Chicago Journal of International Law

Volume 20 | Number 2

Article 18

1-1-2020

When is Cyber Defense a Crime? Evaluating ActiveCyber Defense Measures Under theBudapest Convention

Alexandra Van Dine

When is Cyber Defense a Crime? Evaluating Active Cyber Defense Measures Under the Budapest Convention

Alexandra Van Dine *

Abstract

As cyberattacks increase in frequency and intensity around the globe, private actors have turned to more innovative cyber defense strategies. For many, this involves considering the use of cutting-edge active cyber defense measures—that is, tactics beyond merely erecting firewalls and installing antivirus software that permit cyber defenders to detect and respond to threats in real time. The legality of such measures under international law is a subject of intense debate because of definitional uncertainty surrounding what qualifies as an "active" cyber defense measure. This Comment argues that active defense measures that do not rise to the level of a cybercrime are permissible under international law. Accordingly, it analyzes the Budapest Convention, the only binding international instrument related to cybercrime, and uses its definition of illegal conduct under international law to construct a "stoplight framework" to guide cyber defenders in their actions. Ultimately, this Comment concludes that cyber defenders have a "green light" to use purely passive measures, such as monitoring one's own network traffic, because these measures are highly unlikely to involve conduct the Budapest Convention criminalizes. Active-passive measures, such as attaching code to intruders that tracks them back to their home base, can in some cases be justified under exceptions to the Convention; accordingly, cyber defenders should proceed with caution. Finally, outright active defense measures nearly always rise to the level of offense conduct under the Budapest Convention, and should not be used. This analysis provides needed clarity as to the legality of conduct in cyberspace, and provides cyber defenders with the guideposts they need to confidently innovate in today's complex cyber landscape.

* The author would like to thank everyone involved in crafting and shaping this Comment, especially her faculty advisor, Professor Mary Ellen O'Connell, her editors (Osama Alkhawaja, Whittney Barth, Michael Christ, Mark Cronin, Nyle Hussain, Casey Jedele, Justyna Jozwik, and Brian Pollock), and the entire CJIL Board. Special thanks to Carol, Mark, Jake, and Rob for their love and support. Finally, this Comment would not exist without Michael J. Assante, in whose memory this Comment is offered, and Page O. Stoutland, who first opened the author's eyes to the idea of active cyber defense.

Table of Contents

I. INTRODUCTION

Imagine that you are the systems administrator at a major, multinational power company. Recognizing the vital role your networks play in safely delivering energy to consumers around the world, you are motivated to implement the most state-of-the-art security measures that you can afford.

You then decide to set up a "honeypot"—a part of your system designed to be attractive to attackers and that no one has any legitimate motive to access. Soon, traffic begins to flow, and your dedicated team of cyber defenders monitors it. As time passes, they analyze the traffic to figure out who is intruding, carefully tracing it back to its source when possible. Some of the intruders have masked their locations by routing their activities through multiple IP addresses, and it is impossible to determine their identities. Those intruders are expelled from the system and the firewalls are updated to keep them out.

When intruders can be identified, your defenders have followed them back to their own networks and have investigated those networks in order to learn more about who is accessing the honeypot. After gleaning as much information as possible, a defender shuts off the traffic flowing from that entity in order to stop the attack.

Many of the tactics used in the above scenario are considered to be "active cyber defense" measures.[1] Active cyber defense generally involves cyber defense and security strategies that go beyond simply erecting a firewall or installing antivirus software and allow cyber defenders to detect and respond to threats in real time.[2] These tactics exist on a spectrum that spans everything from active network monitoring to setting cyber traps to retaliatory hacking.[3] Because using these measures may require crossing literal territorial boundaries in cyberspace without a right to be there, whether and how active defense measures can be used at all under international law is a critical question.

International criminal law, as it relates to cyberspace, provides a guidepost as to which actions are and are not permissible—even if taken in self-defense. Although a substantial portion of scholarship examining international law in cyberspace focuses on applying the laws of armed conflict, those analogs are not

[1] *See, for example*, Wyatt Hoffman & Ariel (Eli) Levite, *Private Sector Cyber Defense: Can Active Measures Help Stabilize Cyberspace?*, CARNEGIE ENDOWMENT FOR INT'L PEACE (June 14, 2017), http://perma.cc/CKL9-HE5M.

[2] *See, for example*, CENTER FOR CYBER & HOMELAND SECURITY, GEO. WASH. U., Into the Gray Zone: The Private Sector and Active Defense Against Cyber Threats, 7 (2016) http://perma.cc/SAX8-4LW3.

[3] *Id.* at 9.

useful when addressing intrusions that do not rise to the level of a "use of force."[4] Most active cyber defense tactics do not rise to that level.[5] Moreover, application of the state responsibility doctrine—a central component of the laws of armed conflict—can get very complicated, very quickly in this area. There are no soldiers bearing the flag of the attacking nation, only actions perpetrated by someone sitting behind a computer somewhere in the world, with plenty of tools at his or her disposal to mask his or her location and identity.[6]

This Comment seeks to fill that gap by analyzing and applying international law related to cybercrime, as set forth in the Council of Europe's Convention on Cybercrime (hereinafter "the Budapest Convention" or "the Convention"). The Convention is the only legally binding international instrument delineating when an action in cyberspace becomes a crime. By filtering the active cyber defense discussion through the prism of what constitutes a cybercrime under international law, this Comment articulates a new boundary as to which defensive actions are permissible in cyberspace.

Developing a method to analyze and categorize defensive approaches in this fashion is critical, as the current approach to cybersecurity requires innovation. The frequency of successful cyberattacks—from the WannaCry ransomware attack that struck hospitals in the United Kingdom,[7] to the massive data breach at the U.S. Office of Personnel Management,[8] to the cyberattack against the Wolf Creek Nuclear Operating Corporation[9]—suggests that cyber defenders need to devise more clever defenses. It is generally recognized that attackers have the edge when it comes to agility and innovation,[10] and defenders have long been playing catch-up.

