into force in the United States starting on January 1, 2007. Note that not all signatory states are members of the Council of Europe. In addition to the United States, Japan, Canada and South Africa signed the convention, without, so far, ratifying it.

The Convention on Cybercrime is a rather broad document, whose main objective is, in the words of the convention itself to "pursue [...] a common criminal policy aimed at the protection of society against cybercrime [...] by adopting appropriate legislation and fostering international co-operation." The Convention strives at defining policy in the face of four different types of offenses.

First, under Title 1 (Articles 2 to 6), the Convention specifies that measures should be taken at a national level regarding "the confidentiality, integrity and availability of computer data." This includes "illegal access," "illegal interception," "data interference," "system interference," and "misuse of devices." In other words, the convention establishes that the actions required to undertake most modern cyber-security attacks should be prohibited by national laws. The objective of the Convention on Cybercrime is not to define new legislation, but instead to provide a framework to harmonize existing national laws. Hence additional national legislation might be required to achieve harmonization at an appropriate standard.

Second, the Convention examines computer forgery and fraud, which primarily addresses financial crimes facilitated by the operation or exploitation of computing devices. This area is growing in importance [39], but is mostly outside of the scope of the present paper.

Third, the Convention addresses content protection. Mostly, this part is about controlling the exchange of digital artifacts depicting child pornography.

Fourth, the Convention goes to great lengths to specify possible harmonization of the various copyright laws in the face of digital content distribution and possible abuses.

In terms of international agreements that the Convention aims to foster, the salient point is that extradition procedures of persons charged under the various crimes the Convention defines are relatively clearly specified. In particular, extradition applies to crimes that can be punished by jail terms of a year or "more severe penalties." Essentially, this seems to indicate that felonies (in the American sense of the word) would certainly have to be considered for extradition. However, as is often the case with extradition proceedings, the laws of "both parties" to an extradition proceedings must prescribe punishments for the offense for the accused criminal to be extradited. Specifically, if country A has no laws against a certain type of computer misuse, and an attacker from country A targets country B, which has strong punishments, it is unlikely that country A would accede to country B's extradition request, even if both have ratified the treaty. While such cases may occur, the Convention does create an extradition framework that could be used against remotely located criminals.

The Convention also exhorts its signatories to provide mutual assistance to each other. Specific details include coordinated data collection; and the creation of a continuously available point of contact at each member state to facilitate requests regarding preservation of data, provision of technical advice and collection of legal information.

As can be seen from the above description, the Convention on Cybercrime, while not ignoring potential cyber-attacks on critical infrastructure, primarily focuses on more "common" cyber-threats, such as the development of online crime, and the exchange of illicit contents (copyrighted works, abusive materials...). Indeed the Convention was primarily drafted in response to copyright infringement, distribution of child pornography, and financial crimes aided by online transactions.

Nonetheless, as part of a more general framework prohibiting various acts of unauthorized access, Title 1, Articles 2 to 6 of the Convention, which are outlined above, clearly encompass destructive actions on critical infrastructure. Because these articles do require that laws be passed at a national level, it is appropriate to review some of the most recent proposals in this respect.

2.2 Protecting Cyberspace as a National Asset Act of 2010

In June of 2010, Sen. Joseph Lieberman, joined by Sens. Susan Collins and Tom Carper, introduced the "Protecting Cyberspace as a National Asset Act of 2010" [31]. The bill specifically provisions for what would happen in case of a major crisis due to attacks on critical infrastructure.

In particular, the bill would grant the U.S. President the ability to assume control or deactivate parts of the Internet in emergency situations. Specifically, the President would have the ability to declare a "cyber-emergency." During such emergencies, companies identified as critical infrastructure providers would have to "immediately comply with any emergency measure or action developed" by the Department of Homeland Security, or face fines.

Critical infrastructure providers could encompass a large number of companies, which not only include Internet Service Providers (ISP), but also software manufacturers or search engine operators. The list of companies considered as critical infrastructure providers would be defined by the Department of Homeland Security (DHS).

In other words, any company considered as a purveyor of critical infrastructure, and that relies on "national information infrastructure" could be ordered to deploy emergency measures during a cyber-emergency.

The seemingly far-reaching powers of the bill, partly created by the lack of clear definition of what constitutes "critical infrastructure," other than presence on a DHS list, led to a spirited debate on whether or not the bill was attempting to implement an "Internet Kill Switch," as it was dubbed by the media.

In a post to the "Interesting People" mailing-list [9], former Assistant Secretary for the Department of Homeland Security, Stewart Baker, opined that this bill would primarily be restricted "to a relatively limited set of critical facilities and to the information infrastructure on which they depend." Baker goes on with the following example: "If operators of power grid were [...] to be relying on the Internet and Windows [...], then the authority to order emergency measures would apply to the providers of electric power, to their ISPs,

and to Microsoft. But other users of the Internet [...] will [n]ever be part of the covered infrastructure. Nor will they be subject to the emergency authority." [9]

Whether or not this bill will eventually pass remains, at the time of this writing, unclear. However, the bill's main merit, in our opinion, is to clearly articulate that critical infrastructure may depend on computing infrastructure, and that both types of infrastructures may be equally important during cyber-attacks. Furthermore, the bill gives the right incentives to perform network segmentation, that is, to enforce a strict separation between commercial and critical networks, which, as is argued in Section 5 is a viable option to reduce the impact of cybercrime. Whether the remedy suggested is appropriate, and could not be prone to abuse, is an entirely different debate. But, as Sen. Lieberman himself noted, some other countries (e.g., China, Gulf countries) have the ability to shut down part of their communication infrastructure in case of an emergency. The main counter-argument is that none of the countries with such capabilities is a Western-style democracy, and thus, the option ill fits the situation of the United States.

2.3 Key Escrow and Wiretapping Orders

As was mentioned in the preamble to this section, cryptographic material has long been equated with other dual-use products such as weapons and munitions. This has led to the perception that, much like violent crime could potentially be curbed by regulating weapon sales, thwarting online crime could be made easier by restricting cryptographic tools and algorithms available to Internet users. The rationale is that, deprived of the ability to make communications secret, miscreants would see their activity as a lot easier to control. Regardless of the actual merits of the proposition, legislative efforts in the United States have tended toward cryptographic restrictions; of utmost concern to the technical community are the efforts proposing to enforce *key escrow*.

Key escrow mandates that the parties to an encrypted electronic communication provide a trusted third-party with the key(s) used to encode messages. When talking about key escrow laws, the trusted third-party in question generally means a governmental organization or agency. Closely related to key escrow is the broader notion of digital wiretaps, which, as their name indicates, define a whole class of primitives that can allow a government organization, or law enforcement, to eavesdrop on communications.

Key escrow and digital wiretaps periodically come back in the news as a potential way to combat online miscreants. Anderson [6] documents that back in the early 1990s, the United States Government tried to impose adoption of the "Clipper" chip. This chip was proposed as a replacement for the then-ubiquitous Data Encryption Standard (DES), used for instance in many automated teller machines. Clipper would have been required for any kind of electronic encryption in the United States. Its main feature was to provide a "backdoor" access that could be used to bypass the encryption mechanism and make it possible for authorized parties to eavesdrop on traffic. The idea was that government officials could have access to any encrypted communication they needed to intercept. Following public outrage, plans to deploy the Clipper chip were abandoned.

However, history repeats itself. In 2010, the *New York Times* [49] and the *Guardian* [7] reported similar

legislative efforts under way in both the United States and the United Kingdom. The core idea is that, upon request, Internet Service Providers or software companies should be able to provide deciphered communications to the government. While the mechanisms needed to implement such proposals are not specified at this stage, from the functionality desired, three possible implementations present themselves: (i) key escrow, where the escrow is performed by the government; (ii) key escrow, where the escrow is performed by the service provider (either ISP, or software provider); or (iii) implementation of a backdoor allowing the circumvention of the encryption primitives used altogether, similar to the Clipper chip of the early 1990s.

