At the boundaries of peer production: The organization of Internet security production in the cases of Estonia 2007 and Conficker

Andreas Schmidt

Delft University of Technology, Faculty of Technology, Policy and Management

Article prepared for Telecommunication Policy

DOI 10.1016/j.telpol.2012.02.001, http://www.sciencedirect.com/science/article/pii/S0308596112000249, http://www.journals.elsevier.com/telecommunications-policy/

Abstract

With the emergence of Internet based communication and collaboration, new forms of production have surfaced that are based on openness and non-proprietary resources. The paper analyses the role of open source and peer production elements in the response to the attacks on Estonian Internet services in 2007 and the Conficker botnet in 2008/2009.

While both cases cannot be classified as purely peer-produced security, the two cases of incident response examined here do show some of the characteristics of peer production. By applying certain institutional techniques, the communities balance their need for secrecy with their need to widely share information.

The paper concludes with an explanatory model for the observed results. For appropriate policy outcomes, it suggests greater consideration of the role of social production by researchers and designers of the organisation of Internet security.

Keywords:

Internet security governance, Internet security, Peer production/open source production, Estonia 2007 cyberattacks, Conficker botnet

1 Introduction

Internet security and the security of information systems are no longer narrow technological subjects, but are making news headlines and rising to the top of national and international policy agendas. Increasing numbers of botnets using ever more innovative methods have infected millions of computers worldwide, making them part of an attack-infrastructure. While independent figures on the economic damages of Internet-based crime are not available, these numbers seem to be rising (van Eeten & Bauer, 2008a).

Political response to these Internet-based threats, risks and vulnerabilities has been a mixture of increasing public awareness, private self-regulation, public-private cooperation, creating Internet security groups within traditional state-based security organisations, and fostering international incident-response exercises. The political rhetoric accompanying discussions of Internet security often highlights an alleged lack of response capabilities, institutions and organisations.

But the rise of the Internet not only brings new risks, but also new forms of collaboration among globally distributed teams and communities. The Internet enables individuals to collaborate on the production of intangible goods with virtually no initial capital investments and production costs other than the time and attention allocated by contributors. Ideal-type examples of this mode of production – called open source production (Weber, 2004) or peer production (Benkler, 2006) – are free/open source software (FOSS) projects like Linux and Apache or content production communities like Wikipedia. The characteristics of Internet security and the general advantages of peer production suggest that this production model is likely to be applied in the domain of Internet security. This research aims at providing insights into existing forms of Internet security production and investigates the role of open forms of collaboration therein. It examines the role of peer production in two incident response cases, the Estonian attacks of 2007 and the response to the Conficker botnet in 2008-2009.

The paper begins by operationalizing peer production based on five categories. In the ensuing empirical sections, for each incident, the attack will be briefly described, followed by a depiction of the response activities and an analysis of whether elements of peer production exist therein. An ensuing section summarises the empirical findings. The last section attempts to explain the empirical outcomes and discusses some policy implications of the results.

2 Theoretical approach

2.1 A new form of production for incident response?

Weber characterised open source software projects as a new way of organising production. Absence of participation barriers, voluntariness as a founding principle, and sharing of production resources based on the open-source ideology and a certain property regime found a model of "distributed innovation", which can be more effective than other types of networked collaboration such as platforms or subcontracting (Weber, 2004, p. 231-243). Benkler analysed projects other than FOSS and similarly identified "a new modality of organizing production" utilising changes to transaction costs facilitated by the rise of the Internet (Benkler, 2006, p. 60). This form of production, termed "commons-based peer production", is characterised as "radically decentralized, collaborative, and nonproprietary; based on sharing resources and outputs among widely distributed, loosely connected individuals who cooperate with each other without relying on either market signals or managerial commands" (Benkler, 2006, p. 60). For Weber and Benkler, this form of production is generic and not bound to a specific good such as software. But is it applicable in the field of Internet security, too?

Internet security comprises a plethora of aspects. It has political, technical, and economic dimensions, involves different threats, vulnerabilities, attack and defence technologies and harmed actors, all leading to potentially different conceptualisations of Internet security. This paper cuts through this diversity by focussing on the response to two Internet security incidents, the Estonian DDoS-attacks in 2007 and the Conficker botnet. Internet security incidents are events or developments, caused by human action or technical failure, which put at risk what actors have defined as pivotal to their Internet-related interests. An empirical analysis of these incidents highlights existing modes of providing Internet security.

Security production in the context of Internet security incidents can be defined as any process or activity that assists in the reestablishment of the status quo ante or that helps to mitigate the incident-driven deterioration of the Internet's functionality. Based on literature on IT operations and policing (Crawford, 2006; Kempa, Carrier, Wood & Shearing, 1999; Whitman & Mattord, 2007), incident responses non-exclusively encompass activities such as ongoing monitoring of

systems, detection of the incident, analysis of the ongoing incident and its causes, identification of attacking actors, inventing and applying new hard- or software-based mitigation technologies, the collection and distribution of information required for the aforementioned activities, reconfiguration of systems, developing and distribution of patches, sanctioning of optional perpetrators, or tweaking and developing the operational framework for incident responses.

2.2 Expecting new hybrids of social production

Peer production has been praised for its transformational effects on societal institutions towards increased openness and transparency (Benkler, 2006, p. 379) and on individual virtues such as liberty, creativity, benevolence, and sociability (Benkler & Nissenbaum, 2006). Learning about the limits and the applicability of peer production beyond empirical examples such as FOSS or Wikipedia is important in and of itself.

Beyond the question whether peer production is applied in incident response, there is already evidence to support the assumption that peer production is likely to be present in incident responses. First, the general applicability of open source production in the field of Internet security is already proven by the existence of websites such as phishtank.org, a user-generated phishing intelligence system. The website offers an open data feed with information on phishing attacks. Phishing samples and metadata on the attacks are entered by attacked parties into PhishTank's systems. In addition, the existence of "informal networks of trusted security professionals" has been claimed in other publications (Mueller, 2010; van Eeten & Bauer, 2008b, p. 27).

