CHAPTER II

Essential Terms and Concepts

This chapter contains definitions and explanations of essential terms and concepts for those with a minimal knowledge of cyber security. As a wide readership is anticipated for the book, it is essential that terms and concepts are explained. Those with more experience in cyber security may want to either skim or go to the next relevant chapter, based on your interest. Essential terms are grouped and discussed by category:

1. Types of ethical hackers
2. Definitions and typology of ethical hacking
3. Conventional computer-security model
4. Common methods
5. Other relevant terms

These terms will be used throughout the book and explained further, where relevant to a specific context. Nonetheless, the reader is encouraged to engage with this chapter to ensure a fuller understanding of the ethical-hacking landscape.

2.1 Types of Ethical Hackers

The terminology around ethical hacking is confusing as terms mean different things according to their disciplines, and often

these terms are used interchangeably. For instance, the technical world distinguishes between a hacker and a cracker, whereas the mainstream media lump both terms under the umbrella of hacker. Expressed differently, the distinction is sometimes made by referencing "black-hat," "grey-hat," and "white-hat" hackers. For clarity, these terms are defined below:

Hacker: "A person who delights in having an intimate understanding of the internal workings of a system, computers and computer networks in particular. The term is often misused in a pejorative context, where 'cracker' would be the correct term."[1]

Cracker: "A cracker is an individual who attempts to access computer systems without authorization. These individuals are often malicious, as opposed to hackers, and have many means at their disposal for breaking into a system."[2]

Black-hat hacker: (also referred to as a cracker), is "someone who uses his computer knowledge in criminal activities in order to obtain personal benefits. A typical example is a person who exploits the weaknesses of the systems of a financial institution for making some money."[3]

White-hat hacker: "Although white hat hacking can be considered similar to a black hacker, there is an important difference. A white hacker does it with no criminal intention in mind. Companies around the world, who want to test their systems, contract white hackers."[4] They will test the security of a system, and are often hired to make recommendations to improve such systems.

Grey-hat hacker: "A grey hat hacker is someone who is in between these two concepts. He may use his skills for legal or illegal acts, but not for personal gains. Grey hackers use their skills in order to prove themselves that they can accomplish a determined feat, but never do it in order to make money out of it. The moment they cross that boundary, they become black hat hackers."[5]

People who participate in ethical hacking do not fit neatly into set categories. The differentiation, however, between hackers, crackers, and hat colours plays little importance when looking at these concepts from a legal perspective. Any form of unauthorized access, modification, or impairment of data, network, or computer is a crime. There are no exemptions in most jurisdictions; hackers and crackers alike rely on the discretion of law enforcement as to whether to prosecute or turn a blind eye. Another fallacy in classifying hackers is that an individual falls solely into one definition. Each attack must be individually

characterized, not the individual behind the attack. For example, you might have a hacker who predominantly breaks into systems to learn, sometimes she might even fix a security flaw in a system. The same hacker might also break into a system to collect data on individuals who are actively engaged in child pornography, and then make this data publicly available to law enforcement and the public. Yet this same hacker might also accept a fee to break into a corporation's (one they may view as unethical) database and steal a trade secret that is handed over to a competitor. Each of these examples involves unauthorized access. The difference with each attack goes to intent and motive, and involves the individual's subjective notion of what is ethical or moral. Ethical hacking, therefore, is difficult to define.

2.2 Definitions and Typology of Ethical Hacking

Ethical hacking is also a term that is used interchangeably with hacktivism in the media, but which has a distinct meaning in the computer-science discipline. For example, in the computer sciences "ethical hacking" is used to describe what is known as penetration or intrusion testing (white-hat hacking). Similarly, someone who merely participates in a denial-of-service attack for political reason would not be considered a hacker within the computer-sciences community. This type of action would be more akin to online civil disobedience.

For this book, "ethical hacking" will be used in its broadest sense to include the following activities:

Online civil disobedience: the use of any technology that connects to the Internet in pursuit of a cause or political end.