As pointed out by Anderson [6] among others (see, e.g., [21]), key escrow, and, more generally, wiretapping laws, are unlikely to be very successful in preventing online crime, while they conversely risk alienating regular users. Indeed, while some Internet attacks make use of encrypted communications for coordination between attacking entities, knowing the end points of the communication (i.e., who is talking to whom) is generally the most important part in understanding the conversation semantics; for most protocols, discovering these channels does not require breaking a complex cryptosystem,[1] and indeed, several research efforts have been successful at taking over online criminal networks (see, e.g., [28, 56] among others) without breaking elaborate cryptographic algorithms.

Furthermore, key escrow would only apply to commercial software developed in a specific jurisdiction. Criminals, on the other hand, would be likely to either develop their own software and protocols or use open-source software developed outside of the jurisdiction in question[2], and would avoid the use of any commercial software on which restrictions have been placed. In other words, it seems dubious that key escrow or wiretapping laws would actually be of any help to curb online crime.

3 Threat Models and Associated Policy Challenges

The paper next classifies the different threats facing critical infrastructure. It discusses the seriousness of each of these threats, possible technological remedies, as well as the policy challenges in dealing with these threats. At a high level, it distinguishes between the following families of threats: information theft and espionage on one hand, and disruption of services on the other hand.

3.1 Information Theft, Leaks and Espionage

Information theft encapsulates a large number of security attacks, ranging from theft of intellectual property by company insiders, to industrial espionage, to copyright violations. From a security engineering perspective, all of these attacks share a common trait. The attacker reveals a secret, and subsequently shares it with

[1] An notable exception to that statement occurs when miscreants use anonymous networks, such as Tor [21] to have their machines communicate with each other. In such a case, identifying communication end-points is considerably more cumbersome. Fortunately, anonymity of communications between miscreant-controlled machines is still relatively rare, due to the performance penalties incurred when anonymizing traffic.

[2] This is already the case: most modern software used, or abused, for security attacks is either open-source, or exchanged in underground forums.

unauthorized parties. Note that information theft may also happen inadvertently – it is then categorized as an information "leak."

While they are not direct attacks on the infrastructure, information theft, leaks, and espionage in general pose a grave indirect threat to critical infrastructure. For instance, incidents in Japan in 2006 illustrate the information leak problem quite vividly. The popular Japanese peer-to-peer software Winny has been subject to a family of viruses, collectively referred to as "Antinny." One of the Antinny variants shares the entire contents of the infected users' hard drives, and automatically uploads some files onto the Winny network. With employees working on critical infrastructure taking their work laptops home, and using these laptops to download peer-to-peer files, a disaster was waiting to happen. Indeed, Antinny resulted in catastrophic leaks of information, e.g., nuclear plant data, databases of victims of sex crimes, or classified military data [53]. Once leaked, these secrets cannot be recovered. Along the same lines, the more recent leaks of hundreds of thousands of classified documents through the WikiLeaks site has brought to light a number of ethical and policy concerns linked with the diffusion of classified material.

However important an issue it may be, information theft predates electronic information networks, and, indeed, many laws and regulations already exist to combat the problem. While espionage has long been addressed in terms of national security violations (see, e.g., 18 U.S.C §792–799), more recently, bills tackling commercial espionage have also been signed into law. Specifically, the Economic Espionage Act of 1996 (18 U.S.C. §1831–1839) clearly defines what constitutes information theft of trade secrets and stipulates penalties for perpetrators. In the first actual test of the law, in 2010, a former Boeing employee was convicted under the Economic Espionage Act of 1996, and sentenced to 16 years in prison [41]. Certainly, some of the documents revealed as part of the recent WikiLeaks episode affect national security, and as such, the person(s) who leaked these documents could conceivably be charged under 18 U.S.C §792–799.[3]

As far as information theft goes, in fact, one can argue that the nature of the media concerned (microfilms, digital files, paper, ...) matters less than its contents. As such, it is not surprising that many of the existing laws do not require to be adapted to accommodate for digital transmission of contents.

For copyright violations, the situation is slightly different. As discussed at length in another report [15], the major change brought about by the rise of electronic networks lies in the potential magnitude of the violations. While copyright violations remained very limited in scope until the mid-1990s, the increased ability to replicate and immediately disseminate digital artifacts on a large scale has made copyright violations considerably more significant than in the past. But here again, numerous laws already exist. In particular, the Digital Millennium Copyright Act (17 U.S.C. §§512, 1201–1205, 1301–1332; 28 U.S.C. §4001) is a set of additions and modifications to older copyright laws (collectively referred to as "Copyright Act of 1976") to account for the challenges posed by electronic information networks.

As such, it does not appear that novel laws need to be created, although certain areas remain relatively ill-defined, as evidenced by the judgment rendered by the U.S. Supreme Court in MGM vs. Grokster (545 U.S. 913) which argues that software designed primarily for infringing uses, may be held liable for secondary

[3] Note however, that, for such prosecution to be successful, it should be proven that the person causing the leak deliberately gave this information to a foreign power, which, from a legislative angle, is a much more complex proposition.

copyright infringement. The question of how to determine intent of the software designer is not addressed in the U.S. Supreme Court's decision.

Beyond espionage and copyright laws, one can further argue that, from an international standpoint, laws dealing with the circulation of counterfeit goods or contraband could also apply to information flows. In any case, it does not appear that there needs to be additional legislation crafted specifically for the purpose of electronic information networks.

Last, related to the topic of information leaks is the question of preserving confidentiality of information. This can be absolutely critical for certain types of service providers. For instance, medical providers, including hospitals and pharmacies, must do their utmost to preserve patients' privacy. Hence, laws mandating that security measures be taken to preserve record confidentiality are useful, but can generally be designed at a national level. In the United States, the Health Insurance Portability and Accountability Action (HIPAA) of 1996 (P.L.104-191) fulfills that function. One issue that may arise is when patient records cross borders; for instance, if a US citizen, being treated in the United States, decides to get surgery in Canada. There is, at this point, little harmonization between different national privacy protection laws [48], which may impact companies that outsource some of their data processing services abroad.

3.2 Disruption of Services

Apart from information theft, most cyber-attacks result in disruption of services. This classification is overly broad as it encompasses, for instance, both productivity losses linked to employees being busy deleting email spam messages, as well as distributed denial of service attacks targeting critical function. This section focuses on a subset of attacks that have the worst impact on critical infrastructure.

3.2.1 Insider Threats and Sabotage

Insider threats are directly related to espionage as discussed above. However, here, the focus is on insider threats resulting in destructive actions. Such activities are traditionally referred to as *sabotage* rather than espionage, although the two types of threats usually go hand-in-hand. As is the case for espionage, existing laws define sabotage in the national security realm (e.g., 18 U.S.C. §105) and in the commercial realm (18 U.S.C. §1951, which more generally applies to "Interference with commerce by threats or violence") and do not necessarily need to be amended to encompass electronic sabotage.

3.2.2 Destructive Penetration

Destructive penetration happens in two stages. First, an element of the computing infrastructure is compromised by an external object. For the purposes of this paper, the phrase "external object" encompasses all *malware*, that is, programs written with a malicious intent in mind. More specifically, this paper primarily concerns itself with spyware (code that is manually, but involuntarily, installed by the machine operator), and worms and viruses, which are self-replicating programs that can propagate automatically, or

semi-automatically, over networks of computers.[4] Second, this piece of malware subsequently destroys the machine it has infected.

Usually, deployment of such attacks occurs by contagion. That is, a machine is initially compromised and then seeks other vulnerable machines to infect by probing (randomly or deterministically) portions of the Internet. Once other vulnerable machines have been compromised, they in turn attempt to further the infection by compromising other vulnerable machines. Epidemiological models have been built to describe propagation patterns of such viruses, and have been shown to closely track measurements of the number of compromised systems [36, 55].

Rather than using the Internet, older viruses propagated through physical media such as floppy disks and often contained logic bombs. A logic bomb is a piece of computer code that remains harmless until a certain trigger, such as a clock reading, is activated. Logic bombs could cause serious harm to compromised systems, for instance reformatting their hard drive, or corrupting boot sectors. While the harm was reversible (e.g., by reinstalling a fresh operating system), such attacks usually incapacitated the targeted machine for a period of time.