Second, the underlying transaction cost-based explanatory model of the peer production indicates that peer production could occur in incident response processes. Extending the seminal Coase argument on firms (Coase, 1937), Benkler states that actors will prefer social exchange and production mechanisms over markets, and hierarchies if transaction costs and the motivations of potential contributors allow for superior efficacy (Benkler, 2004b, p. 306). The widespread distribution of actors and technical artefacts involved in security incidents inevitably requires distributed collaboration. Internet security production, at least partly, resembles "decentralized information gathering and exchange" (Benkler, 2006, p. 373). The peer production/open source process offers advantages in costs and efficacy in coordinating such dispersed and advanced knowledge and managing distributed innovation. (Weber, 2004, p. 264-272) Furthermore, pre-

field-research talks with Internet security experts suggested that views dominant in FOSS communities with regard to sharing and collaboration are also shared there. So both motivational base and relative transaction costs appear to be favourable for the peer production approach.

However, the applicability of peer production in incident response would probably be diminished by the potential need for secrecy. Weber predicts altered modes of open source production. The definitional prerequisite for openness in open source/peer production, the motivational similarities among participants in FOSS projects and security production, and the traditional tendency towards secrecy in the domain of security are indeed "conflicting pressures" that "form an interesting (if not very well controlled) experiment for generalizing open source." (Weber, 2004, p. 271)

2.3 Operationalisation and data collection

Analysing whether an existing production system uses elements of peer production requires an operationalization thereof. A collaborative production system can be called peer production if it matches the following criteria: A) Collaboration among participating actors is highly decentralised; B) Resources and information required during the production process are shared ad-hoc and in an unrestricted way; C) Produced goods are non-proprietary and hence can be reused and adapted by anybody involved in the production system; D) Produced goods are shared and accessible in non-market-based ways; E) Collaboration among actors is not based on managerial or hierarchical commands.

Based on this operationalisation, peer production in incident response exists if at least some products necessary and specifically designed for incident response meet the aforementioned criteria, or if collaboration in either one or more of the aforementioned subprocesses follows the peer/open source production model.

Two major Internet security incidents, Estonian 2007 cyberattacks and the Conficker botnet, have been analysed to answer the question about the potential application of peer/open source production as a mode of security production. Data on the cases has been gathered by desk research and semi-structured qualitative interviews (n=22) with actors involved in incident handling in the respective cases. This material is used to identify core activities, products and services of the collaborative incident response.

3 Estonia 2007

3.1 The attack

In the midst of a major domestic conflict between the Estonian government with the Russian minority, and increasingly strained diplomatic relations with bordering Russia, Internet services of major Estonian organisations were exposed to unusually high Internet traffic starting on April 27, 2007. What was soon identified as a major Distributed Denial-of-Service (DDoS) attack, resulted in malfunction or non-availability of services such as websites, email or online banking of Estonian political bodies, media, ISPs, and banks.

On Russian-language web forums, instructions and scripts on how to harm Estonian servers were published, along with pleas to run those scripts at a certain point of time. In a later phase of the attacks, botnets replaced humans as the emitters of overloading web traffic. In contrast to the media frenzy and the contemporary political perception, the actual damages in monetary terms were rather mild. Of all critical infrastructures, bank websites were down for local customers for a mere 90 minutes. Web defacements of political institution websites caused some political embarrassment, but never endangered the state's governing capacities.

3.2 Providing security

The ongoing attacks were soon addressed by an Estonian security community of mainly operational Internet and IT security experts, who were supported by a wider international technical community. Response activities can be divided into three categories: discovery and monitoring, problem analysis, and mitigation.

Discovery, analysis and response to DDoS attacks and web defacements are part of the day-to-day activities in the operation of Internet services. By 2007, the existence of DDoS threats was common knowledge among security professionals; many techniques and common procedures had been developed for discovering and mitigating DDoS attacks.

DDoS attacks represent a form of seizing IT systems. Either an irregularly high number of data packets congest all the data pipes, overstretching the resources of network components, or data requests that exploit vulnerabilities bring down the servers providing Internet services. Detecting

these effects on the operational level is every-day IT operations. Discovering that a substantial part of Estonia's web-services were flooded with malicious distributed requests, however, required the exchange of information among Estonian operators of the organisations affected.

The analysis of the ongoing problem was not the most pressing issue of the response team, given the technical simplicity of DDoS attacks and their frequent appearance. What made the Estonian attacks in 2007 unique was that the Internet services of an entire country (albeit a small one with just 1.4m inhabitants) were simultaneously flooded with data packets and web requests, as well as the 3-week duration of the attacks.

Mitigating DDoS attacks requires the application of one or more response techniques. Upscaling servers, offering a temporarily stripped down website, granting or denying access to the website to certain ranges of IP addresses, increasing bandwidth between targets and their ISPs or backbones, routing DDoS traffic to sinkholes – all these techniques help to keep web services online.

A DDoS attack aimed at overstretching web-server capacities can be countered by a reconfiguration of components on the network perimeter of an organisation. For attacks flooding the network routes to an organisation's infrastructure a different defence approach is more promising: Malicious packets are dropped by any conveying intermediary between the attacking and the attacked ends. This approach requires collaboration with actors controlling parts of the Internet infrastructure that are conveying packets from the attacking systems to the target systems.

The second phase of attacks was based on botnets, with bot-infected drones scattered on machines located in numerous countries. Mitigation of botnets emitting DDoS packets is in the short run restricted to the aforementioned mitigation techniques. If a botnet is operational for weeks and is dedicated solely to a specific DDoS attack, the defending actors would probably want to take down the botnet itself or take over its command-and-control system. In the case of the Estonian attacks, however, the second botnet-based phase of the attacks lasted for too short a time.