Hacktivism: the clever use of technology that involves unauthorized access to data or a computer system in pursuit of a cause or political end.[6]

Penetration/intrusion testing: is a type of information-systems security testing on behalf of the system's owners. This is known in the computer-security world as ethical hacking. There is some argument, however, as to whether penetration testing must be done with permission from a system's owners or whether a benevolent intention suffices in the absence of permission. Whether permission is obtained or not, however, does not change the common cause: improving security.

Vulnerability discovery: is the process of finding weaknesses and ways in a network, device, or within an organization that are

capable of being exploited by others (sometimes for nefarious reasons). Vulnerability discovery is often done with the authorization of the owner/operator of a network or device, but not always.

Counterattack: is also referred to as hackback or strikeback. Counterattack is when an individual or organization who is subject to an attack on their data, network, or computer takes similar measures to attack back at the hacker/cracker. For example, when an individual or organization is subject to a denial-of-service attack, that organization might initiate their own denial-of-service attack on the responsible party's website.

Security activism: is similar to penetration/intrusion testing in that the cause is to improve security. Security activism goes beyond mere testing of security, however, to gather intelligence on crackers, and to launch active attacks to disrupt criminal online enterprises. One example is the taking down of a botnet (see definition below).

My definition of ethical hacking potentially includes all the above, though ethical-hacking incidences are, like most things, contextual and fact-specific. I have chosen not to require that an act be "legal" as all the case studies discussed in this book are captured as illegal under hacking provisions that adopt a strict liability approach. For example, when access or use of data, network, or computer is unauthorized it is captured under criminal provisions. Some jurisdictions, as will be seen in the book, require intent to commit a criminal act, other jurisdictions have hacking provisions with no mention of intent or motive. The absence of intent in a criminal provision is known as a strict liability offence.

Ethical hacking, then, is the non-violent use of a technology in pursuit of a cause, political or otherwise, which is often legally and morally ambiguous.[7]

The use of a technology that resulted in acts of violence or physical harm would fall outside the scope of ethical hacking. Cyber jihadism, controversially, is included under this definition if the actions do not result in violence or physical harm, though arguably this is difficult to measure.

2.3 Conventional Computer-Security-Threat Model

The conventional computer-security model is adopted whereby threatening events impinge on vulnerabilities to cause harm. Safeguards are then used to prevent or ameliorate that harm. At

least that is the theory. Some ethical-hacking incidences do not cause harm, or at least not in the conventional way. Nonetheless, these incidences are often treated as falling within the standard computer-security-threat model. More fully:

Threat: A threat is a circumstance that could result in harm or damage and may be natural, accidental, or intentional. A party responsible for an intentional threat is referred to as an attacker.

Threatening event: A threatening event is an instance of a generic threat (such as malicious code) that may cause harm or damage.

Harm: Harm is anything that has deleterious consequences, which includes injury to persons, damage to property, financial loss, loss of value of an asset, and loss of reputation and confidence. Harm arises because a threatening event impinges on a vulnerability.

Vulnerability: A vulnerability is a feature or weakness that gives rise to susceptibility to a threat. Vulnerabilities exist in software and hardware.

Exploit: An exploit is the implementation, in software, of a vulnerability.

Safeguard: A safeguard is a measure intended to avoid or reduce vulnerabilities. Safeguards may or may not be effective and may be subject to countermeasures.

A functioning cyber-security ecosystem has an attack-safeguard-countermeasures cycle. Increasingly, as will be seen in the book, there is the need to identify and remedy threats and vulnerabilities before attacks may be instigated.

The question becomes to what extent does ethical hacking challenge the conventional computer-security-threat model? This question is explored in a variety of contexts, drawing on case studies throughout the book.

2.4 Common Methods Used in Ethical Hacking

The following section provides explanations of some of the most common methods used in ethical hacking.