Modern (i.e., post-2001[5]) computer worms and viruses have generally had relatively mild effects on the machines they compromised. Indeed, worms are generally used to provide control over the compromised machine to an attacker; the machine so compromised is subsequently used to attack other machines while making it more difficult for the attacker to be identified, as discussed in the next subsection.

The reason for the apparent milder effect of modern malware compared to its ancestors is primarily economic. With the increased monetization of online crime [39], compromised machines and their contents (e.g., financial and other confidential information) have become a commodity that can be traded in underground economies. Destroying these machines would translate into a loss of revenue for the miscreants behind the compromise, and is therefore unlikely to be a frequent option.

However, nothing prevents a determined attacker from unleashing a worm or virus that compromises a large number of machines, with the objective of incapacitating some critical infrastructure, by destroying compromised machines, rather than exploiting them for profit. Such a scenario is described by Weaver and Paxson [60], who imagine that a computer worm could be programmed to propagate in 15 minutes or less over the entire Internet, making reaction impossible at human timescales. The worm imagined by Weaver and Paxson would proceed to reflash the BIOS of the computers it targets. The BIOS is a small program that is immediately run after powering on the machine. This program typically resides in on-board memory, and altering it can make it impossible for the computer to boot. Thus, BIOS reflashing with a malicious intent

[4] The terms "worms" and "viruses" are often used interchangeably. A sometimes-adopted definition distinguishes between fully automated propagation of harmful code (worms) and propagation initiated by manual action such as opening an email attachment (viruses). Because the distinction is irrelevant to the the discussion in this paper, the terms "worm" and "viruses" are synonyms in this paper.

[5] 2001 marks the emergence of the "Code Red" worm [55] which was the first modern, automated large-scale worm to rapidly propagate over the entire Internet. The Morris Worm from 1988 actually was the first program to harness the power of the Internet to propagate on a large-scale, but, back in 1988, the Internet hosted a comparatively small pool of machines, and at the time, almost none of the critical infrastructure was connected to it.

would effectively destroy part of the targeted computer's hardware, and make recovery only possible after time-consuming and expensive efforts (e.g., integrated circuit replacement). Weaver and Paxson argue that economic damages incurred by such a destructive worm would far exceed the tens of billions of dollars [60].

Such devastating threats have yet to be realized in practice. However, the recent incidents linked to the Stuxnet worm suggest that some classes of attackers may not necessarily worry about collateral damages and attempt to infect as many systems as possible in order to destroy a few specific targets. Indeed the Stuxnet worm tries to reprogram some very special purpose hardware, but, at the same time, aims to propagate to large numbers of machines. By analogy to epidemiology, one could model Stuxnet as a rather aggressive virus for a certain portion of the population (i.e., those machines running the specifically targeted hardware), but one that mostly relies on asymptomatic carriers to propagate.

As in epidemiology, it is extremely difficult to detect the original infection point – i.e., the "patient zero." For many of the most famous modern worms, the original infection point was never identified. For instance, neither the original infection point or the authors of the Slammer worm [35], which is discussed in more detail below, were ever found. Likewise, while the CodeRed worm propagation rate was *a posteriori* modeled using epidemiological models [55], its original infection point, as well as its authors, remain a mystery. In fact, for most modern worms, it is almost impossible to find out the original compromise, unless the author of the worm makes a mistake that can lead investigators on its trail. For instance, the author of the Sasser [58] and Netsky worms [57] was identified, but only after he boasted of his feat to acquaintances. In other words, attribution of authorship is usually done through traditional police investigation, rather than technological discovery.

Furthermore, even when the author can be identified, he or she is unlikely to be in the same jurisdiction as his/her victims. For instance, in the case of the Sasser and Netsky worms, the perpetrator was a German citizen, residing in Germany. Because he was under 18 at the time, he only received probation and community service, per German law. Given that both worms had global impact, he would undoubtedly have faced a more severe sentence in some of the countries where machines were compromised.

Intrusion Detection. Due to the difficulties in pinpointing the origin of such infections, technological defenses have mostly been focused on preventing infections from impeding one's network. In other words, the technological strategy to defend against destructive penetration has mostly been one of containment, rather than prevention or deterrence. More precisely, defenses have traditionally been focusing on intrusion detection, both at the end-host (computer) level and at the network level.

At the end-host level, anti-viruses are the main defense against malware, and are now part of all major computing infrastructures. However, classifying programs as malware or as legitimate is not an easy task and is still a very active area of research. There are essentially two complementary approaches being used: signature-based detection of malware and behavioral analysis. Most anti-virus programs operate with a database of so-called "signatures" of known bad programs. Signatures typically refer to binary patterns present in memory or on a disk when known bad files are copied to the machine or run on its processor. Anti-viruses periodically check each file present on the system against those signatures. The process is rel-

atively fast but is vulnerable to minor changes in the harmful programs. For instance, two programs may have exactly the same semantics, yet have considerably different binary footprints, a problem known as *polymorphism*. Worse, given that modern computer architectures allow for self-modifying code, a program may change its behavior while running and easily escape signature-based detection. In addition, the number of signatures of known bad programs has been exponentially increasing over time [13]. Due to the limitations of signature-based detection, behavioral analysis (e.g., [29, 42]) has recently been proposed. The idea is to have programs run in a confined environment before being allowed to run on the machine proper, and to check their behavior in the confined environment against known undesirable properties. The difficulty in deploying behavioral analysis is its computational overhead. Programs have to run a certain amount of time before they can be declared as "safe," which precludes any fast detection.

At the network level, the intrusion detection literature is extremely rich (see [8] for an overview), and intrusion detection systems can roughly be divided into two categories. Similar to anti-viruses discussed above, signature-based intrusion detection systems, such as Snort [44], basically match incoming network traffic patterns with a database of known signatures of harmful traffic. Traffic identified as malicious may either be denied entrance to the protected network, or may raise an alert sent to the network administrator. The main issue with signature-based intrusion detection is that, much like signature-based malware detection, it is essentially powerless against attacks that do not fit a particular signature. For instance, slight variations of existing attacks may result in slightly different traffic patterns, which in turn may allow for evading intrusion detection systems.

A different approach, advocated by the Bro intrusion detection system [43] is to build state machines that track the status of a connection and only raise alarms when several conditions are met. The main advantage of such stateful intrusion detection systems is that they can detect attacks that do not conform to a very specific traffic signature. On the other hand, stateful intrusion systems tend to require more processing resources than their signature-based counterparts.

Trusted Platform Module and Attestation. A more recent line of work to provide technological defenses against destructive penetration is that of *attestation* (see for instance [45, 51, 52]). The idea is to check all programs running on a machine, against an existing database of "valid" programs.[6] If unwanted software (e.g., a virus) resides on the machine, it will not pass this check and will not be authorized to run. One can see that the main conceptual difference with the intrusion detection schemes discussed above is that by default, programs are *not* allowed to run in such environments, and must instead show that they are benign before being allowed to run. The risk then becomes that a legitimate program is mistakenly not authorized to run; but as pointed out by Saltzer and Schroeder in their seminal paper on access control [47], such a "failsafe default" is actually desirable – indeed, the machine does not become more vulnerable due to a failure of the security system.

[6] Attestation actually refers to the ability of a machine to provide proof, to an inquiring party, that it is running a certain program. By a slight abuse of terminology, attestation is used to refer to the entire body of research devoted to providing primitives that ensure correctness of the programs being run on a given platform.

The main technological challenge is to ensure that the hardware (or software) in charge of performing the checks is itself trustworthy. The Trusted Platform Module (TPM) is one possible solution. TPM are hardware chips that, through a combination of cryptographic primitives and secure storage can provide measurements of files present on a system with guaranteed integrity (that is, the measurements cannot be altered by an adversary) [46]. A number of modern computers (e.g., IBM Lenovo laptops, recent Apple MacBook laptops, ...) include a TPM chip, but those are so far relatively unused. A second challenge is to ensure that the database against which measurements are to be checked remains secure. While securing the contents of such a database seems possible (for instance by requiring that only authenticated parties add new entries[7]), TPM chips have policy implications that exceed the realm of security. For instance, certain vendors could collude to prevent inclusion in this database of trusted programs written by their competitors [50].