Summing up, responding to the attacks mainly encompassed applying the aforementioned anti-DDoS techniques, providing an ad-hoc communication and data exchange platform, conveying

relevant information from the source (the target) to an intermediary with the ability to reduce malicious web traffic; that is, ISPs in Estonia and abroad.

3.3 Applying the framework

After this brief description of the incident and the response activities, the elements of the peer production framework described above are now applied to the case.

3.3.1 Decentralisation of collaboration

Effective countering of the attacks was mainly achieved by the Estonian Computer Emergency Response Team (CERT EE), the IT departments of affected Estonian organisations, and by local Internet Service Providers (ISPs), complemented by what is called the global Internet security community.

Collaboration among Estonian technical experts was partly ad-hoc, and partly followed pre-existing paths. Starting in the late 1990s, national Internet security experts and government bodies worked closely in task forces for securing the partly Internet-based Estonian national elections and local banks started sharing attack information.

The Estonian Internet security community was expecting an attack on Estonian Internet and IT services because of the public unrest. This community is based on a mix of professional and social ties, which allowed for an instant set-up of an incident-targeting ad-hoc task force, once the start of web attacks became apparent.

The task force mainly collaborated via Internet, though some physical meetings occurred in the course of the three weeks. CERT EE provided the infrastructure of collaborative tools. While these channels existed before the incident, the number of participants grew from some 30 to almost 200 in the course of the incident. Access to these communicational platforms was granted by CERT-EE. Applicants who wanted access to these platforms had to prove that mitigation of the problems would require their participation.

The inclusion of traditional security forces in the collaborative effort enabled the technical community to gain insights about future attacks planned on Russian forums. Security providers mentioned that they are regularly watching these forums, too.

The Estonian cause was supported by an international Internet security community. Technical experts have formed a wide number of permanent security discussion groups, usually based on e-mail lists. However, a detailed list of communities or global mailing lists involved in the response activities could not be compiled, as policies of some of these lists require their names to be kept secret. Access to these communities requires recommendations by one or more existing members and usually something to bring in by the candidate.

Even though these different international security communities have no administrative control over the Estonian technological infrastructure, many members offered support. Some Estonian experts were introduced to the wider international community on mailing lists or in physical meetings. The support of the international community allowed for moving the line of defence away from attacked systems closer to bot-infected drones by either taking down or blacklisting attacking systems.

3.3.2 Sharing of input resources

The second characteristic of peer production is unrestricted sharing of resources and inputs required for the core products of the response efforts. As described above, the main activities of security production were applying anti-DDoS techniques, providing a communication and data exchange platform, and conveying attack information to owners of Internet components.

As described in the previous section, sharing of input resources only happened within the boundaries of the communities involved. A second barrier to broad sharing lies within the community: attack data is usually only shared directly with the owners of certain IP address ranges. The overstretched CERT EE outsourced this task to a familiar third-country CERT.

Within Estonia, exchange of attack data happened via the platforms provided by CERT EE. This allowed local ISPs to block inbound traffic from specific sources. Later, the Estonian community switched from a whack-a-mole approach, which tried to address any micro-incidents within the larger countrywide incident, to a more all-encompassing approach of blocking certain traffic from abroad. While this approach made parts of the Estonian Internet infrastructure inaccessible from abroad, it helped to reduce downtimes for Estonian citizens significantly.

Effective security measures were eventually implemented by operational staff of infrastructure owners. The actual measures depended on the specific attack patterns and the given network

architecture, and thus were unique for every network. Nevertheless, sharing information and discussing solutions on the collaborative platforms provided by CERT EE eased the implementation of mitigating means.

3.3.3 Non-proprietarity of produced goods

The third characteristic of peer production is the non-proprietarity of the goods produced, that is their distribution and sharing is not stymied by legal access restrictions. On the contrary, FOSS communities enforce openness and disallow private, exclusive appropriation of their products with specific copyright licences. The proprietarity of security production differs among respective incident response activities.

Applying anti-DDoS techniques only alters existing infrastructures by reconfiguring existing systems, replacing or supplementing them with new components; it doesn't alter the ownership of existing technological artefacts. These reconfigurations result in systems that are more resilient against DDoS attacks and allows service providers and users to continue to offer and receive assured services. The re-established security is as exclusive or non-exclusive as the services it affects.

Discussing mitigation techniques, configuration scripts and the like creates new information that is openly shared within the community, which also implies the exclusion of non-members. The recipes how to re-establish assured system functionality of partly proprietary services are openly shared within the community and freely reusable and alterable. However, sharing of enhanced knowledge and tools built on previously shared items is not enforceable by the community. It lacks institutions akin to the General Public License used in Open Source communities to enforce re-sharing.

The same holds true for CERT EE providing an access-controlled communication and data exchange platform. It is a non-commercial service, a combination of software, hardware and configuration settings. Even though the latter two are proprietary, the communication and exchange platform could be emulated at probably modest cost given the low price tag for hosting, the possibility of scraping existing content and transferring it into a new system. The emulatability of communication and exchange services prevents its provider to exert authority based on the ownership and technical control of these systems.

From a birds-eye perspective, the entirety of counter-incident efforts of the community resemble a service to re-establish a secure state of the Internet, characterised by the absence of loss of data integrity, confidentiality and functionality of web-based services. Internet security conceptualised in this manner resembles a by definition non-appropriatable public good.

3.3.4 Non-marketness of production

The fourth characteristic of commons-based peer production is that the production of a good is not based on market incentives, that is persons involved in the production process don't seek monetary compensation for their time spent in the first place.