SQL injection: Defacing a website involves the insertion of images or text into a website. This is often done via a SQL (structured query language) injection. A SQL injection is an attack in which computer code is inserted into strings that are later passed to a database.[8] A SQL injection can allow someone to target a database giving them

access to the website. This allows the person to deface the website with whatever images or text they wish.

DNS hijacking: DNS (domain name system) hijacking allows a person to redirect web traffic to a rogue domain name server.[9] The rogue server runs a substitute IP address to a legitimate domain name. For example, www.alanna.com's true IP address could be 197.653.3.1, but the user would be directed to 845.843.4.1 when they look for www.alanna.com. This is another way of redirecting traffic to a political message or image.

Adware: Adware refers to any software program in which advertising banners are displayed as a result of the software's operation. This may be in the form of a pop-up or as advertisements displayed on the side of a website, such as on Google or Facebook.

Phishing: Phishing refers to the dishonest attempt to obtain information through electronic means by appearing to be a trustworthy entity.

Ransomware: Ransomware is a type of malicious software that prevents the user from accessing or using their data (often through encrypting the data), whereby a fee must be paid or service performed before the user's data is decrypted.

Malware: A simplistic definition of malware is malicious software. Malware, for the purpose of this research, is defined as potentially harmful software or a component of software that has been installed without authorization to a third-party device.[10]

Virus: A virus is a "block of code that inserts copies of itself into other programs." Viruses generally require a positive act by the user to activate the virus. Such a positive act would include opening an email or attachment containing the virus. Viruses often delay or hinder the performance of functions on a computer, and may infect other software programs. They do not, however, propagate copies of themselves over networks. Again, a positive act is required for both infection and propagation.[11]

Worm: A worm is a program that propagates copies of itself over networks. It does not infect other programs, nor does it require a positive act by the user to activate the worm. It replicates by exploiting vulnerabilities.

Zero day: Zero day is an exploit or vulnerability that is exploited against a target on the day on which public awareness of the existence of the vulnerability occurs (i.e., zero days have elapsed between the awareness and the use). These vulnerabilities are

typically considered to be the most valuable as the utility and value of an exploit or vulnerability markedly decreases once it is known, as vendors produce patches or users reconfigure their systems to ameliorate the effect of the vulnerability.[12]

Back door: A back door is a method of accessing a computer program or network that circumvents security mechanisms. Sometimes a programmer will install a back door so that the programmer can accesses the program to perform security patches, troubleshoot, or monitor use. Attackers, however, can also use back doors that they discover (or install themselves) as part of an exploit.[13]

Distributed denial of service (DDoS): A DDoS attack is the most common form of online civil protest. A denial-of-service attack is distributed when multiple systems flood a channel's bandwidth and/or flood a host's capacity (e.g., overflowing the buffers).[14] This technique renders a website inaccessible.

DDoS attacks are performed with a botnet, with several of the compromised computers sending packets to the target computer simultaneously. A DDoS attack may also be distributed by use of peer-to-peer nodes.[15] The importance of botnets is explained below.

A botnet is comprised of core elements.[16] They are defined below for clarity and will be re-examined in more specific contexts in the analysis that follows this section.

Botnet: A botnet is a collection of compromised computers that are remotely controlled by a bot master.

Compromised computer: The term "compromised computer"[17] is commonly used interchangeably, and in some cases wrongly, in the literature with "zombie," "bot," and "bot client," which confuses hardware with software, creates inconsistency of usage, and may be confusing to users. Herein, a "compromised computer" is a computer that is connected to the Internet (an internet is any network of any size that uses the protocol TCP/IP, and the Internet is the largest such internet)[18] and on which a bot is installed.[19] The computer is thus said to be compromised.