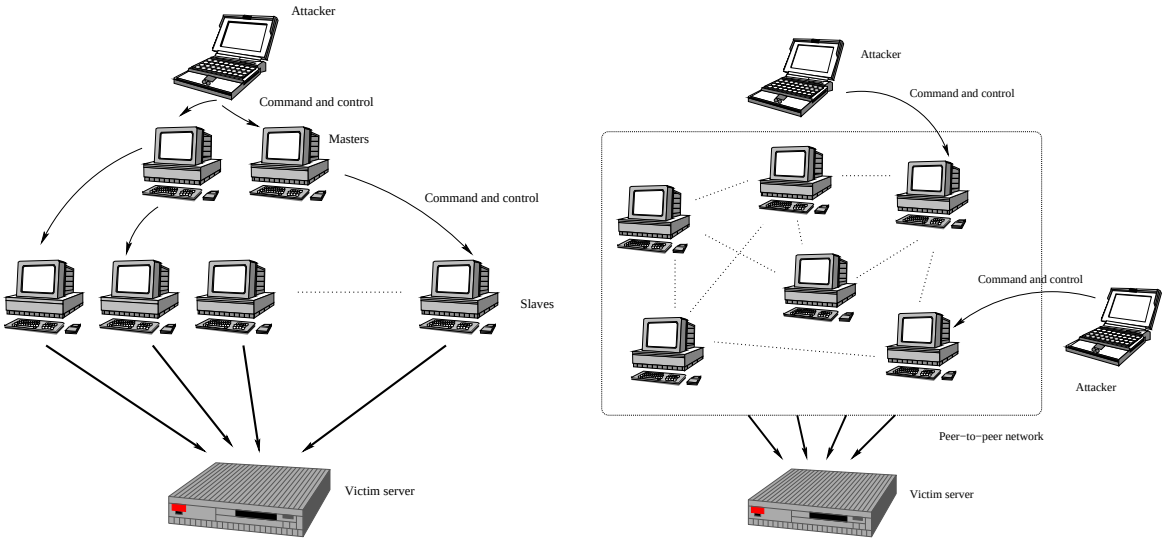
3.2.3 Distributed Denial of Service.

Distributed denial of service (DDoS) attacks are closely related to destructive penetration discussed above, and are one of the major scourges to afflict the Internet in recent years. Most attacks proceed in two phases. First, a number of vulnerable hosts on the Internet are compromised by malware, in a manner identical to what was described earlier. But instead of destroying the hosts so compromised, the attacker instead installs programs that give him/her full control over these victim machines. Subsequently, the attacker communicates orders to these compromised machines, for instance, sending a large number of individually innocuous requests to a target, in an effort to momentarily incapacitate the victim.

Figure 1 illustrates how networks of compromised machines are used to perform denial of service attacks. In the early 2000s, which correspond to the early days of modern distributed denial of service attacks, compromised hosts were usually federated in a hierarchy separated between slaves and masters – the decision to grant master status to a compromised machine was generally dependent on its availability, reliability and performance [34]. Figure 1(a) shows that an attacker would transmit orders to a handful of masters, which would each relay the orders to a number of slaves. Eventually all slaves would receive the attacker's command and would send large amounts of traffic to the intended victim. The advantage of such hierarchical protocols is two-fold: the attacker is guaranteed relative anonymity through indirection, and the dissemination of command and control messages is made much more scalable to a large number of compromised hosts.

Figure 1(b) shows a more modern infrastructure, as described for instance in [22]. Contrary to past efforts, modern denial of service attacks usually do not rely on networks of machines solely compromised for the purpose of launching such attacks. Instead, miscreants compromise a large number of machines, federate these machines into peer-to-peer networks using modified versions of known protocols (e.g., Kademlia [32]), and then sell "timeshares" of computing time on the resulting computing infrastructure to other miscreants. In essence, this is nothing different from cloud computing, except that both the "cloud" providers and

[7] In a different context, the AppleStore, in which all applications are vetted by Apple prior to being available, gives a proof of concept that such databases are viable.



(a) Hierarchical organization (c. mid-2000s) (b) Peer-to-peer organization (c. late-2000s-early 2010s)

Figure 1: **Denial of Service infrastructures.** Denial of service infrastructures have evolved from hierarchical networks specifically designed for distributed denial of service attacks (a), to more general networks, usually organized using peer-to-peer protocols that can perform denial of service attacks as one of many "services" they provide (b). (Other such services include hosting infrastructure for fraudulent websites, dissemination of spam email, among others.)

their users are all operating in complete illegality, sharing and bartering access to machines that do not belong to them. These compromised machines are often referred to as "bots," and the resulting networks of compromised hosts, or "botnets" have been extensively studied in recent research (see, e.g., [4, 27, 28, 56] among others).

Regardless of the operating mode of the attacker, the victim usually can only trace the source of the attack back to the compromised hosts that flooded it with unwanted traffic. On the other hand, identifying the villain that commanded the attack is a lot thornier, as modern botnets are very often used by many different parties, and are accessed through multiple layers of indirection. For instance, attackers can resort to onion routing [21] to communicate with the botnet anonymously,[8] or can, more simply, talk to their bots through intermediate, proxy servers, to shield their identity – in either case, it is near impossible to re-identify the origin of the attack.

As a case in point, there is still no consensus on who was behind the cyber attacks of June 2010 that took down prominent United States and South Korean websites [61]. Due to obvious geopolitical connections, North Korea was suspected, but, for the technological reasons outlined above, no evidence thus far has credibly pointed to any identifiable perpetrator.

Whenever attack attribution (i.e., identifying the perpetrator) can be performed accurately, existing agreements such as discussed in Section 2 could then be used to prosecute and/or punish. This is similar to

[8]Communication between different bots, on the other hand, is usually not anonymous, due to performance reasons.

the dissemination of illegitimate pornographic goods (e.g., child pornography). While it may be difficult to identify the producer of certain pictures with high accuracy, when an author or perpetrator is found, international conventions exist to bring them to justice. However, considering how difficult attack attribution may be, such laws may generally be hard to apply. Instead, it may be worth exploring conventions that could make it harder for attackers to carry out their deeds, as discussed next.

3.2.4 Collateral damages due to other attacks

Some, if not most, of the attacks outlined above sometimes result in collateral damages for third parties that are not, a priori, involved in the attack itself. One such example was that of the "attack" that crippled essential navigation services offered by the Port of Houston (Texas) [30]. The Port of Houston was not the intended victim of the attack. Rather, one chatroom user, based in the United Kingdom, allegedly launched a denial of service attack on another user, based in South Africa. Systems used by the Port of Houston infrastructure happened to be on the path between the attacker and the target, and were flooded with traffic, which made them unresponsive.

In that specific case, the machine from which the attack originated was clearly identified, and its owner was prosecuted in the United Kingdom, where he resided at the time, under the Computer Misuse Act of 1990 (1990 (c. 18)). He was ultimately acquitted, since doubts about whether or not his machine had been exploited by another person to conduct the attack were never alleviated in court.

Another example is that of the Slammer worm that was mentioned earlier. Slammer essentially shut down large portions of the Internet due to the overload it imposed on the network infrastructure. This is all the more remarkable considering that Slammer, itself, is not a particularly harmful program. Its only goal is to replicate as quickly as possible over the network; but other than this replication primitive, the worm is actually completely harmless. However, the worm ended up being so successful at propagating that the network traffic it generated overloaded packet routers that are essential to transmit Internet traffic. Routers overloaded due to Slammer stopped processing appropriately routing messages coming from other routers, which furthered network instability even more, as unprocessed messages were reissued, adding to the already overloaded state of the network.