Most interviewees stated that their core motivation was to do good, and to assist in securing the Internet. Some of the Estonian interviewees also stated that they wanted to protect their country and its technological infrastructure. Another motivation mentioned for their involvement in the mutual support system was reciprocity; the fear that they could be attacked and would then depend on the help of others.

3.3.5 Absence of managerial command or hierarchies

One could argue that most of the persons involved in the Estonian case are employees of companies affected by the DDoS attacks and their job description most likely includes some degree of responsibility for their organisations' networks. The commitment of persons from unaffected third countries, on the other hand, cannot be explained as a product of job requirements. It appears unlikely that a technical expert in a third country would have to fear repercussions within his company for not supporting mitigation efforts in a remote country.

Even within Estonia, the role of hierarchies in explaining individual motivation is ambiguous. The technical experts involved in the case enjoyed a great deal of operational freedom. The question to interviewees what would happen if they dropped out and stopped their engagement in security communities was along the lines that: the bad guys would win, it would be an Internet full of botnets, spam and malware, no one could step in. Despite this perception one could argue that the informal community would be replaced by a mix of inter-organisational agreements and legal provisions as soon as the community's output would fall below what is rendered as politically acceptable. In addition, sharing of information is quite often restricted by statute.

4 Conficker

4.1 The attack

In late 2008, a Computer worm later called Conficker infected millions of computers, exploiting a vulnerability in components of the Microsoft Windows operating system (National Institute of Standards and Technology, 2011). It was one of the first pieces of malware that came with self-installing, self-propagating features like the worms in the late 1990s (Symantec, 2009, p. 1), turning computers into drones of a botnet. In the underground economy, botnets provide technological platforms for criminal activities.

In an attempt to pre-empt the application of simple defence techniques, the botherders avoided static URLs for their command and control (CnC) servers by using an algorithm to compute some 500 domain names every day. Applying a technique called *HTTP rendezvous* (Microsoft, 2009, p. 96), the botherder uses the same algorithm, registers these domains, prepares the CnC mechanisms on these servers and thus enables the bots to establish a connection with their CnC servers. Conficker C further decreased the dependency on central CnC mechanisms. It came with an enhanced peer-to-peer (P2P) technique and an altered domain computing algorithm, which used a vastly increased pool of 50 000 potential domain names per day. (For descriptions of technical details see Microsoft, 2009, p. 95-99)

Conficker has so far only been used for installing rogue security software on an unknown number of PCs (Microsoft, 2009, p. 99). Apart from the cleaning costs, the actual damages appear to be relatively low, though no exact figures exist here. Without the joint collaboration of the Internet security community, Conficker would have become a fully operational botnet with several million drones, ready to be used for criminal activities. Currently, communication between bots and CnC servers is effectively interrupted. However, some 1 to 3 millions of computers remain infected with the Conficker malware (Hogben, Plohmann, Gerhards-Padilla & Leder, 2011, p. 28). Most of them are located in BRIC countries.

4.2 Providing security

Re-establishing a secure state after the creation of a botnet involves cleaning the infected computers and dismantling or containing the attack capabilities of the botnet.

There are a number of general anti-botnet strategies (Hogben et al., 2011; Stanković & Simić, 2009). The first fundamental approach is to disinfect affected PCs and harden non-infected PCs. To achieve this, security patches for exploited, vulnerable software components need to be updated. In addition, and given the existing structure of Windows PC security, signature files for security products like AV software or intrusion detection systems (IDS) need to be updated. Theoretically, this strategy would be sufficient if applied to any PC connected to the Internet worldwide. The second fundamental approach is to interrupt the communication between the bots and the botherder, rendering a botnet non-operational and economically useless.

Major response activities against the Conficker botnet included providing updates for software and AV or IDS signatures, analysing the malware, and collecting and analysing the botnet's traffic.

Microsoft released a critical security update (Microsoft, 2008) before Conficker started exploiting it a month later in November 2008. Security vendors had their signature updates in place soon after Conficker emerged. However, even though updates had been available in a timely manner, many infrastructure owners hadn't applied them months after their respective releases. This usual pattern among users renders anti-botnet approaches useless that only focus on security updates requiring manual user intervention.

Analysis of the fundamental design of the botnet includes analysis of the activities of the binary on the infected machines and the communication of the binary with remote computers on the Internet. Reverse engineering of the malware's binary code often is the only way for the responding side to understand the techniques of the botnets so that they can design appropriate mitigation techniques. When responding actors eventually figured out logic of the botnet, they aimed at cutting the communication lines between the millions of bots and their CnC servers.

To understand the activities of the bots and the size of the botnets, so called sinkholes were created. Sinkholes are central repositories storing traffic data sent from the bots to alleged *HTTP rendezvous* domains registered and controlled by the responding actors. These sinkhole databases contain information about Conficker infected machines and their attempts to contact alleged rendezvous domains.

A take-down procedure for all CnC domains and servers were not feasible with a conventional organisational approach, as the bot-herders used *HTTP rendezvous* with first 500 (Conficker A and B) and then 50 000 (Conficker C) potential domains per day to hand over new instructions to the bots. Applying a response approach never applied before on such a scale, the community successfully started seizing the domains before the bot-herders could register them. The vast majority of Conficker A/B hasn't been able to contact *HTTP rendezvous* servers, download the latest payload and upload itself to a more recent variant (Giles, 2009).

4.3 Applying the framework

Analogous to the Estonian case, this section applies the elements of the peer production framework to the Conficker case. It provides an analysis whether the five characteristics of peer production can be found in the major activities of the anti-Conficker endeavour, that is, security updates, malware analysis, sinkholing and domain registration.

4.3.1 Decentralisation of collaboration

The relevant work against the Conficker incident was the result of a collaboration of security providers, the vendor of the affected OS platform, individual and corporate Windows users, academic and non-academic security researchers, ICANN, registry operators responsible for some 110 Top-Level Domains (TLDs), and independent affiliations of Internet security experts. The bulk of the work was carried out by an ad-hoc collaborative network, which called itself the Conficker Working Group (CWG) and comprised all of the aforementioned actors minus users, country-codes TLD registry operators and some independent researchers.