Bot: A bot is software that is capable of being invoked from a remote location in order to provide the invoker with the capacity to cause the compromised computer to perform a function.[20] Botnets have a modular structure whereby modules (bots) may be added or taken away from each bot to add to it new exploits and capabilities. This ensures a botnet master's ability to rapidly respond to technical measures set up to infiltrate and take down the botnet.[21]

Bot server and command-and-control (C&C) source: C&C refers to the communications infrastructure of a botnet. A botnet master issues commands and exercises control over the performance of bots. Bots fetch data from a pre-programmed location, and interpret that data as triggers for action and instructions on what function to perform. The pre-programmed location is known as the bot server or C&C source. C&C is achieved by means of a bot server. The term "server" refers to any software that provide services on request by another piece of software, which is called a client. The bot requests and the server responds. Where the client is a bot, the server is reasonably enough called a bot server. Common bot servers are IRC servers, HTTP servers, the DNS (by means of TXT records), peer-to-peer nodes, cloud nodes, and increasingly devices otherwise known as the Internet of things (e.g., Xbox).

Traffic between the C&C source and its bots may be in clear or encrypted form. For example, IRC is an open-network protocol that can also be used with SSL (Secure Sockets Layer). SSL enables the establishment of an encrypted channel. Where the C&C of a botnet occurs in IRC alone, the information is openly available for viewing and tracking. When SSL is used in conjunction with IRC, the information is encrypted and is, therefore, not visible to anyone who lacks access to the relevant decryption key. For the purpose of clarity, there will be no further reference to the term "bot server" here unless in a quote. Rather, "C&C source" will be the term used throughout.

Multihoming: Involves the configuration of a domain to have several IP addresses. If any one IP address is blocked or ceases to be available, the others essentially back it up. Blocking or removing a single IP address, therefore, is not an effective solution to removing the content. The content merely rotates to another IP address.

Dynamic DNS: A service that enables the domain name entry for the relevant domain name to be updated very promptly, every time the IP address changes. A dynamic DNS provider enables a customer to either update the IP address via the provider's web page or using a tool that automatically detects the change in IP address and amends the DNS entry. To work effectively, the time-to-live value for the DNS entry must be set very short, to prevent cached entries scattered around the Internet serving up outdated IP addresses. Chapter 7 will explore DNS policy to prevent dynamic DNS being used by botnet masters.

Fast flux: A particular, dynamic DNS technique used by botnet masters whereby DNS records are frequently changed. This could be every five minutes.[22] Essentially, large volumes of IP addresses are rapidly rotated through the DNS records for a specific domain. This is similar to dynamic DNS tactics. The main difference between dynamic DNS and fast flux is the automation and rapidity of rotation with a fast-flux botnet.[23] Some fast-flux botnets rotate IP addresses every five minutes, and others every hour. Introducing a policy whereby IP addresses are not allowed to quickly rotate at the DNS level will be explored in chapter 7.[24]

Distributed Command and Control (or super botnets): A type of botnet that draws on a small botnet comprised of fifteen to twenty bots. The botnet herders may have anywhere from 10,000 to 250,000 bots at their disposal but use a select few for a particular purpose. The smaller botnet is then used to issue commands to larger botnets (hence the term "distributed command and control").[25]

Encryption: Encryption is the conversion of plain text into "ciphertext," encrypted information. Encryption acts to conceal or prevent the meaning of the data from being known by parties without decryption codes. Botnet instructions commonly use encryption. Encrypted instruction can then not be analyzed, making investigating, mitigation, and prevention much more difficult. Public-key cryptography is often used. In public-key cryptography, a twin pair of keys is created: one is private, the other public. Their fundamental property is that, although one key cannot be derived from the other, a message encrypted by one key can only be decrypted by the other key.