This quest to match the creativity and agility of cyberattackers has involved the use of active defense measures. These strategies permit defenders to detect and expel intruders from networks faster and might deter illegitimate access more effectively. This outcome is preferable for large, for-profit corporations because

[4] *See* Alexandra Perloff-Giles, *Transnational Cyber Offenses: Overcoming Jurisdictional Challenges*, 43 YALE J. INT'L L. 191, 202–03 (2018).

[5] *See id.* at 204 ("For most transnational cyber offenses…the offense does not constitute an Article 51 'armed attack' or a 'resort to armed force'…").

[6] *See id.* at 203.

[7] *See* Lily Hay Newman, *The Ransomware Meltdown Experts Warned Us About is Here*, WIRED (May 12, 2017, 2:03 PM), http://perma.cc/A3J5-6WKK.

[8] *See* Brendan I. Koerner, *Inside the Cyberattack That Shocked the US Government*, WIRED (Oct. 23, 2016, 5:00 PM), http://perma.cc/Y7FY-DE5F.

[9] *See* Nicole Perlroth, Hackers are Targeting Nuclear Facilities, Homeland Security Dept. and F.B.I. Say, N.Y. TIMES, July 6, 2017, at B5.

[10] *See* Alyza Sebenius, *Writing the Rules of Cyberwar*, THE ATLANTIC (June 28, 2017), http://perma.cc/ZR8J-QR9G.

relying on the processes of international or domestic law to cure violations after the fact can be unsatisfying. The financial and reputational impacts of these attacks are difficult to fully remedy. Once the personal data of millions of people is leaked, or the power grid has been shut off, it is difficult to recover the full cost of the cyber incident. Empowering system administrators and operators to identify and address intrusions in real-time would be more effective at stopping an attack before these consequences occur.[11]

As professionals increasingly explore active cyber defense as a solution to these problems, an analysis of how to do so in a way that comports with international law is extremely important. As it stands, "[e]ven though counterstrikes are currently of questionable legality, counterstrikes have already been occurring on the internet over the last decade, initiated by both government and private actors."[12] Providing guidance to those private actors is of particular importance, as "[t]he development of the [i]nternet is essentially market-led and driven by private and government initiatives" and "the private sector continues to play a very important role in the expansion and development of the [i]nternet."[13] Guidance in the private sector is sorely needed, and this Comment contributes to that conversation.

Section II of this Comment discusses the relevant international law related to cybercrime as set forth in the Budapest Convention. This Section analyzes activities the Convention requires signatories to criminalize that are relevant to the types of actions taken as part of an active defense strategy. In order to develop as accurate an understanding as possible, this Section draws upon guidance documents produced by the Council of Europe to aid in interpreting the Convention.

In Section III, this Comment defines the term "active cyber defense" and proposes a spectrum of cyber defenses. This Comment, based on a survey of the active defense literature, divides this spectrum into three categories of defenses. First, passive measures are those that, despite their inclusion under the active defense umbrella, do not involve taking any external action. They are deployed internally on an entity's own network. Second, there are "active" passive measures—defenses that may be set up and operated on an entity's own network, with occasional external consequences. Finally, there are active defense measures. These are purely external to the network and are targeted and deployed specifically to end an attack or an intrusion. International law has different implications for each of these categories.

[11] *See* Jay P. Kesan & Carol M. Hayes, *Mitigative Counterstriking: Self-Defense and Deterrence in Cyberspace*, 25 HARV. J. L. & TECH. 429, 474 (2012).

[12] *Id.* at 475.

[13] *International Telecommunications Union Res. 102*, THE PLENIPOTENTIARY CONFERENCE OF THE INTERNATIONAL TELECOMMUNICATIONS UNION (2014), http://perma.cc/TGD5-ZGE8.

Finally, in Section IV, this Comment will apply those laws to the proposed cyber defense spectrum and distinguish between lawful defenses and unlawful cybercrimes. This application suggests that cybersecurity professionals are almost always justified in employing passive defense measures. Indeed, these are rarely even implicated by the Convention, as they operate entirely internally to an entity's network. As for active-passive measures, their permissibility depends upon whether they qualify as one of three potential defenses suggested as justified by the Convention. Finally, purely active defense measures are almost never permissible under the Convention, and therefore are generally unlawful under international law.

Categorizing measures in this way should help to clarify the boundaries within which cyber defenders must work when it comes to innovating and advancing cyber defense.

II. THE BUDAPEST CONVENTION

The Budapest Convention,[14] which entered into force in 2004, is the only binding international instrument related to cybercrime.[15] It was created to articulate a "common criminal policy aimed at the protection of society against cybercrime," and specifically intends "to deter action directed against the confidentiality, integrity, and availability of computer systems, networks and computer data as well as the misuse of such systems, networks, and data by providing for the criminalization of such conduct."[16] The Convention sets forth the powers and procedures that states have in investigating, prosecuting, and punishing these crimes.[17] Sixty-one states have ratified it, including Australia, Canada, Israel, Japan, the U.S., and most countries in the European Union.[18]

The Convention approaches these goals from three different angles. First, it standardizes the domestic criminal law related to cybercrime in states that are party to (and therefore bound by) the Budapest Convention (hereinafter "States Party"). Second, it motivates the creation of the necessary criminal procedural laws to investigate and prosecute cybercrime within States Party. Finally, it establishes an agile international cooperation regime.[19] It defines nine discrete offenses: illegal access, illegal interception, data interference, system interference, misuse of

[14] Convention on Cybercrime, Nov. 23, 2001, T.I.A.S. 13174, E.T.S. No. 185, http://perma.cc/4KKP-2YM7 [hereinafter Budapest Convention].

[15] *Budapest Convention and Related Standards*, COUNCIL OF EUR., http://perma.cc/C34X-EUJF.