In the beginning, the response activities were far less coordinated. From October 2008 to February 2009, from the discovery of the vulnerability until the establishment of the CWG, the response to Conficker was loosely coupled, based on discussion in several mailing lists. (Compare Piscitello, 2010; Rendon Group, 2011) However, the innovative combination of P2P botnet communication and *HTTP rendezvous* required new and unprecedentedly intense collaboration not only among the security community, but also with actors controlling decisive elements of the Internet.

Social ties in the Internet community led to a meeting in mid-February 2009 which would form the core of the CWG (Rendon Group, 2011, p. 18), a virtual ad-hoc task force that is neither

affiliated with an existing organisation nor a legal entity by itself. Characteristics of the group are its consensual governance model, the establishment of a managing core group and a division of labour by a number of more or less autonomously operating subgroups. Some 250-300 members have used the collaborative systems ad-hoc set up by a volunteer group of security experts (Rendon Group, 2011, p. 21). Access is granted if at least two existing members supported the admission of a new member and no opposition could be observed.

By expanding and reorganising the response team, the challenges posed by the technological features of the malware could be met. At the same time, the anti-Conficker response remained accessible for non-members.

4.3.2 Sharing of input resources

Sharing of the informational resources necessary for providing security varied among the different activities of security production. Predominantly, it has only been made accessible within the restricted, non-incident related security communities, and, since its foundation, within CWG and its subgroups.

For updates of vulnerable software and malware definition files, it is common usage in the security community to share binaries of malware. Tools and techniques for analysing malware, their functionality and logic however remain to a substantial extent proprietary; Anti-Virus companies are building up patent libraries of techniques and technologies in their domain.

Relatively easy access to malware binaries and low in-advance capital cost allows anyone interested to analyse the Conficker malware; time and high expertise are the only prerequisites.

Sinkholing per definition requires access to traffic data of bots and – in the case of Conficker – rendezvous domains. Several actors in the group engaged in sinkholing independently, and most security vendors have such facilities in place. Groups of security experts privately designed and operated a sinkhole system for Conficker, which stores millions of records of traffic between Conficker bots and alleged rendezvous domains.

4.3.3 Non-proprietarity of produced goods

In a collaborative response endeavour, in which activities are bundled into interdependent work domains, one such sub-domain requires the outputs of another sub-domain as its inputs. As an

example: domain name registries need lists of Conficker domains, which are the outcome of reverse-engineering of the respective algorithm by malware analysts, while these analysts need access to malware binaries collected by honeypots and might also need information gathered by sinkholes for clues about the malware logic.

Security updates were as proprietary as they could possibly be: developed and released by Microsoft as compiled Windows executables. Only Microsoft has access and the legal right to alter Windows source code. Probably related to Microsoft's anti-counterfeiting efforts, the bulk of the approximate 3 million infected Windows computers is still located in piracy-prone BRIC countries (Shadowserver Foundation, 2011).

Similarly proprietary are the signature updates provided by AV and security vendors. Notable exceptions are the rules for intrusion detection systems (IDS) provided by two junior academics under a Creative Commons license (Leder & Werner, 2009, p. 8). Using them, IDSs can detect traffic in their networks originating from Conficker bots.

The analysis of the Conficker malware and the botnet has probably been the most open element in the whole response endeavour. Relevant papers were published by SRI International, a non-profit research institute, and the Honeynet Project and Research Alliance, a non-profit organisation with voluntary researchers.

A particular element of knowledge derived from botnet analysis, however, remained inside the circles of the security community and was not published for the public at large: how to remotely take over the bots by exploiting a vulnerability of the Conficker malware. The researchers thereby followed the responsible disclosure policy, which states that information shall not be widely published if it allows malware authors to easily commit illicit hacks. Any security expert that does not follow this principle risks reputational damage in the overall community, with possible consequences for their and their introducers' access to vetted mailing-list-based communities.

A similar mix of openness and closure can be observed with sinkholing of botnet traffic. Sinkholing was performed by several actors independently. After the establishment of the CWG, actors involved agreed on a unified sinkhole under the auspices of the Shadowserver Foundation. The collected traffic data has been shared with anybody requesting for it, a positive background

check by the community provided. Network operators with infected machines received lists of affected URLs. Control of the overall systems that collect, compute and analyse sinkholed data remain with a group of voluntary experts.

Ownership of Conficker domains resides with registry operators; control over them isn't of much use other than directing traffic to sinkholes. Malevolent actors could make use of Conficker sinkholes and their data as they reveal vulnerabilities.

4.3.4 Non-marketness of production

The production of security updates happened within the market system. Conficker was the first worm in years and had the potential to turn into a similar disaster as the worms in the early 2000s. Microsoft therefore had a significant incentive to bring the Conficker worm under control. They were one of the drivers for connecting the relevant actors to join forces. AV vendors' activities likewise appear to market-driven.

The incentives for many of the remaining actors, especially those involved in studying the malware and sinkholing traffic data, have been different and did not follow direct pecuniary calls. Most interviewees responded that their participation and voluntary contribution of time and occasionally even money (e.g. for privately registering domains) was based on the motivation to make the Internet more secure, do something against the criminals or to address an interesting itch to scratch. These motivations are very similar to those in open source communities.

Most of the sinkholing is based on voluntary contributions, just like so many efforts in the Conficker response. Sinkhole data is provided to interested actors in the community for free.

4.3.5 Absence of managerial commands or hierarchies

Inside the CWG, no particular actor or organisation had authoritative control over the activities of other actors. While the core group could steer the CWG into certain directions, its internal decision making process was based on unanimity, and their leadership appears to be accepted by non-core group members as it increased the effectiveness and efficiency of the group.