Proxy servers: Proxy servers refer to a service (a computer system or an application) that acts as an intermediary for requests from clients by forwarding requests to other servers. One use of proxy servers is to get around connection blocks such as authentication challenges and Internet filters. Another is to hide the origin of a connection. Proxy servers obfuscate a communication path such that user M connects to a website through proxy server B, which again connects through proxy server Z, whereby the packets appear to come from Z not M. Traceback to Z yields information of an additional hurdle, however, as packets also appear to come from B. Other proxy servers such as Tor are anonymous. Tor is also known as an onion router. Tor is described as follows:

> Tor protects you by bouncing your communications around a distributed network of relays run by volunteers all around the world: it prevents somebody from watching your Internet connection from learning what sites you visit, and it prevents the sites you visit from learning your physical location.[26]

Tor is described as onion routing due to the use of multiple layers of proxy servers, similar to the multiple layers of an onion. Tor is used by users in heavily Internet-censored countries, like China and Iran, to access blocked websites, as well as by some criminals to prevent law enforcement from traceback to the source.

Virtual private network (VPN) service: A VPN is a network that uses a public telecommunications infrastructure (usually the Internet) to connect remote sites or users together.[27] This connection allows secure access to an organization's network. Instead of a dedicated, real-world connection such as a leased line, a VPN uses virtual connections "routed through the Internet from an organization's private network to the remote site or employee."[28] VPN is made secure through cryptographic tunnelling protocols that provide confidentiality by blocking packet sniffing and interception software. VPN is used by many companies and government agencies, as well as by cybercriminal gangs such as will be seen in section 2.6 with the Mariposa botnet.

Rootkits: Rootkits are software or hardware devices designed to gain administrator-level control and sustain such control over a computer system without being detected.[29] A rootkit is used to obscure the operation of malware or a botnet from monitoring and investigation.

Peer-to-peer (P2P) communications: P2P "is any distributed network architecture composed of participants that make a portion of their resources (such as processing power, disk storage or network bandwidth) directly available to other network participants, without the need for central coordination instances."[30] Famous botnets such as Waledac, Torpig, and Mariposa use P2P protocol as their backup C&C. A P2P network relies on the capacity of multiple participants' computers, each of which has both client and server capabilities. This differs from conventional client-server architectures, where a relatively low number of servers provide the core function of a service or application.[31] Such networks are useful for many purposes, such as sharing of scientific information among researchers, file-sharing

of videos and music, and for telephone traffic. P2P operates on peer nodes.[32] P2P may be used to send content in clear or encrypted format. The ad hoc distribution of P2P makes it an ideal server location for C&C. The use of P2P channels allows an additional layer of rapid IP-address fluctuation. For this reason, botnets that use P2P channels are seen as offering the equivalent of "double fast-flux." The diagrams in figure 1 explain a botnet.

In step 1, the botnet herder needs to install bots on computers and thereby acquire compromised computers in order to build his/her botnet.

In step 2, the botnet master then makes content available to the bots, which causes them to perform actions. The botnet master may or may not be the botnet herder who builds the botnet. The botnet master could, for example, hire the use of the botnet.

There are three ways of using a botnet to perform a denial-of-service attack:

Make the botnet. In the first, a person would have to physically make a botnet through painstaking hours of labour, as it would involve compromising several hundred, if not thousands, of computers. This type of botnet would require the botnet master to have a high level of computer skills.