[16] Budapest Convention, *supra* note 14, at 2.

[17] *See id.*

[18] *Chart of Signatures and Ratifications of Treaty 185*, COUNCIL OF EUR., http://perma.cc/57D7-XPBF.

[19] *See generally*, Budapest Convention, *supra* note 14.

devices, computer-related forgery, computer-related fraud, child pornography-related offenses, and offenses related to copyright.[20]

The Convention further delineates several procedural law issues, including expedited preservation of stored data, expedited preservation and partial disclosure of traffic data, production order, search and seizure of computer data, real-time collection of traffic data, and interception of content data.[21] It also calls for constructing a network that operates twenty-four hours a day, seven days a week to facilitate rapid assistance among signatories "for the purpose of investigations or proceedings concerning criminal offences related to computer systems and data, or for the collection of evidence in electronic form of a criminal offence."[22]

Additionally, the Council of Europe published several guidance documents to aid in the interpretation of the Convention. Although these documents "[do] not constitute [instruments] providing an authoritative interpretation of the Convention," they "might be of such a nature as to facilitate the application of the provisions contained therein."[23] The Council also explains that "Guidance Notes represent the common understanding of the parties to this treaty regarding the use of the Convention."[24] Accordingly, they are relevant to the present analysis, and even set forth several key definitions.[25]

[20] COUNCIL OF EUR., *Explanatory Report to the Convention on Cybercrime* (Nov. 23, 2001), ¶ 18, http://perma.cc/A6XF-647V [hereinafter Explanatory Report].

[21] *Id.* at ¶ 19.

[22] Budapest Convention, *supra* note 14, at art. 35.

[23] Explanatory Report, *supra* note 20, at 1.

[24] Cybercrime Convention Committee, *T-CY Guidance Note #3: Transborder Access to Data (Article 32)*, at 3 (Dec. 2–3, 2014), http://perma.cc/494T-7EHG.

[25] The Explanatory Report, for example, defines "computer system" as

> a device consisting of hardware and software developed for automatic processing of digital data. It may include input, output, and storage facilities. It may stand alone or be connected in a network with other similar devices [sic] "Automatic" means without direct human intervention, "processing of data" means that data in the computer system is operated by executing a computer program . . . A computer system usually consists of different devices, to be distinguished as the processor or central processing unit, and peripherals. A "peripheral" is a device that performs certain specific functions in interaction with the processing unit, such as a printer, video screen, CD reader/writer or other storage device.

Explanatory Report, supra note 20, at ¶ 23.

"[C]omputer program" is defined as "a set of instructions that can be executed by the computer to achieve the intended result." *Id.*

"[N]etwork" is defined as

All of the procedures established by the Convention are limited by a concern for preserving human rights, including those enshrined in the European Convention for the Protection of Human Rights and Fundamental Freedoms, the U.N. International Covenant on Civil and Political Rights, and other similar instruments.[26] All measures taken pursuant to the Convention must adhere to the international legal principle of proportionality.[27]

Notably, however, countries such as Russia, China, and India, among others, have not ratified the Convention.[28] These countries are large, geopolitically powerful, and active in cyberspace, so their unwillingness to ratify the Convention might be perceived as weakening the Convention's impact. Their rationales for refusing to ratify tend to fall into one of two categories. First, they object to ratifying something when they have not participated in its drafting process.[29] Second, they consider the treaty to be an infringement on sovereignty.[30] In the specific cases of Russia and China, in addition to objecting on both of these bases, these states have long been reticent to participate in cooperation or other information or intelligence sharing when it comes to cyberspace.[31]

A. Overview of Offenses and Remedies under the Budapest Convention

The Budapest Convention requires States Party to adopt legislation or other measures that criminalize intentional commission of certain offenses. These include, as relevant to the topic of active cyber defense: illegal access to computer systems, illegal interception of data, data interference, system interference, misuse of devices, computer-related forgery, and computer-related fraud. This Comment

> an interconnection between two or more computer systems. The connections may be earthbound (e.g., wire or cable), wireless (e.g., radio, infrared, or satellite), or both. A network may be geographically limited to a small area (local area networks) or may span a large area (wide area networks), and such networks may themselves be interconnected . . . What is essential is that data is exchanged over the network.

Id. at ¶ 24.

[26] *See* Budapest Convention, *supra* note 14, at art. 15.

[27] *Id.* Proportionality encompasses the idea "that a State's acts must be a rational and reasonable exercise of means towards achieving a permissible goal, without unduly encroaching on protected rights of either the individual or another State." Emily Crawford, *Proportionality*, *in* MAX PLANCK ENCYCLOPEDIA OF PUBLIC INTERNATIONAL LAW ¶ 1 (2011), http://perma.cc/YJ8E-VB5C.

[28] Joyce Hakmeh, *Building a Stronger International Legal Framework on Cybercrime*, CHATHAM HOUSE (June 6, 2017), http://perma.cc/TJT5-MMQB.

[29] *Id.*

[30] *Id.*

[31] *Id.*

only discusses these offenses as they are most relevant to the types of actions undertaken as part of mounting an active cyber defense.