But social norms influenced the behaviour of people involved in the response. As described in the section before the last, one of the research papers analysing Conficker was published in two versions, one for the public and one for the restricted community. The institution of responsible

disclosure affected the behaviour of contributors, akin to an implicit command by the community.

The domain registration process was characterised by the networked, non-coercive approach. Most of the global registry operators complied with ICANN's requests to pre-emptively register or take down Conficker domains. Only a few maintained a legalist stance, insisting on court orders before taking down Conficker domains. (Piscitello, 2010, p. 9) Chinese governmental registry operators received lists from mostly U.S.-based security experts, who they had never met or heard of before, to block or take down .cn Conficker domains. This process required some degree of trust on both sides. (See also Rendon Group, 2011, p. 19)

Some interviewees complained about a lack of authority in the field of Internet security in general. They view Conficker as a botnet that might still be used for substantial criminal activities in the future, and would prefer the botnet to be destroyed by remote vaccination of infected machines. However, there is no decision-making or technical advisory body that could legitimately decide to virtually break into millions of computers to install *beneware* or healing fixes onto them.

5 Conclusion: Social production of Internet security

The aim of this empirical analysis has been to identify the role of peer production in Internet security incident response. It is clear that in both of the cases studied, incident response did not consistently rely on all five of the elements of peer production. Some aspects of the five elements, however, were present in both cases. Table 1 summarizes the results.

In the Estonian case, the overall incident response activity was by and large a *decentralized* effort, with specific tasks located at certain actors. Task sharing was negotiated or emerged ad-hoc and did not follow the ruling of any sort of central decision-making body. Some degree of centrality can be found in the role of CERT EE as information hub. For the Conficker incident, the response was a collaborative endeavour by a wider and more diverse set of actors compared to the Estonian case. Security production first happened somewhat independently, coordinated via mailing lists. It later merged into a virtual ad-hoc organisation, complemented by some

external actors. The Conficker response was far more global in nature; the bulk of the Estonian response was mostly of local origin.

Restrictions on *sharing of input resources* – information in this case – were significant in both cases. The first kind of barrier encircles the Estonian ad-hoc incident community and the wider global mailing-list communities. A second barrier is the community-wide rule to share sensitive data only directly and on a need-to-know basis. For the Conficker case, access to informational input resources was often restricted to vetted community members. Other than in the close social circles of the Estonian case, information was occasionally shared between actors who had not built up trust-based relationships before. Some aspects of the security production entirely relied on proprietary, non-shared resources.

As to the characteristic of *non-proprietarity*, the results are mixed in several dimensions. First, one could argue that in both cases the response produced an increase of reliability, a reduction of vulnerability of Internet components and thus a reduced risk of damages for Internet users; that is, an increase in Internet security. This increased security has all the characteristics of a public good. Second, the security increase was achieved by a number of intermediary products as described in the empirical sections. In the Conficker case, there are proprietary OS and signature updates on the one hand and relatively open analysis and research on the other.

The last two characteristics, the *absence of markets and hierarchies* in the organisation of production, are related as they picture the motivation of an individual contributor. In both cases, contributors acted voluntarily, beyond their job descriptions and uncompensated. The categorisation can be methodologically tricky when the lines of professional life and private interest blur. When non-participation in the community effort would probably have led to negative professional repercussions for contributors, the shadow of hierarchy (Scharpf, 1997) or the hidden hand can be assumed as motives. Based on that, the response within Estonia was at least to some extent influenced by their professional responsibilities, whereas motivations for contributions from the international IT security communities were mostly intrinsic. In the

Conficker case, the situation is slightly different, as the aforementioned differences between the

Table 1. Elements of peer production in the two incidents

	Estonia 2007	Conficker
Decentrality	Yes	Yes
Input resources sharing	Hybrid (access layers)	Hybrid (access layers)
Non-proprietarity	Hybrid	Hybrid (updates – research)
Non-marketness	Yes	Hybrid
Non-hierarchicalness	Hybrid	Yes

production of updates, malware analysis and bot-traffic data-warehouses indicate.

6 Discussion: Explanations and policies

6.1 Explaining social production of Internet security

As the empirical sections have shown, the overall incident response is the result of mixed activities pursued in the market, hierarchical and social frameworks. The study does not reveal hard quantitative figures about the overall relevance of social production as compared with the deliveries of markets or hierarchies. Nevertheless, qualitatively important contributions were based on social exchange mechanisms.

The results, the existence of social production and of non-pure elements of peer production, can be explained by the analytical model as developed by Benkler and Weber. According to their reasoning, the selection of peer- or open source production over market-, firm- or state-based frameworks is the result of superior effectiveness of social sharing, facilitated by favourable "motivational effects and … transaction costs" (Benkler, 2004b, p. 277).

Neither the data nor the research approach of this study allow for exact statements about the causes of the actual shape of the production model. One could however hypothesize that the observed instances of the social production model resemble an institutional design that tries to incorporate some major advantages of the ideal-type peer production/open source model, while

at same time factoring in the need for secrecy. In terms of transaction costs, the observed hybrid, not-quite-peer-production flavour of social production reduces the risks of intrusion by malevolent bad guys, who seek to nullify the communities' defence efforts. On the other hand, it keeps transaction costs of secrecy relatively low by using community-based vetting procedures and remaining somewhat permeable to new contributors. The open source access-for-all policy and the swift trust model, which is based on assumed common values and goals (Adler & Heckscher, 2006, p. 24; Osterloh & Rota, 2004, p. 13-16), is replaced by a system of limiting access to its infrastructures, vetting potential contributors and sharing on a need-to-know basis only. While secrecy thwarts one source of peer production effectiveness – the unplanned, unrestricted use of resources by high numbers of agents with diverse talents and skills – security communities can still leverage relatively low-cost permeability to new contributors to take advantage of external "information gains" (Benkler, 2002, p. 407-423).