As discussed before, finding the originator of an attack is usually very difficult and may not even be possible. However, this does not necessarily matter when dealing with collateral damage. Essentially, collateral damage mostly occurs due to what economists call negative externalities. That is, a network A may suffer from an overload of traffic, because an unprotected network B is letting attackers take full advantage of it. While network B's owners are not actively malicious, they can be viewed as negligent, in that they facilitate the attack taking place.

Negative externalities are a problem well-known to economists. For instance, pollution is an example of negative externality. A given factory may increase its output and make more profit in the process, but the increased output increases pollution and thus penalizes everyone else. The classical solution to negative externalities is to impose "side payments" that essentially re-align economic incentives. In the factory

example, taxing polluters proportionally to the amount of pollution they produce is one way to re-align their economic incentives toward a more desirable outcome. In the context of network security, ISPs have the economic leverage to choose who they receive traffic from. In the example given above network A could realize that a vast amount of undesirable traffic transits through network B and accordingly renegotiate the peering agreements (or transit agreements) it has in place to govern its connectivity with B.

As a case in point, a host of measurement studies indicated that from April 2008 onward, taking down a particularly suspicious Internet Service Provider (McColo) could significantly reduce the overall volume of email spam at the time. The suspicious service provider was eventually disconnected in November 2008 from the rest of the network by its peers – not by legal action but essentially being punished for being a "polluter" [18] – and ended up bankrupt. Through this incident, market forces were at play. Indeed, the community essentially became fed up with a misbehaving member and imposed sanctions by itself, although it did take more than seven months before decisive action was taken. It is plausible to imagine that the same outcome could have been obtained through legal enforcement, but the question remains whether legal pressure is necessary as a complement to existing economic incentives. Since negative externalities are in play, new legal incentives may be useful in rectifying undesirable economic incentives.

More generally, security analysts need to be alert to the possibility that attacks on commercial networks can have dire side effects on critical infrastructure. The Port of Houston incident discussed above is one such example of critical infrastructure suffering from an unrelated attack. The insight, here, is that segmentation of networks and services is absolutely needed to provide isolation. Indeed, the only reason why the Port of Houston suffered from the attack is because its essential services shared connectivity with the commercial Internet. This illustrates the point that vital services should be isolated from the rest of the network – possibly running on physically separate infrastructure.

Here, international conventions may not help as much as national best practices imbedded in law. For instance, the United States military has rules that demand that military functions be separated from civilian functions and use different physical networks for classified data in order to minimize the risk of leaks between confidential channels and open channels [6, 12]. One could conceivably imagine drawing inspiration from these policies to enact similar practices for critical infrastructure. This paper returns to the notion of network separation in Section 5, but note that the aforementioned Liebermann bill discussed earlier provides incentives to perform such separation, as critical infrastructure providers would want to minimize the amount of infrastructure that can be seized during an emergency situation.

4 Practical Instances of Threats to Critical Networks

The paper next relates the threats discussed above to existing critical networks. It namely looks at financial infrastructure, air traffic control systems, hospitals, and power and water systems in turn. While far from exhaustive, the discussion in this section chiefly aims at providing some brief, practical examples of threats faced by critical infrastructure. This discussion refrains as much as possible from speculating about hypothetical threats, and instead looks at potential vulnerabilities that seem relatively glaring, considering what

is known about these infrastructures and about past attacks on similar systems.

4.1 Financial Infrastructure

The financial infrastructure is obviously an important piece of critical infrastructure. It faces essentially two very different types of threats: attacks on the information exchanges it supports, and attacks on its computing infrastructure.

4.1.1 Attacks on Information Exchanges

These attacks first include all types of leaks or thefts of confidential data, ranging from individual banking account information to insider information leading to illicit trades. This body of attacks usually comes down to one major vector: infiltration of personal computers by malware. How such infiltration is done, and can be mitigated, is discussed in Section 3.2.2.

In the United States, all publicly traded companies have a legal impetus to ensure that their financial operations are protected and cannot afford to be negligent. The Sarbanes Oxley Act of 2002 (Pub. L. 107–204, 116 Stat. 745) imposes that each publicly traded company issue a yearly report on its activities and compliance. Per Section 404, the report must include a report from management on the company's "internal control over financial reporting." The report should contain in particular, a statement of management's responsibility for establishing and maintaining internal controls and maintaining an adequate internal control structure and procedures for financial reporting. The report must also "contain an assessment, as of the end of the most recent fiscal year of the Company, of the effectiveness of the internal control structure and procedures of the issuer for financial reporting." (15 U.S.C. §7262a). In layman's terms, every year, management has to certify they guarantee information security (including, but not limited to, preventing conflicts of interest) of their financial reports, as well as evaluating the effectiveness of the information security technologies they use. Fraudulent reports are subject to civil and criminal penalties.

Social Engineering. An additional vector of attack frequently used on financial information flows is that of social engineering. So-called phishing scams, in which villains entice victims to reveal confidential banking information under false pretenses [20] have been a flourishing criminal enterprise. Numerous measurement studies have shown the impact of phishing and have been useful in describing attacker tactics [37, 38, 40]. The main takeaway from these works is that infrastructure protection is best realized upstream. That is, pressure needs to be applied to network service providers that facilitate such scams, for instance, careless domain name registrars who do not check their registrants' identities, preferring to make a profit even when their customers are engaging in illegal acts. This is another example of negative externality that could be addressed by legal remedies.

Social engineering is also used in stock manipulation attacks. The idea of these attacks is to send vast amounts of spam email advertising that some (usually little-known) penny stock is going to rise. Unaware investors may indeed notice a rise in the stock price (due to the attacker and their associates purchasing

significant quantities of it while they are sending these spam emails). This rise substantiates the information "disclosed" in the email, and victims may be enticed to buy some of this stock. Of course, the entire scheme is nothing more than a modern variant of the "pump and dump" scam, where the attacker quickly rids herself of all of her stocks as soon as enough victims have partaken in the scam. Such scams are facilitated by email, may account to up to 15% of all spam, and have been quite successful, at least until the mid-2000s [10, 24].

There are two complementary approaches to combat the problem of stock manipulation through email scams. One can try to curb spam, which has so far been a relatively vexing problem (despite advances in filtering techniques which have arguably reduced the magnitude of the harm), or one can monitor stock exchanges for irregularities. In the United States, the latter squarely falls within the purview of the Security Exchanges Commission, which, in our opinion, already has at its disposal the legal and investigative arsenal necessary to identify and rectify such problems.

User education may also be considered, but as pointed out in recent work [16, 54] it is unlikely to be effective. User education is indeed facing an uphill battle when it is competing with apparently adverse financial incentives; that is, even though people are told something is a scam, they may still elect to participate in it because of the (illusory) potential for monetary gain.

4.1.2 Attacks on Computing Infrastructure

Due to the attack vectors discussed above, most efforts aimed at securing financial infrastructure have focused on end-user protection, through technologies designed to minimize infiltration or social engineering.

Comparatively little has been done at the infrastructural level (e.g., stock exchange), in the way of introducing novel technology. To be sure, network protection through firewalls, network redundancy, and capacity overprovisioning are all part of the technological arsenal used to prevent infrastructural mishaps. Yet, the key point is that financial actors are much more worried about the timeliness and reliability of their trades, than they are worried about threats to their computing infrastructure. In particular, for many large trading firms, the main goal is to have as low a latency as possible in issuing their trading orders. For this reason, such large firms build dedicated network lines to communicate their orders. As argued in Section 3, such a segmentation is useful is preventing collateral damages due to unrelated network attacks.

However, there is a large number of private investors and small institutions that do not have the resources to implement such dedicated networks, and are instead relying on commodity Internet services for their transactions. As shown by the relative success of the pump and dump scams discussed earlier, an attacker managing to compromise a majority of these private or semi-private investors (using malware, for instance) could certainly have an impact on the stock market as a whole.

Perhaps more worrisome is the problem of software reliability. The May 2010 "flash crash" of the New York Stock Exchange, which lost 600 points in minutes before recovering almost as quickly is attributed not to the act of a malevolent person, but to a computer-based decision algorithm causing a large investor to send sell orders at a frantic pace [11].