These offenses are further punished in their inchoate form via Article 11, which requires States Party to criminalize both "aiding and abetting" and "attempting" the delineated offenses.[32] Article 12 provides for corporate liability for these crimes.[33] Although individuals who commit these crimes may be subject to deprivation of liberty, that punitive option seems unavailable under the Convention for pursuing corporate liability.[34]

The Convention also provides guidance on jurisdiction, noting that States Party have jurisdiction over offenses committed within their respective territories, on a ship flying the state's flag, on an aircraft registered under the laws of the State Party, or by a national of the State Party if the offense is criminalized in the State where the crime is committed or if it "is committed outside the territorial jurisdiction of any State."[35] States Party may exercise jurisdiction in accordance with their own domestic law, and the Convention provides a procedure by which jurisdictional conflicts might be resolved.[36]

In exchange for the promise to criminalize these offenses, the Convention provides extensive processes and procedures for mutual assistance and information sharing.[37] Not only must States Party implement domestic legislation criminalizing the enumerated offenses, they must do the same with regard to ensuring the ability to meet mutual assistance obligations under the Convention.[38] Moreover, States Party may forward information obtained in the course of their investigations to other states in order to assist them in carrying out the Convention.[39] However, States Party may insert provisions into their implementing legislation delineating when they will refuse cooperation, along with other conditions. [40]

States Party may only undertake two specific actions without authorization from another Party. First, States Party may "access publicly available (open source) stored computer data, regardless of where the data is located geographically."[41]

[32] Budapest Convention, *supra* note 14, at art. 11.

[33] *Id.* at art. 12.

[34] *Id.* at art. 13.

[35] *Id.* at art. 22.

[36] Budapest Convention, *supra* note 14, at art. 22.

[37] *See id.* at arts. 25–26.

[38] *Id.* at art. 25.

[39] *Id.* at art. 26.

[40] Budapest Convention, *supra* note 14, at art. 25.

[41] *Id.* at art. 32.

Second, States Party may "access or receive, through a computer system in [their territories], stored computer data located in another Party" as long as the Party obtains the consent of the person with the lawful authority to disclose that information.[42]

In summary, the Convention requires States Party to criminalize certain delineated offenses in exchange for assurances of help in bringing those who commit those offenses to justice. It attempts to construct an investigatory framework that respects national sovereignty while still incentivizing cooperation over self-help.

B. Offenses

Depending on how they are developed and executed, many potential components of an active cyber defense strategy could rise to the level of offenses prohibited by the Budapest Convention. In order to understand where the Convention draws this line, this Section further details the relevant offenses and the behavior they target. Specifically, those offenses are illegal access to computer systems, illegal interception of data, data interference, system interference, misuses of devices, computer-related forgery, and computer-related fraud. This Section will also discuss the inchoate form of these offenses, as well as potential corporate liability.

1. Illegal Access to Computer Systems

The Budapest Convention criminalizes illegal access to computer systems,[43] including "mobile phones or 'smart' phones, PDAs, tablets, and other [systems] that produce, process, or transmit data."[44] A computer system is defined in the Convention as "any device or group of interconnected or related devices, one or more of which, pursuant to a program, performs automatic processing of data."[45]

According to the Explanatory Report, "illegal access" encapsulates "dangerous threats to and attacks against the security … of computer systems and data."[46] It further explains that "security" encompasses the systems' and data's confidentiality, integrity, and availability[47] and clarifies that "mere unauthorized intrusion[s]" like hacking should be illegal as well, as those intrusions can

[42] *Id.*

[43] Budapest Convention, *supra* note 14, at art. 2.

[44] Cybercrime Convention Committee, *T-CY Guidance Note #1: On the Notion of "Computer System"* at 3 (Dec. 2012), http://perma.cc/S78P-VYHC.

[45] Budapest Convention, *supra* note 14, at art. 1.

[46] Explanatory Report, *supra* note 20, at ¶ 44; *see* Budapest Convention, *supra* note 14, at art. 2.

[47] Explanatory Report, *supra* note 20, at ¶ 44.

compromise data confidentiality or even embolden hackers to commit more serious offenses in the future.[48] The Report further defines "access" as "the entering of the whole or any part of a computer system" and clarifies that "it does not include the mere sending of an e-mail message or file to that system."[49] According to the Report, the method of entry is irrelevant, as long as the system is entered via some connection point.[50]

When it comes to actually criminalizing conduct, the Explanatory Report provides that States Party are welcome to take a broad approach and criminalize hacking in general.[51] Alternatively, they may narrow the definition of criminal behavior using the qualifications noted in the second sentence of Article 2[52]—namely, "infringing security measures, [with] the *intent* of obtaining computer data or other dishonest intent, or in relation to a computer system that is connected to another computer system."[53]

2. Illegal Interception of Data

The illegal interception of non-public transmissions of computer data to, from, or within a computer system using technical means is treated as a crime under the Budapest Convention.[54] This "[i]llegal interception" provision aims to protect data privacy and is intended to mimic the violation of privacy that occurs via wiretaps and recordings of telephone conversations in the physical world.[55]

The Explanatory Report clarifies that all forms of electronic transfer can give rise to an Article 3 offense. According to the Report, interception by "technical means" involves "listening to, monitoring or surveillance of the content of communications, [] the procuring of the content of data either directly, through access and use of the computer system, or indirectly, through the use of electronic eavesdropping or tapping devices."[56] Interception can also involve recording.[57]

"Technical means," according to the Report, include "technical devices fixed to transmission lines as well as devices to collect and record wireless communications" and "may include the use of software, passwords, and codes."[58]

[48] *Id.*

[49] *Id.* at ¶ 46.

[50] *Id.*

[51] Explanatory Report, *supra* note 20, at ¶ 50.

[52] *Id.*

[53] Budapest Convention, *supra* note 14, at art. 2 (emphasis added).

[54] *Id.* at art. 3.

[55] Explanatory Report, *supra* note 20, at ¶ 51.

[56] *Id.* at ¶ 53.