Another, somewhat related explanation for the existence of some elements of social production in incident response lies in the set-up costs and marginal transaction costs for non-social production frameworks. From an organisational and institutional perspective, effective incident response is a problem of gathering data, analysing information, and coordinating action worldwide among numerous actors. Market-based production requires a high degree of crispness of the products and services to be delivered. This also holds true to a lesser extent for firm- or state-based provisioning. It takes time and substantial resources to gather the information required for setting up such frameworks. While traditional security-providing actors were still getting their act together in 2007 and 2009, a global community of experts with mixed motivations – acting as volunteers, as geeks, as employees, as entrepreneurs, as governmental clerks – had already emerged to address existing security issues. In the long run, the social exchange might phase out once traditional security actors are willing to invest in the set up and maintenance costs for a mainly market- and state-based security provisioning. Alternatively, traditional security institutions might be content to merely gain a decisive role in a security network that still contains elements of social production.

6.2 Implication for Internet security policies

Policy-wise, the results of this study suggest reconsideration of some previous assumptions about Internet security. First, contrary to frequent statements in policy discourses, life in Estonia hardly

came to a standstill in 2007. The most severe impact of the attacks probably was the disabling of the online services of Estonia's largest bank for some 90 minutes, and limited functionality for several hours. (See also Ottis, 2009)

Second, when Internet security regulators aim at transforming the institutional space of Internet security provisioning by introducing monetary or hierarchical incentives, the risk of crowding out intrinsic motivations (Benkler, 2004a, p. 321-326) needs to be considered. Instead of being complementary, these incentives can demotivate potential contributors and thereby stall voluntary cooperation. In the incidents described, security production relied on some elements of peer production. Hence, adding regulations that force previous contributors to behave in certain ways might actually decrease security when incentives and motivations are not properly aligned.

Third, contributions by social communities can be deterred. While there are community-driven initiatives nurtured by intrinsic motivation for botnets and phishing, no such effort has emerged against child abuse imagery. As a result, the average take-down time of child abuse imagery is roughly 719 hours, compared to 4.3 hours for phishing websites (Moore et al., 2009, p. 16). Moore explains these take-down failures by noting that national police were given a monopoly for countering abuse imagery, police rarely cooperate internationally, and the monopoly thwarted businesses from addressing this issue. One could hypothesize that the rigorous anti-abuse-imagery legislation combined with harsh social sanctions on anybody suspected to have had contact with such imagery, has disincentivised community-driven initiatives to overcome the problem. The abuse imagery takedown failure is a showcase of how Internet security might suffer if security communities are entirely superseded by traditional governmental institutions.

Fourth, the efficacy of social production elements within the overall response can, at least theoretically, be increased by raising new opportunities for openness and by installing a standing infrastructure to support ad-hoc incident response activities. A community that defaults to secrecy loses one of the key efficacy drivers of peer production: that every person can chose the area and work package she wants to contribute. Defaulting to security was partly driven by efficiency considerations and lack of resources to evaluate the advantages of secrecy on a more granular level. Granularly opening security data and collaborative communities might lead to the kind of innovations that are driving high-level Open Data policies in countries such as France (data.gouv.fr), the United Kingdom (data.gov.uk), the United States (data.gov) and the European

Union (du Preez, 2011; Huijboom & Van den Broek, 2011). This certainly increases the set-up costs for organising response in general, but could decrease marginal set-up costs of ad-hoc organisations for dedicated incidents and therefore decrease transaction costs for peer-producing initiatives.

Last but not least, Internet security has to some extent the characteristics of a public good. Proponents of economics of Internet security have frequently stated that it is because of this characteristic regulatory intervention to overcome the mismatch between incentives to act and externalities of non-acting is required. (Moore et al., 2009, p. 9) It would be this alleged tragedy-of-the-commons-like situation that requires regulators to go after control points (Moore et al., 2009, p. 9). However, social production relies on motivations of potential contributors. These motivations do not necessarily match the incentives of ISPs, software vendors, or registrars. While an ISP might only have little or at best modest incentives to do something about malware-infected client computers (van Eeten & Bauer, 2008b), a security engineer can regard a botnet as a personal challenge. Therefore, the incentive-based models of cyber-security activities need to be supplemented by the social dimension of the Internet security communities. The wealth of networks addresses some of the risks of computer networks.

7 References

Adler, P. S., & Heckscher, C. (2006). Towards collaborative community. In P. S. Adler & C. Heckscher (Eds.), *The firm as a collaborative community: Reconstructing trust in the knowledge economy* (pp. 11-106). Oxford, New York: Oxford University Press.

Benkler, Y. (2002). Coase's penguin, or, Linux and the nature of the firm. *Yale Law Journal*, *112*(3), 367–445. Retrieved from http://www.yalelawjournal.org/images/pdfs/354.pdf

Benkler, Y. (2004a). Peer production of survivable critical infrastructures. Proceedings of the Telecommunications, Policy, and Research Conference held at George Mason University Law School in Arlington, Virginia, Oct 1-3, 2004. Retrieved from http://web.si.umich.edu/tprc/papers/2004/340/Benkler%20Critical%20Infrastrcutures.pdf

Benkler, Y. (2004b). Sharing nicely: On shareable goods and the emergence of sharing as a modality of economic production. *Yale Law Journal*, *114*(2), 273–359.

Benkler, Y. (2006). The wealth of networks: How social production transforms markets and freedom. New Haven, London: Yale University Press.

Benkler, Y., & Nissenbaum, H. (2006). Commons-based peer production and virtue. *Journal of Political Philosophy*, *14*(4), 394–419.