In sum, protecting the financial infrastructure seems to be best achieved by a combination of three

factors: improving software reliability, enticing network participants to perform network segmentation as much as possible (which, again, is an economically sound proposition since it also results in lower latency for traders), and combining these mandates with incentives for the rest of the network to avoid sending harmful traffic.

4.2 Air Traffic Control Systems

The second piece of critical infrastructure examined is Air Traffic Control (ATC) systems. Compared to other navigation systems, there is, in fact, nothing inherently specific about securing ATC systems against vulnerabilities, other than the large amount of traffic and data they continuously process.

However, recall the earlier example of the Port of Houston suffering collateral damages from an a priori unrelated distributed denial of service attack. This attack led to an undesirable outcome because the Port of Houston was sharing infrastructure pieces (links and routers) with the commercial Internet. It is far from clear that there was a technical necessity behind this: various other, special purpose networks could have been used to communicate navigation data to ships.

Along the same lines, it seems relatively unreasonable to have the critical functions for navigation systems connected to the Internet. In fact, while most of the ATC systems are not connected to the Internet directly, some machines are connected both to internal ATC networks, as well as to the commercial Internet. As a result of this coupling of part of the ATC network with the Internet, those ATC systems can be compromised by anybody with an Internet connection – and in fact, sobering incidents have already demonstrated the feasibility of such attacks [33]. Hence, it seems clear that here too, segmenting the network is likely the most natural way of making attackers' task considerably harder. Here network segmentation, makes sense, as there is no operational imperative for connecting critical navigation systems to the Internet: navigational and positional data is fed to these systems by different networks.

4.3 Hospitals

Hospitals offer a particularly complex environment, which relies on multiple critical components. For instance, hospitals have long had backup power generators, in case the main power lines shuts down. From a cybersecurity perspective, hospitals, much like financial institutions, face two key issues: preservation of confidential information and infrastructure availability.

Preserving confidential information is particularly important, and is in fact mandated by law in may countries (e.g., HIPAA in the United States) as discussed in Section 3. Hence, access control models have been designed to ensure that as few information leaks as possible do occur. In the UK, for instance, the British Medical Association completely revamped its access control mechanisms in the mid-to-late-1990s in an effort to address multiple privacy violations [6]. However, the difficulty here resides in devising access control models that are flexible enough for doctors, nurses, and other treating personnel to be able to do their job without any obstacle, while at the same time providing strong privacy safeguards. Unfortunately, convenience and security are often at odds: flexible access control means that the risk of illegitimate accesses

(e.g., employees peering out of curiosity into the medical file of a celebrity to which they should not have access) is increased. Periodic audits of all accesses can help mitigate these risks and could be in fact mandated by law. In other words, laws stipulating that hospitals should be in charge of enforcing information security on their records are desirable, but the need for international agreements beyond harmonizing data protection laws is uncertain.

Infrastructure availability is a different issue. Hospitals not only rely on critical computer support, but also on critical physical facilities. In a particularly deplorable incident, a rogue employee attempted to compromise HVAC control systems of the hospital in which he was employed as a security guard [25]. This was simply done by having unfettered physical access to the computers governing the HVAC flows. The employee was only apprehended after boasting of his plans on YouTube. This hospital should have been forced, by law, to have much stronger access control policies in place for its critical physical systems; in particular, the security guard should never have been able to access these machines.

Computing infrastructure, itself, is likely to be a lot less critical in the case of hospitals than in some of the other critical infrastructure (e.g., ATC networks discussed above). Indeed, at a fundamental level, most health care can be performed without the aid of networked computers. The main use for computer networks in hospital environments is tied to digital records, which itself can be problematic if these records are inaccessible or tampered with. Having back-up systems (e.g., paper-based medical charts, or even electronic records stored on devices that are not connected to any network) seems appropriate to ensure continuity of care, even in face of adverse circumstances.

4.4 Cyber-Physical Systems

Power, water, and gas distribution systems, at first glance, do not seem to rely on computing infrastructure as much as other critical functions. However, most control systems (e.g., dams, power grid controllers) are nowadays governed by computing systems, hence the term "cyber-physical systems" to generally describe these infrastructures.

Cyber-physical system operators are generally loathe to replace working computing equipment, and risk catastrophic faults in the process. Indeed, numerous pieces of equipment and hardware used in cyber-physical systems are very expensive and have a long lifespan (tens of years) compared to the computing equipment with which they interface. As a result, the computing infrastructure used is often obsolete. It is not uncommon to find, in 2010, Windows 98 or 2000 machines operating certain pieces of equipment (sensors, actuators), simply because there is no software available for these specific pieces of equipment on more modern operating systems. Considering that these computers provide the required functionality, the fact that they may be vulnerable to a myriad of attacks is often nothing more than a distant preoccupation in the minds of their operators.

Additionally, cyber-physical systems generally have a high degree of physical security (e.g., barbed wires and walls). This creates a misconception among cyber-physical systems operators that their systems are safe. However, in recent years, for cost-efficiency reasons, cyber-physical systems have become in-

creasingly interconnected with the business networks of the companies operating them; and in turn to the Internet [26], which makes them vulnerable to remote attacks. Recent research has examined, by simulation and field experiments, the destructive potential of cyberattacks (see, e.g., [5, 14]), and the results have been sobering.

The question is how does one move forward? Firstly, there needs to be legal demands on the operators of critical cyber-physical systems to provide adequate levels of security to their assets. Much like physical assets such as pipelines are regulated (e.g., in the United States, by the Pipelines and Hazardous Materials Safety Administration, PHMSA), the computing infrastructure that interfaces with them must be part of a certification process. From a technical standpoint, it is absolutely necessary to revert to network segmentation (or separation) of these assets. In a vast majority of cases, there is absolutely no technical reason why a machine governing a dam sensor would have to be somehow connected to the same intranet as the main office.

Unfortunately, physical separation is itself not sufficient, as the Stuxnet worm demonstrated. This worm can indeed use several vectors of propagation, including propagation via USB drives. In other words, even a machine disconnected from all networks can be infected, for instance if a contractor is (unknowingly or not) carrying the virus on an infected USB drive that they plug into the machine. Thus, network segmentation is much more complex than simply disconnecting a network cable – thorough access control has to be enforced at all levels. In particular, in the above example, machines serving critical functions should essentially not be able to install any data or programs from an external input.

Research in multi-level and multi-lateral security (see, e.g., [6, chap. 8–9]) has produced a number of formal guidelines to prevent contagion of systems with unverified inputs, which could be readily used to devise access control policies of such critical systems. Likewise, the use of Trusted Platform Module technology discussed above could also be a solution to ensure that only trusted software is installed and run on critical functions.

5 Discussion

This study has examined the major cybersecurity threats on critical infrastructure, by looking at a taxonomy of attacks, and considering practical instances of critical systems. It has also discussed technical factors facilitating or mitigating these threats, and have considered the interactions between these technical factors and some of the policy decisions that could be made. This section summarizes the findings and subsumes them as recommendations for building legal frameworks.

Because international conventions and laws heavily rely on national laws, it is necessary to separate potential legal remedies between those that can (or should) be deployed at a national level, and those that are more suited to international frameworks.

5.1 National Level

A common thread throughout this analysis is that of network segmentation. This paper has argued that critical resources should not be connected to large commercial networks such as the Internet. However, there may be strong economic incentives for doing so. Consider an electricity provider: connecting electricity meters to the accounting department allows significant savings in the invoicing process, compared to having employees driving around to physically inspect these meters. However, the accounting department itself is likely to be directly connected to the Internet, giving a possible path to attackers to access critical resources. This is the direction to which calls for a "smart grid" are pointing.

Thus, it is particularly important that laws at a national level re-align economic incentives by subjecting computing infrastructure used in critical functions to the same level of scrutiny as the rest of the infrastructure. While network segmentation may not necessarily be strictly feasible in practice, it is of utmost importance that operators of critical functions be incentivized to develop the most secure practices, *before* incidents happen. In this respect, laws mirroring Sarbanes Oxley (or, to a lesser extent HIPAA) should be devised for critical functions.