[57] *Id.*

[58] *Id.*

Apparently, the "technical means" qualification was intended to avoid over-criminalization.[59]

Article 3 offenses apply to "non-public" *transmissions* of computer data—that is, the transmission, and not the data, is what is non-public. Indeed, the data may well be public information that parties wish to communicate confidentially, or even data "kept secret for commercial purposes."[60] Employee communications also fall under an umbrella of "non-public" transmissions,[61] but domestic law can provide some legitimate cover for intercepting these communications.[62] In such a case, interception would take place "with right."[63] As to the transmission itself, States Party have the option of requiring that the communication take place between remote computer systems (as opposed to within a single computer system or between two systems belonging to the same person).[64]

Finally, the Convention requires that an interception be committed "intentionally" and "without right" for criminal liability to attach.[65] The intercepting person is justified in his or her action, for example, if acting with the authorization of the transmission's participants or if "surveillance is lawfully authorized in the interests of national security or the detection of offences by investigating authorities."[66] Furthermore, common commercial practices like using "cookies" are not intended to be criminalized, as these interceptions do not occur "without right."[67]

3. Data Interference

"The damaging, deletion, deterioration, alteration, or suppression of computer data" is considered a criminal offense under the Budapest Convention.[68] Data interference as an offense is meant to protect computer data and programs from intentional infliction of damage in a way similar to the protections enjoyed by physical objects.[69] The legal interest at stake "is the integrity and the proper functioning or use of stored computer data or computer programs."[70] The Report

[59] Explanatory Report, *supra* note 20, at ¶ 53.

[60] *Id.* at ¶ 54.

[61] *Id.*

[62] *Id.* at ¶ 58.

[63] Explanatory Report, *supra* note 20, at ¶ 58.

[64] *Id.* at ¶ 55.

[65] *Id.* at ¶ 58.

[66] *Id.*

[67] Explanatory Report, *supra* note 20, at ¶ 58.

[68] Budapest Convention, *supra* note 14, at art. 4.

[69] Explanatory Report, *supra* note 20, at ¶ 60.

[70] *Id.*

clarifies that "damaging" and "deteriorating" as used in Article 4 specifically refer "to a negative alteration of the integrity or of information content of data and programmes."[71] According to the Report, "deletion" of data is meant to have an equivalent meaning to the destruction of a physical object.[72] "Suppress[ion]" means "any action that prevents or terminates the availability of the data to the person who has access to the computer or the data carrier on which it was stored," and "alteration" involves "the modification of existing data."[73]

The Report clarifies that this offense covers "[t]he input of malicious codes, such as viruses and Trojan horses," [74] as well as any resulting changes in data;[75] however, activities considered common or inherent in network design or commercial operating practices are not criminalized, as such acts are done "with right."[76] These activities could include "the testing or protection of the security of a computer system authorised by the owner or operator, or the reconfiguration of a computer's operating system that takes place when the operator of a system acquires new software."[77] Moreover, modifying traffic data in order to facilitate anonymous communications or to ensure secure communications (as with encryption) is considered "a legitimate protection of privacy" and is therefore undertaken "with right."[78] Parties are permitted, however, to "criminalise certain *abuses* related to anonymous communications," for example, when they are used to facilitate the commission of crimes.[79]

4. System Interference

The Budapest Convention criminalizes "the serious hindering . . . of the functioning of a computer system by inputting, transmitting, damaging, deleting, deteriorating, altering, or suppressing computer data."[80] According to the Explanatory Report, the legal interest at stake is ensuring the proper functioning of computer and telecommunication systems.[81] The Report defines "hindering" as "actions that interfere with the proper functioning of the computer system,"

[71] Explanatory Report, *supra* note 20, at ¶ 61.

[72] *Id.*

[73] *Id.*

[74] A "Trojan horse" is defined as "a type of malware that is often disguised as legitimate software." *What is a Trojan Virus? - Definition*, KASPERSKY, http://perma.cc/N5JE-9CVZ.

[75] Explanatory Report, *supra* note 20, at ¶ 61.

[76] Explanatory Report, *supra* note 20, at ¶ 62.

[77] *Id.*

[78] *Id.*

[79] *Id.* (emphasis added).

[80] Budapest Convention, *supra* note 14, at art. 5.

[81] Explanatory Report, *supra* note 20, at ¶ 65.

and clarifies that such hindrance must be sufficiently serious "to give rise to criminal sanction."[82]

States Party are permitted to "require a minimum amount of damage to be caused in order for the hindering to be considered serious."[83] The drafters themselves considered it "serious" when sending data to a system in a way that had "a significant detrimental effect on the ability of the owner or operator to use the system, or to communicate with other systems."[84] They further noted that "spamming," or sending messages "in large quantities or with a high frequency" should only rise to a level meriting criminal sanction when sent intentionally and in a way that seriously hinders communication.[85] Parties are left to determine on their own how seriously a system must be hindered for the act to be punishable by criminal law.[86]

5. Misuse of Devices

Under the Budapest Convention, "the production, sale, procurement for use, import, distribution or otherwise making available of" devices and computer programs "designed or adapted primarily for the purpose of committing" illegal access, illegal interception, data interference, or system interference is a crime.[87] These are devices and computer programs that include computer passwords or access codes by which any part of a computer system could be accessed.[88] Article 6 further calls for the criminalization of *possessing* such items with the *intent* to commit the delineated offenses.[89] Further, Article 6 clarifies that its text "shall not be interpreted as imposing criminal liability where [there is no] purpose of committing" the delineated offenses.[90] The Convention lists "authorized testing or protection of a computer system" as an example of when this might be the case.[91]

Article 6 targets the black market for the various tools required to perpetrate cyberattacks and intrusions.[92] By criminalizing the acquisition of such tools, the

[82] *Id.* at ¶¶ 66–67.

[83] *Id.* at ¶ 67.

[84] *Id.*

[85] Explanatory Report, *supra* note 20, at ¶ 69.

[86] *Id.*

[87] Budapest Convention, *supra* note 14, at art. 6.

[88] *Id.*

[89] *Id.*

[90] *Id.*

[91] *Id.*

[92] Explanatory Report, *supra* note 20, at ¶ 71; Budapest Convention, *supra* note 15, at art. 6; Explanatory Report, *supra* note 20, at ¶ 71.