Coase, R. H. (1937). The nature of the firm. *Economica*, *4*(16), 386–405. doi:10.1111/j.1468-0335.1937.tb00002.x

Crawford, A. (2006). Networked governance and the post-regulatory state? Steering, rowing and anchoring the provision of policing and security. *Theoretical Criminology*, *10*(4), 449. doi:10.1177/1362480606068874

du Preez, D. (2011, December 12). European commission launches open data strategy. *computing.co.uk,* Retrieved from http://www.computing.co.uk/ctg/news/2131718/european-commission-launches-strategy-europe

Evron, G. (2008). Battling botnets and online mobs. Estonia's defence efforts during the internet war. *Georgetown Journal of International Affairs*, 9(1), 121–126.

Giles, J. (2009, June 12). The inside story of the conficker worm. *Newscientist*. Retrieved from http://www.newscientist.com/article/mg20227121.500-the-inside-story-of-the-conficker-worm.html

Herzog, S. (2011). Revisiting the estonian cyber attacks: Digital threats and multinational responses. *Journal of Strategic Security*, *4*(2), 4.

Hogben, G., Plohmann, D., Gerhards-Padilla, E., & Leder, F. (2011). *Botnets: Detection, measurement, disinfection & defence*. European Network and Information Security Agency (ENISA). Retrieved from http://www.enisa.europa.eu/act/res/botnets/botnets-measurement-detection-disinfection-and-defence/at_download/fullReport

Huijboom, N., & Van den Broek, T. (2011). Open data: An international comparison of strategies. *European Journal of Epractice*, *12*(March/April 2011). Retrieved from http://www.epractice.eu/en/document/5290090

Kempa, M., Carrier, R., Wood, J., & Shearing, C. (1999). Reflections of the evolving concept of 'private policing'. *European Journal on Criminal Policy and Research*, *7*(2), 197–223. doi: 10.1023/A:1008705411061

Landler, M., & Markoff, J. (2007). In Estonia, what may be the first war in cyberspace. *International Herald Tribune*. Retrieved from http://www.iht.com/articles/2007/05/28/business/cyberwar.php

Leder, F., & Werner, T. (2009, March 30). *Know your enemy: Containing Conficker. To tame a malware* (rev1 ed.) (The Honeynet Project). Retrieved from http://www.honeynet.org/files/KYE-Conficker.pdf

Microsoft (2008, October 23). Microsoft security bulletin MS08-067. *Microsoft TechNet* (Version 1.0) [Web page]. Retrieved from http://www.microsoft.com/technet/security/bulletin/ms08-067.mspx

Microsoft (2009, November). *Microsoft security intelligence report* (Vol. 7, January - June 2009). Microsoft. Retrieved December 29, 2009, from http://www.microsoft.com/sir

Moore, T., Clayton, R., & Anderson, R. (2009). The economics of online crime. *Journal of Economic Perspectives*, *23*(3), 3-20.

Mueller, M. (2010). Networks and states: The global politics of Internet governance. Cambridge, Mass.: MIT Press.

National Institute of Standards and Technology (2011). *Vulnerability summary for CVE-2008-4250*. National vulnerability database, national cyber-alert system [Web page]. (Original work published October 23, 2008) Retrieved from http://web.nvd.nist.gov/view/vuln/detail?vulnId=CVE-2008-4250

Osterloh, M., & Rota, S. (2004). Trust and community in open source software production. *Analyse & Kritik*, (26), 279–301.

Ottis, R. (2009). *Conflicts in cyberspace: Evgeny Morozov on cyber myths*. Retrieved from http://conflictsincyberspace.blogspot.com/2009/06/evgeny-morozov-on-cyber-myths.html

Piscitello (2010). *Conficker summary and review*. ICANN. Retrieved from https://www.icann.org/en/security/conficker-summary-review-07may10-en.pdf

Rendon Group (2011). *Conficker working group: Lessons learned* (Report created in June 2010, commissioned by the Department of Homeland Security). Retrieved from http://www.confickerworkinggroup.org/wiki/uploads/Conficker_Working_Group_Lessons_Learned_17_June_2010_final.pdf

Scharpf, F. W. (1997). Games real actors play: Actor-centered institutionalism in policy research. Boulder, Colorado: Westview Press.

Shadowserver Foundation (2011). *Conficker*. [Web page] Retrieved from http://www.shadowserver.org/wiki/pmwiki.php/Stats/Conficker

Stanković, S., & Simić, D. (2009). Defense strategies against modern botnets. International Journal of Computer Science and Information Security, 2(1). Retrieved from http://arxiv.org/abs/0906.3768

Symantec (2009). *The downadup codex. A comprehensive guide to the threat's mechanics* (2.0 ed.). Retrieved from http://www.symantec.com/content/en/us/enterprise/media/security_response/whitepapers/the_downadup_codex_ed2.pdf

Tikk, E., Kaska, K., & Vihul, L. (2010). *International cyber incidents - Legal considerations*. Tallinn, Estonia: Cooperative Cyber Defence of Excellence (CCD COE).

van Eeten, M. & Bauer, J. (2008). *ITU study on the financial aspects of network security: Malware and spam*. ICT Applications and Cybersecurity Division, Policies and Strategies Department, ITU Telecommunication Development Sector. Retrieved from http://www.itu.int/ITU-D/cyb/cybersecurity/docs/itu-study-financial-aspects-of-malware-and-spam.pdf

van Eeten, M. J. V., & Bauer, J. M. (2008). *Economics of malware: Security decisions, incentives and externalities*. OECD publishing. OECD Science, Technology and Industry Working Papers 2008/1. doi:10.1787/241440230621

Weber, S. (2004). *The success of open source*. Cambridge, MA: Harvard University Press.

Whitman, M. E., & Mattord, H. J. (2007). Principles of incident response and disaster recovery. Boston, Mass.: Thomson Course Technology.