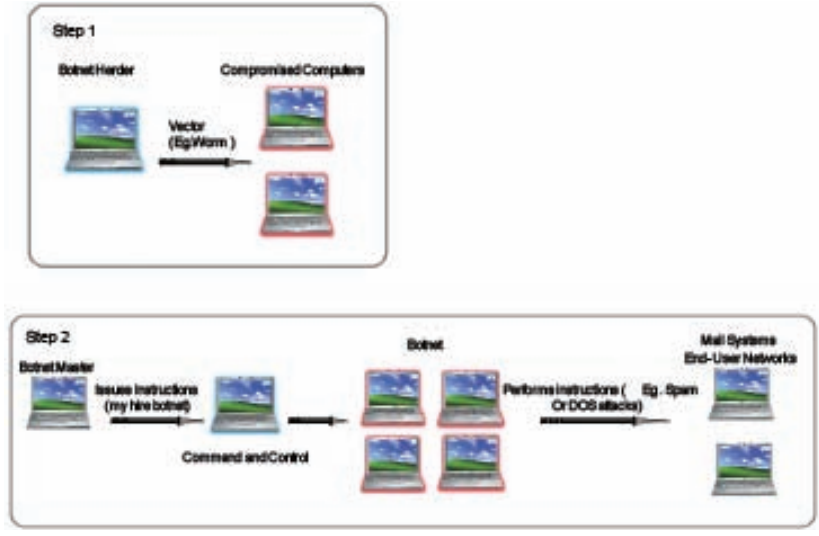

Figure 1. Steps in Procuring and Using a Botnet.

Hire/rent a botnet. The second type involves simply hiring someone to execute a denial-of-service attack. This requires no computer skills, but for the ability to use a search engine (such as Google). Bot-agent design and bot delivery have become a commoditized service industry.[33] A small botnet is sufficient to launch an effective denial-of-service attack causing much damage, and costs as little as US$200 for a twenty-four-hour attack.[34] A person does not require any special computer skills to use a botnet to commit a crime. Figure 10 later in the book sketches the commercialization of denial-of-service attacks with a botnet. The customer would merely specify the targeted website to attack, pay a nominal fee of US$200, and a denial-of-service attack would be launched for twenty-four hours against the website.

LOIC. The third type is where the user allows their device to become part of a botnet for the purpose of participating in a DDoS protest with LOIC (as expressed above, Low Orbit Ion Canon) or similar software. LOIC is a free software program and is used for most of the denial-of-service attacks performed by members of Anonymous, for example. Use of LOIC requires minimal computer skills. One googles LOIC, downloads the software with a click, types in the URL (e.g., www.paypal.com), and presses start. The denial-of-service attack then commences and people from all over the world may join in using LOIC.

2.5 Other Relevant Terms

Cloud: The cloud is a term for web-based applications and data-storage solutions. Companies such as Google, Microsoft, Yahoo, and Amazon are among the many companies that offer cloud computing services for individuals, corporations, and governments to store and access their data online, on the cloud.[35]

Internet of things (IoT): The IoT refers to "the network of physical devices, vehicles, home appliances, and other items embedded with electronics, software, sensors, actuators, and connectivity which enables these things to connect, collect and exchange data."[36] IoT sees traditionally non-Internet-connected devices or objects becoming connected to Internet-connected devices in a network, thereby rendering such devices or objects to be monitored and controlled.

API keys: An application programming interface (API) key "is a code passed in by computer programs calling an application

programming interface...to identify the calling program, its developer, or its user to the website. API keys are used to track and control how the API is being used, for example to prevent malicious use or abuse of the API (as defined perhaps by terms of service)."[37]

Surface Web: "The Surface Web is the portion of the Web that has been crawled and indexed (and thus searchable) by standard search engines such as Google or Bing via a regular web browser."[38]

Deep Web: The Deep Web refers to non-indexed websites (websites which do not appear in your web page browser's search engine results). This can include publicly accessible online databases, pay-to-access databases, subscription-based services, and webpages located behind password-protected web pages.[39]

Dark Web: Deeper than the Deep Web, the dark Web is accessible via software such as Tor, which enables users to anonymously connect to web pages. The dark Web is a space for political dissidents, whistle-blowers, and journalists to communicate with others, but it is also a space for cybercriminals to operate in due to its somewhat anonymizing features and degrees of anonymity.[40]

Internet protocol (IP) address: "An Internet Protocol address (IP address) is a numerical label assigned to each device connected to a computer network that uses the Internet Protocol for communication. An IP address serves two principal functions: host or network interface identification and location addressing."[41]