Convention's drafters aimed to cut off the problem at the source.[93] The Report defines "distribution" as "the active act of forwarding data to others," and defines "making available" as "the placing online devices for the use of others."[94] The drafters also intended "making available" to encompass "the creation or compilation of hyperlinks in order to facilitate access to such devices."[95] "Computer program," as used in Article 6, "refers to programs that are for example designed to alter or even destroy data or interfere with the operation of systems, such as virus programs, or programs designed or adapted to gain access to computer systems."[96]

After extensive debate, the drafters elected *not* to restrict the category of devices to "those which are designed exclusively or specifically for committing offenses."[97] In their view, this would be too narrow a category and, as a result, would make it more difficult to meet prosecutorial burdens of proof. This would effectively nullify the offense's criminalization.[98] Moreover, dual-use devices would have been excluded, despite presenting a similar threat. On the other hand, the drafters also rejected the idea of including *all* devices, including those both illegally and legally produced.[99] They settled on making the "intent" prong of the offense dispositive for imposing punishment and permitting States Party to decide how many devices are necessary to establish criminal liability.[100]

The drafters did not intend to criminalize possession of devices that are "produced and put on the market for legitimate purposes."[101] As an example, they suggested that those products made "to counter-attacks against computer systems" would fall into this category.[102] They manifested this intent by further requiring that there be evidence of specific intent to use the device to commit an offense described earlier in the Convention.[103]

6. Computer-Related Forgery

"[T]he input, alteration, deletion, or suppression of computer data, resulting in inauthentic data with the intent that it be considered or acted upon for legal

[93] Explanatory Report, *supra* note 20, at ¶ 71.

[94] *Id.*

[95] *Id.* at ¶ 72.

[96] *Id.*

[97] *Id.* at ¶ 73.

[98] *Id.*

[99] Explanatory Report, *supra* note 20, at ¶ 73.

[100] *Id.* at ¶¶ 73, 75.

[101] *Id.* at ¶ 76.

[102] *Id.*

[103] *Id.*

purposes as if it were authentic" is treated as a crime under the Budapest Convention.[104]

This offense was intended to "parallel" the offense of forging documents in the physical world.[105] The thinking behind this provision was that "[m]anipulations of [ ] data with evidentiary value may have the same serious consequences as traditional acts of forgery if a third party is thereby misled."[106] Article 7 further defines "computer-related forgery" as involving the unauthorized creation or alteration of stored data "so that they acquire a different evidentiary value in the course of legal transactions," citing the fact that such transactions rely on the "security and reliability of electronic data."[107] Because "national concepts of forgery vary greatly," the drafters "agreed that the deception as to authenticity refers at minimum to the issuer of the data, regardless of the correctness or veracity of the contents of the data."[108] That being said, States Party are permitted to apply the term "authentic" not only to the data's issuer, but to the data's genuineness as well.[109]

Because the data referred to in this provision is equivalent to a document with legal effects, "[t]he unauthorized 'input' of correct or incorrect data brings about a situation that corresponds to the making of a false document."[110] Accordingly, further alterations, deletions, or suppression would roughly equate to falsifying a *real* document.[111] The Explanatory Report defines "alterations" as "modifications, variations, [and] partial changes;" "deletions" as "removal of data from a data medium;" and "suppression" as "holding back [or] conceal[ing] [ ] data."[112] "For legal purposes" refers "to legal transactions and documents which are legally relevant."[113] Under Article 7, States Party are also permitted to require some kind of dishonest intent in order for criminal liability to attach.[114]

7. Computer-Related Fraud

The Budapest Convention criminalizes "the causing of a loss of property to another person by: (a) any input, alteration, deletion or suppression of computer

[104] Budapest Convention, *supra* note 14, at art. 7.

[105] Explanatory Report, *supra* note 20, at ¶ 81.

[106] *Id.*

[107] *Id.*

[108] *Id.* at ¶ 82.

[109] *Id.*

[110] *Id.* at ¶ 83.

[111] *Id.*

[112] *Id.*

[113] *Id.* at ¶ 84.

[114] *Id.* at ¶ 85.

data; (b) any interference with the functioning of a computer system, with fraudulent or dishonest intent of procuring, without right, an economic benefit for oneself or for another person."[115]

The purpose of Article 8 is to "criminalize any undue manipulation in the course of data processing with the intention to effect an illegal transfer of property."[116] Article 8(b) specifically addresses actions like "hardware manipulations, acts suppressing printouts and acts affecting recording or flow of data, or the sequence in which programs are run."[117]

Manipulations—whether of data, systems, hardware, or otherwise—under this section "are criminalized if they produce a direct economic or possessory loss of another person's property and the perpetrator acted with the intent of procuring an unlawful economic gain for himself or for another person."[118] The Report defines "loss of property" as "includ[ing] loss of money, tangibles, and intangibles with an economic value."[119] The intent required under the Article "refers to the computer manipulation or interference causing loss of property to another," and the offense further requires some "fraudulent or other dishonest intent to gain an economic or other benefit for oneself or another."[120] This limitation ensures that general, non-fraudulent commercial practices that simply happen to be economically detrimental to one party and beneficial to another are not included in this offense.[121]

8. Inchoate Offenses and Corporate Liability

Article 11 of the Budapest Convention requires States Party to criminalize both "aiding and abetting" and "attempt[ing]" the delineated offenses.[122] Separately, Article 12 provides for corporate liability for these crimes.[123] Although individuals who commit these crimes may be subject to deprivation of liberty, that punitive option appears inapplicable in cases of corporate liability under Article 13 of the Convention.[124]

The Explanatory Report provides helpful guidance on these provisions. The Convention requires States Party to criminalize aiding and abetting the

[115] Budapest Convention, *supra* note 14, at art. 8.

[116] Explanatory Report, *supra* note 20, at ¶ 86.

[117] *Id.* at ¶ 87.

[118] *Id.* at ¶ 88.

[119] *Id.*

[120] *Id.* at ¶ 90.

[121] *Id.*

[122] Budapest Convention, *supra* note 14, at art. 11.