Interestingly, laws granting extraordinary powers of seizure of critical infrastructure, such as the "Protecting Cyberspace as a National Asset Act of 2010" may, quite counter-intuitively, be a step in the right direction. By creating an extremely undesirable situation in a time of crisis (take-over of private entities by government), they would create strong incentives for private entities that could be subject to seizure to adopt reasonable security policies and avoid being affected.

5.2 International Level

At an international level, the situation is a little bit different. While national laws should be used to incentivize targets to better defend themselves against attacks, one could imagine international laws being devised to bring to justice the perpetrators of the attacks.

However, as this paper has pointed out in various places, attack attribution is a difficult, if not impossible, task, unless the attacker makes a critical mistake. Thus, rather than focusing on devising laws to punish perpetrators, a more useful legal framework would ensure that economic incentives do not help miscreants.

Currently, there is almost no incentive for a domain name registrar to thoroughly check the business in which its customers engage. As such, some registrars tend to offer services to customers of questionable repute (see, e.g., [17] for a small case study on a particular instance of the problem), even when it has been made clear that dubious practices were at play. Likewise, there is almost no incentive for Internet Service Providers to ensure that the end hosts to which they provide network access are properly secured and cannot be abused by miscreants. If a given ISP's network is infested by bots, the ISP may simply ignore them as long as they do not contribute exaggerated levels of traffic.[9] In fact, there is almost no reason for the ISP to take action. Thus, it would appear quite useful to have regulations that rectify such negative economic

[9] If, on the other hand, the bots were sending massive amount of traffic, the ISP would likely take action on economic grounds, as bot traffic would be detrimental to traffic from other customers and potentially threaten the ISP's revenue stream.

externalities.

Trachtman [59] proposes an interesting analogy with maritime cargo. According to the International Maritime Organizations' International Ship and Port Facility Security Code, "each vessel's flag state is responsible for reviewing and certifying its security plan." [59] While Trachtman points out limitations of drawing too close a parallel with data networks – notably the infeasibility of inspecting every single data packet "at the border" for traces of potential security attacks, similar conventions could be drafted for data networks.

For instance, one could conceivably impose fines on networks that have been shown to transmit large amounts of denial of service traffic over time, even though the attackers may not originate from these networks. Likewise, one could regulate more strongly domain name registrars to ensure that they at least exert due diligence in verifying the legitimacy of their customers. Interestingly, market forces tend to lead to a certain degree of self-regulation, as shown by the McColo episode discussed earlier [18]. However, laws could be used to catalyze these market forces and lead to action before attacks occur regularly.

Such international laws, because they aim at the very core of the problem – poorly aligned economic incentives allowing cyberattacks to fester – would be very hard to pass, as they likely would mean a loss of revenue for many organizations without any counterpart. Nevertheless, given the increased connectivity between critical functions, stronger incentives for improving security at all levels must be put in place. Failing that, networking technology may actually become a vector for disaster.

Acknowledgments

This paper was prepared as part of the Advanced Methods of Cooperative Security Program, with generous support from the John D. and Catherine T. MacArthur Foundation and the Yamamoto-Scheffelin Endowment for Policy Research. This work greatly benefited from interactions with the Center for International and Security Studies at Maryland (CISSM). In particular, discussions with John D. Steinbruner and additional editorial suggestions and changes offered by Jonas E. Siegel considerably improved this manuscript.

References

[1] Guidelines & procedures, including the initial elements, 2007. Available online at http://www.wassenaar.org/guidelines.

[2] M. Abadi. Logic in access control. In *Proc. 18th IEEE Annual Symposium on Logic in Computer Science*, Ottawa, ON, Canada, June 2003.

[3] M. Abadi, M. Burrows, B. Lampson, and G. Plotkin. A calculus for access control in distributed systems. *ACM Transactions on Programming Languages and Systems*, 15(4), September 1993.

[4] B. Acohido. Are there 6.8 million – or 24 million – botted PCs on the Internet? http://lastwatchdog.com/6-8-million-24-million-botted-pcs-internet/. Last accessed September 16, 2010.

[5] Saurabh Amin, Xavier Litrico, S. Shankar Sastry, and Alexandre M. Bayen. Stealthy deception attacks on water scada systems. In *Proceedings of the 13th ACM international conference on Hybrid systems: computation and control*, HSCC '10, pages 161–170, Stockholm, Sweden, 2010.

[6] R. Anderson. *Security Engineering: A Guide to Building Dependable Distributed Systems*. Wiley Computer Publishing, New York, NY, 2 edition, 2001.

[7] C. Arthur. Plan to store Britons' phone and internet data revived, October 2010. http://www.guardian.co.uk/technology/2010/oct/20/internet-phone-data-plan-revived.

[8] R. Bace and P. Mell. Intrusion detection systems, 2001.

[9] S. Baker. Re: Lieberman bill lets president take emergency control of the Internet. Post made to the Interesting People mailing-list. Archive available online at http://seclists.org/interesting-people/2010/Jun/53.

[10] R. Boehme and T. Holz. The effect of stock spam on financial markets, April 2006. Working paper available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=897431.

[11] G. Bowley. Lone $4.1 billion sale led to 'flash crash' in may, October 2010. https://www.nytimes.com/2010/10/02/business/02flash.html.

[12] D. Carr. Military aims to collapse networks, maintain security. *Defense Systems*, August 2009. http://www.defensesystems.com/Articles/2009/07/29/Cyber-Defense-Encryption.aspx.

[13] Sang Kil Cha, Iulian Moraru, Jiyong Jang, John Truelove, David Brumley, and David G. Andersen. SplitScreen: Enabling efficient, distributed malware detection. In *Proc. 7th USENIX NSDI*, San Jose, CA, April 2010.

[14] R. Chabukswar, B. Sinopoli, G. Karsai, A. Gianni, H. Neema, and A. Davis. Simulation of network attacks on SCADA systems. In *Proceedings of the First Secure Control Systems Workshop, Cyberphysical Systems Week*, Stockholm, Sweden, April 2010.

[15] N. Christin. Peer-to-peer networks: Interdisciplinary challenges for interconnected systems. In M. Dark, editor, *Information Assurance and Security Ethics in Complex Systems: Interdisciplinary Perspectives*, 2010.

[16] N. Christin, S. Egelman, T. Vidas, and J. Grossklags. It's all about the Benjamins: Incentivizing users to ignore security advice. In *Proceedings of IFCA Financial Cryptography'11*, Saint Lucia, March 2011. To appear.

[17] N. Christin, S. Yanagihara, and K. Kamataki. Dissecting one click frauds. In *Proc. ACM CCS'10*, Chicago, IL, October 2010.

[18] R. Clayton. How much did shutting down McColo help? In *Proceedings of the Sixth Conference on Email and Antispam (CEAS)*, July 2009.

[19] Council of Europe. Convention on cybercrime, November 2001. `http://conventions.coe.int/Treaty/EN/Treaties/Html/185.htm`.

[20] R. Dhamija, J. D. Tygar, and M. Hearst. Why phishing works. In *Proc. ACM CHI*, pages 581–590, April 2006.

[21] R. Dingledine, N. Mathewson, and P. Syverson. Tor: The second-generation onion router. In *Proceedings of the 13th USENIX Security Symposium*, August 2004.

[22] D. Dittrich and S. Dietrich. P2P as botnet command and control: a deeper insight. In *Proceedings of the Third International Conference on Malicious and Unwanted Software (Malware)*, October 2008.

[23] N. Falliere, L. Murchu, and E. Chien. W32.Stuxnet dossier, November 2010. `http://www.symantec.com/content/en/us/enterprise/media/security_response/whitepapers/w32_stuxnet_dossier.pdf`.

[24] Laura Frieder and Jonathan Zittrain. Spam works: Evidence from stock touts and corresponding market activity. *SSRN eLibrary*, 2007.

[25] K. Jackson Higgins. Security guard busted for hacking hospital's HVAC, patient information computers, July 2009. `http://www.darkreading.com/insider-threat/167801100/security/privacy/218300006/index.html`.