URL: "A Uniform Resource Locator (URL), colloquially termed a web address, is a reference to a web resource that specifies its location on a computer network and a mechanism for retrieving it. A URL is a specific type of Uniform Resource Identifier (URI), although many people use the two terms interchangeably. URLs occur most commonly to reference web pages (http), but are also used for file transfer (ftp), email (mailto), database access (JDBC), and many other applications."[42]

Notes

1. RFC 1392 Internet Users Glossary.
2. RFC 1392 Internet Users Glossary.
3. Hacking Alert, "White Hat and Grey Hat Hacking."
4. Hacking Alert, "White Hat and Grey Hat Hacking."
5. Hacking Alert, "White Hat and Grey Hat Hacking."
6. Samuel 2004.

7. Samuel 2004.
8. Security Spotlight 2010.
9. Security Spotlight 2010.
10. Clarke 2009.
11. Pfleeger 2006.
12. Oremus 2013.
13. Rouse 2007.
14. For more, see Wikipedia, "Denial of Service Attack (distributed)."
15. Athanasopoulos, Anagnostakis, and Markatos 2006.
16. Solomon and Evron 2008.
17. The term "compromised computer" has been selected over the term "compromised device." A computer may be as little as a processor (a personal computer will often contain multiple processors) or may be the world's largest computer. The term "computer" is used here to refer to any computing device, even if it is commonly called by some other name, and includes current and future devices with computing capabilities which may be connected to the Internet, including mobile phones, tablets, surveillance cameras, controllers for ADCs (analogue-digital converters) such as monitoring water levels, etc. For this reason, Clarke, for example, prefers "device." I have chosen "compromised computer," however, because it reflects the terminology used in computer science and information studies on botnets.
18. TCP/IP is often used as a single acronym when in fact it references two key protocols. TCP refers to transmission control protocol. TCP is a connection-oriented protocol that establishes a communication channel, known as a data stream, between two network hosts. IP refers to Internet protocol and is an identification and addressing scheme that, in the case of the latter, links distinct numerical labels as IP addresses. See Pfleeger and Pfleeger 2006.
19. A computer may still be compromised in the absence of a botnet master. Where a controller is gone but where a botnet continues to infect computers, it is referred to as an "orphan botnet." See Gutman, "The Commercial Malware Industry."
20. Modified definition of Clarke's (2009), where he defines bots as "(Generally, a program that operates as an agent for a user or another program. More specifically:) software that is capable of being invoked remotely in order to perform a particular function."
21. Dunham and Melnick 2009.
22. See "How Fast-Flux Service Networks Work" at http://www.honeynet.org/node/132.
23. Dunham and Melnick 2009.

24. L. Gaaster, *GNSO Council Issues Report on Fast Flux Hosting*, March 31, 2008, available at https://gnso.icann.org/sites/default/files/filefield_5868/gnso-issues-report-fast-flux-25mar08.pdf.
25. Barakat and Khattab, "A Comparative Study of Traditional Botnets Versus Super-Botnet," in INFOSEC 2010.
26. Tor Project, "Anonymity Online." Tor is available at https://www.torproject.org. There are many other types of anonymizing proxy servers and similar technologies, such as Phantom Access Agent.
27. Wikipedia, "Virtual Private Network."
28. Tyson 2010.
29. Pfleeger and Pfleeger 2006.
30. The author looked at many different definitions of P2P and found the Wikipedia definition had the best description; see Wikipedia "Peer-to-peer."
31. Clarke 2004.
32. Oram 2001.
33. Ollmann, cited in Greenberg 2010.
34. Ollmann, cited in Greenberg 2010.
35. Soghoian 2009.
36. Wikipedia, "Internet of Things."
37. Wikipedia, "Application Programming Interface Key."
38. Rudesill, Caverlee, and Sui 2015.
39. Rudesill, Caverlee, and Sui 2015, 8.
40. Rudesill, Caverlee, and Sui 2015, 8.
41. Wikipedia, "IP Address."
42. Wikipedia, "URL."

Page left blank intentionally