[123] *Id.* at art. 12.

[124] *Id.* at art. 13.

commission of any offense listed in Articles 2–10, but does not require the same for attempts.[125] This is because not all of those offenses lend themselves to an "attempt" framework, and the legal systems within some States Party limit the situations in which attempt can be punished.[126] As a result, attempt is only criminalized for those offenses under Articles 3, 4, 5, 7, 8, 9(1)(a), and 9(1)(c).[127] However, States Party have the option to make a reservation refusing to criminalize attempt at all, or to only criminalize it in some cases.[128] The Explanatory Report notes that this provision aims to maximize ratification of the Convention while respecting legal traditions inherent in the States Party.[129]

As to aiding and abetting, liability attaches "where the person who commits a crime established in the Convention is aided by another person who also intends that the crime be committed."[130] For example, a service provider without any criminal intent to assist in the transmission of harmful code would not be considered to be aiding and abetting a cyberattack.[131]

Regarding corporate liability, the Explanatory Report explains that Article 12 is "intended to impose liability on corporations, associations and similar legal persons for the criminal actions undertaken by a person in a leading position within such legal person, where undertaken for the benefit of that legal person."[132] The Article also suggests that such an entity would be liable where "a leading person fails to supervise or control an employee or an agent of the legal person, where such failure facilitates the commission by that employee or agent of one of the offenses established in the Convention."[133]

Paragraph 1 of Article 12 sets forth four conditions that must be met in order to establish corporate liability. An offense described in the Convention must be committed first, by a person, second, with a leading position, third, who is acting within the scope of his or her authority, and fourth, for the benefit of the legal person.[134] The Report defines a "person who has a leading position" as "a natural person who has a high position in the organization, such as a director."[135]

[125] Explanatory Report, *supra* note 20, at ¶¶ 118–19.

[126] *Id.* at ¶¶ 118, 120.

[127] *Id.* at ¶ 120.

[128] *Id.* at ¶ 122.

[129] *Id.*

[130] *Id.* at ¶ 119.

[131] Explanatory Report, *supra* note 20, at ¶ 119.

[132] *Id.* at ¶ 123.

[133] *Id.*

[134] *Id.* at ¶ 124.

[135] *Id.*

Such persons generally have "a power of representation or an authority to take decisions or to exercise control."[136]

The Article also provides for the imposition of liability when the crime is committed by a person "acting under the legal person's authority"—that is, "one of its employees or agents acting within the scope of their authority."[137] In order to attach liability in those circumstances, three conditions must be fulfilled. First, an offense must have been committed by an employee or agent of the legal person. Second, the offense must have been committed for the legal person's benefit. Third, the failure of the "person with a leading position" to supervise the employee or agent must have made the offense possible.[138]

The Explanatory Report notes that "failure to supervise should be interpreted to include failure to take appropriate and reasonable measures to prevent employees or agents from committing criminal activities on behalf of the legal person," and it sets out a few factors that can be used to evaluate what constitutes an "appropriate and reasonable" measure.[139] Those factors include "the type of the business, its size, the standards or the established business best practices."[140] The Report is careful to note that "[t]his should not be interpreted as requiring a general surveillance regime over employee communications."[141]

9. Sanctions

As to the sanctions put in place to penalize the criminalized offenses, the Convention requires States Party to implement punishments that are "effective, proportionate and dissuasive" and include the possibility of a term of imprisonment for natural persons.[142] Sanctions available for legal persons should be similarly "effective, proportionate, and dissuasive," and may be "criminal, administrative, or civil" in nature.[143] States Party must "provide for the possibility of imposing monetary sanctions on legal persons."[144] Generally, States Party have discretion "to create a system of criminal offenses and sanctions that is compatible with their existing national legal systems."[145]

[136] *Id.*

[137] Explanatory Report, *supra* note 20, at ¶ 125.

[138] *Id.*

[139] *Id.*

[140] *Id.*

[141] *Id.*

[142] *Id.* at ¶ 128.

[143] Budapest Convention, *supra* note 14, at art. 13; Explanatory Report, *supra* note 20, at ¶ 129.

[144] Explanatory Report, *supra* note 20, at ¶ 129.

[145] *Id.* at ¶ 130.

C. Additional Relevant Concepts in the Budapest Convention

The Convention repeatedly states that the conduct it prohibits is conduct done "without right."[146] The Explanatory Report notes that this phrase "reflects the insight that the conduct described is not always punishable per se, but may be legal or justified *not only in cases where classical legal defenses are applicable, like consent, self-defense, or necessity*, but where other principles or interests lead to the exclusion of criminal liability."[147] Therefore, when it comes to implementing these principles in domestic law, the Report suggests that actions taken "without right" might refer to those taken "without authority" or those taken outside the scope of existing legal defenses.[148] The "authority" referred to in the Report can be conferred by any number of entities—the legislature, for example, or the executive, among others.[149] This suggests that States Party are able to permit certain cyber activity if it occurs in the context of an established legal defense, excuse, or justification.[150] The Report further clarifies that accessing a computer system intended to be freely and openly available to the public is always done "with right."[151]

Furthermore, the Explanatory Report clarifies that the Convention's framers did not intend to criminalize "legitimate and common activities inherent in the design of networks or legitimate and common operating or commercial practices."[152] How such exceptions would work within various domestic legal systems is, per the Report, a decision left to each individual State Party.[153]

The Explanatory Report also takes up the issue of Article 32, which permits States Party to access certain types of data without authorization.[154] According to the Report, the issue of unilateral access was discussed at length by the Convention's drafters, who considered in detail the instances in which such unilateral action would be permissible.[155] Ultimately, they concluded that preparing a comprehensive legal regime in this area was not possible at the time of the Convention's writing.[156] As a result, the two situations in Article 32 where

[146] *See generally* Budapest Convention, *supra* note 14.