[26] Vinay M. Igure, Sean A. Laughter, and Ronald D. Williams. Security issues in scada networks. *Computers & Security*, 25(7):498 – 506, 2006.

[27] S. Kandula, D. Katabi, M. Jacob, and A. Berger. Botz-4-sale: Surviving organized DDoS attacks that mimic flash crowds. In *Proceedings of the 2nd USENIX Symposium on Networked Systems Design & Implementation (NSDI'05)*, pages 287–300, Boston, MA, May 2005.

[28] C. Kanich, C. Kreibich, K. Levchenko, B. Enright, G. Voelker, V. Paxson, and S. Savage. Spamalytics: An empirical analysis of spam marketing conversion. In *Proceedings of the Conference on Computer and Communications Security (CCS)*, Alexandria, VA, October 2008.

[29] Engin Kirda, Christopher Kruegel, Greg Banks, Giovanni Vigna, and Richard A. Kemmerer. Behavior-based spyware detection. In *Proceedings of USENIX Security*, pages 273–288, August 2006.

[30] J. Leyden. UK teenager accused of 'electronic sabotage' against US port – houston, we have a problem. *The Register*, October 2003. `http://www.theregister.co.uk/2003/10/06/uk_teenager_accused_of_electronic/`.

[31] J. Lieberman, S. Collins, and T. Carper. Protecting cyberspace as a national asset act of 2010, June 2010. U.S. Senate bill. Online at `http://hsgac.senate.gov/public/index.cfm?FuseAction=Files.View&FileStore_id=4ee63497-ca5b-4a4b-9bba-04b7f4cb0123|pdf`.

[32] P. Maymounkov and D. Mazières. Kademlia: A peer-to-peer information system based on the XOR metric. In *Proceedings of the 1st International Workshop on Peer-to-Peer Systems (IPTPS'02)*, pages 53–65, Cambridge, MA, February 2002.

[33] E. Mills. Report: Hackers broke into faa air traffic control systems, May 2009. http://news.cnet.com/8301-1009_3-10236028-83.html.

[34] J. Mirkovic, S. Dietrich, D. Dittrich, and P. Reiher. *Internet Denial of Service: Attack and Defense Mechanisms*. Pearson Education, Upper Saddle River, NJ, 2005.

[35] D. Moore, V. Paxson, S. Savage, C. Shannon, S. Staniford, and N. Weaver. Inside the Slammer worm. *IEEE Security and Privacy*, 1(4):33–39, July 2003.

[36] D. Moore, C. Shannon, and J. Brown. Code-Red: a case study on the spread and victims of an internet worm. In *Proceedings of 2nd ACM/USENIX Internet Measurement Workshop*, pages 273–284, Marseille, France, November 2002.

[37] T. Moore and R. Clayton. Examining the impact of website take-down on phishing. In *Proceedings of the Second APWG eCrime Researcher's Summit*, Pittsburgh, PA, October 2007.

[38] T. Moore and R. Clayton. Evil searching: Compromise and recompromise of internet hosts for phishing. In *13th International Conference on Financial Cryptography and Data Security*, Barbados, February 2009.

[39] T. Moore, R. Clayton, and R. Anderson. The economics of online crime. *Journal of Economic Perspectives*, 23(3):3–20, Summer 2009.

[40] T. Moore, R. Clayton, and H. Stern. Temporal correlations between spam and phishing websites. In *2nd USENIX Workshop on Large-Scale Exploits and Emergent Threats (LEET '09)*, Boston, MA, April 2009.

[41] T. Mrozek. Former boeing engineer sentenced to nearly 16 years in prison for stealing aerospace secrets for china. United States Attorney's Office Central District of California. Press release No. 10-027. http://www.justice.gov/usao/cac/pressroom/pr2010/027.html.

[42] James Newsome and Dawn Song. Dynamic taint analysis: Automatic detection, analysis, and signature generation of exploit attacks on commodity software. In *Proceedings of Network and Distributed Systems Security Symposium*, February 2005.

[43] V. Paxson. Bro: A system for detecting network intruders in real-time. *Computer Networks*, 31(23–24):2435–2463, December 1999.

[44] M. Roesch. Snort: lightweight intrusion detection for networks. In *Proceedings of the 13th USENIX Systems Administration Conference (LISA'99)*, Seattle, WA, November 1999.

[45] Ahmad-Reza Sadeghi and Christian Stüble. Property-based attestation for computing platforms: caring about properties, not mechanisms. In *Proceedings of the 2004 workshop on New security paradigms*, NSPW '04, pages 67–77, New York, NY, USA, 2004. ACM.

[46] Reiner Sailer, Xiaolan Zhang, Trent Jaeger, and Leendert van Doorn. Design and implementation of a tcg-based integrity measurement architecture. In *Proceedings of 13th Usenix Security Symposium*, San Diego, CA, August 2004.

[47] J. Saltzer and M. Schroeder. The protection of information in computer systems. *Communications of the ACM*, 17(7), July 1974.

[48] C. Saul. Beyond HIPAA: International operations may subject companies to additional privacy laws, September 2003. http://www.martindale.com/members/Article_Atachment.aspx?od=111407&id=34760&filename=asr-34762.pdf.

[49] C. Savage. U.S. tries to make it easier to wiretap the Internet, September 2010. http://www.nytimes.com/2010/09/27/us/27wiretap.html.

[50] S. Schoen. Trusted computing: Promise and risk. Technical Report 20031001, Electronic Frontier Foundation, October 2003. Available at http://www.eff.org/files/20031001_tc.pdf.

[51] Arvind Seshadri, Mark Luk, Elaine Shi, Adrian Perrig, Leendert van Doorn, and Pradeep Khosla. Pioneer: Verifying integrity and guaranteeing execution of code on legacy platforms. In *Proceedings of ACM Symposium on Operating Systems Principles (SOSP)*, pages 1–16, October 2005.

[52] E. Shi, A. Perrig, and L. Van Doorn. Bind: a fine-grained attestation service for secure distributed systems. In *Security and Privacy, 2005 IEEE Symposium on*, pages 154 – 168, May 2005.

[53] The Asahi Shimbun. Leaks spur splurge for new SDF computers, March 2006. http://www.asahi.com/english/Herald-asahi/TKY200603090159.html.

[54] F. Stajano and P. Wilson. Understanding scam victims: Seven principles for systems security. Technical Report UCAM-CL-TR-754, Cambridge University, August 2009. To appear in Communications of the ACM.

[55] S. Staniford, V. Paxson, and N. Weaver. How to 0wn the Internet in your spare time. In *Proceedings of the 11th USENIX Security Symposium*, pages 149–167, San Francisco, CA, August 2002.

[56] B. Stone-Gross, M. Cova, L. Cavallaro, B. Gilbert, M. Szydlowski, R. Kemmerer, C. Kruegel, and G. Vigna. Your botnet is my botnet: analysis of a botnet takeover. In *Proceedings of ACM CCS'09*, Chicago, IL, October 2009.

[57] Symantec Corp. W32.Netsky.B@mm, February 2004. http://www.symantec.com/security_response/writeup.jsp?docid=2004-021812-2454-99.

[58] Symantec Corp. W32.Sasser.Worm, April 2004. http://www.symantec.com/security_response/writeup.jsp?docid=2004-050116-1831-99.

[59] J. Trachtman. Global cyberterrorism, jurisdiction, and international organization. In M. Grady and F. Parisi, editors, *The law and economics of cybersecurity*. Cambridge University Press, 2006.

[60] N. Weaver and V. Paxson. A worst-case worm. In *Proceedings (online) of the Third Annual Workshop on Economics and Information Security (WEIS'04)*, Minneapolis, MN, May 2004. Available at http://www.dtc.umn.edu/weis2004/weaver.pdf.

[61] M. Williams. Was North Korea behind the DDOS attack? *PC World*, July 2009. http://www.pcworld.com/businesscenter/article/168219/was_north_korea_behind_the_ddos_attack.html.

[62] Y. Yasuhara. The myth of free trade: the origins of COCOM 1945–1950. *The Japanese Journal of American Studies*, 4, 1991.