[147] Paul Rosenzweig, *International Law and Private Actor Cyber Defense Measures*, 50 Stan. J. Int'l L. 103, 108–109 (2014) (citing Explanatory Report, *supra* note 20, at ¶ 38).

[148] Explanatory Report, *supra* note 20, at ¶ 38.

[149] *Id.*

[150] Rosenzweig, *supra* note 147, at 109.

[151] Explanatory Report, *supra* note 20, at ¶ 47.

[152] *Id.* at ¶ 38.

[153] *Id.*

[154] Budapest Convention, *supra* note 14, at art. 32.

[155] Explanatory Report, *supra* note 20, at ¶ 293.

[156] *Id.*

permission for unilateral access is granted are the two situations in which all drafters agreed that it would be permissible.[157]

The Report outlines these two situations: 1) when the data is publicly available anyway, and 2) when "the Party has accessed or received data located outside of its territory through a computer system in its territory, and it has obtained the lawful and voluntary consent of the person who has lawful authority to disclose the data to the Party through that system."[158] The person with lawful authority varies based upon the circumstances. According to the Report, one example would be a service provider who has the authority to retrieve data from a person's email and voluntarily disclose it to law enforcement officials.[159]

The Council further published a Guidance Note specifically related to Article 32.[160] The Note characterized Article 32(b) as "an exception to the principle of territoriality" and explained that it "permits unilateral transborder access without the need for mutual assistance under limited circumstances."[161] It clarifies that "transborder access" means "to unilaterally access computer data stored in another Party without seeking mutual assistance."[162] Article 32(b) may only be used if it is "known where the data are located," as it references "stored computer data located in another Party."[163] Therefore, to invoke the Article in the first place, one must know: 1) where the data is located, and 2) that it is located in the jurisdiction of another Party to the Convention.[164] The Note is explicit that if the data's location is unknown, or if the data is stored domestically or within a non-Party, then Article 32(b) does not apply.[165] Regarding the section's required "consent" element, the Note clarifies that the term means that the person being asked to disclose data cannot be forced to do so, nor deceived in order to induce consent.[166]

However, the Note also specifies that Article 32(b) is only to be applied within the context of criminal investigations conducted pursuant to Article 14.[167] Article 14 imposes the obligation upon States Party to implement legislation or other measures to ensure that criminal investigations can take place for the

[157] *Id.*

[158] *Id.* at ¶ 294.

[159] *Id.* at ¶ 294.

[160] Cybercrime Convention Committee, *supra* note 24, at 3.

[161] *Id.*

[162] *Id.* at 4.

[163] *Id.* at 3, 6.

[164] *Id.* at 6.

[165] *Id.*

[166] Cybercrime Convention Committee, *supra* note 24, at 6.

[167] *Id.* at 5.

offenses specifically criminalized in the Convention.[168] Authorities are to apply the same standards under Article 32(b) that they would domestically. Therefore, if a disclosure would not be permitted domestically, it would not be permitted under Article 32(b).[169]

III. Defining Active Cyber Defense

This Section constructs a working definition of active cyber defense. To do so, this Section reviews the general status of cybersecurity, discusses the various definitions, merits, and drawbacks of active cyber defense as put forth in existing literature, and defines a spectrum of cyber defensive measures. The legality of conduct on this spectrum will be analyzed in greater detail in Section IV.

A. The Cybersecurity Landscape

Nearly every internet-connected global citizen—from multinational corporations, to governments, to individuals—is vulnerable to malicious cyber activities. Cybersecurity measures aim to keep hackers from accessing "assets belonging to or connecting to an organization's network."[170] Intuitively, this means that cyber defenders work to deny network access to would-be intruders in order to protect the data contained therein.[171]

These types of attacks are perpetrated using malware. "Malware," an abbreviated form of "malicious software," is designed and used to access and, in many cases, harm a computer.[172] It can be deployed in a variety of ways—sometimes through direct attacks, but other times using stealthier means.[173] These stealthier means can include "phishing" or "spearphishing,"[174] which involve

[168] Budapest Convention, *supra* note 14, at art. 14.

[169] Cybercrime Convention Committee, *supra* note 24, at 7.

[170] *What is Cyber Security?*, FireEye Resources, http://perma.cc/T427-2UTZ.

[171] A "network" is a "system that transmits data between users," including devices belonging to those users (like phones, tablets, and computers), as well as equipment connecting those devices (like servers and routers). *Definition of: network*, PCMag Encyclopedia (2019), http://perma.cc/V4KQ-9PJX. A "server" is "[a] computer system in a network that is shared by multiple users," and a "router" is a device on a network that forwards information from one network to another. *See Definition of: server*, PCMag Encyclopedia (2019), http://perma.cc/6BM3-4VFA.; *Definition of: router*, PCMag Encyclopedia (2019), http://perma.cc/3U54-UL4L. A "system" can be conceived of as "[a] group of related components that interact to perform a task." *Definition of: system*, PCMag Encyclopedia (2019), http://perma.cc/8JS5-5X5B.

[172] *What is malware and how can we prevent it?*, Norton Security Center, http://perma.cc/LC6Y-G6QF.

[173] *Id.*

[174] "Phishing" occurs when an attacker uses a fake email sent to company employees to gain access to an otherwise protected system. "Spearphishing" occurs when an attacker specifically targets an

sending fake e-mails to employees to get malware onto an otherwise protected system, "watering holes,"[175] or any number of other methods. Part of the challenge cyber defenders face is the ever-evolving nature of these malware deployment methods.

Hackers[176] seek access to networks and systems almost constantly. In 2017 alone, 159,700 cyber incidents impacted businesses around the world, making it the "worst year ever" in terms of data breaches and cyberattacks.[177] These intrusions have a variety of purposes, ranging from activism to criminal activity to espionage, and even to acts of war. The largest-scale cyberattacks in recent years have tended to focus on the theft of personal data or intellectual property.[178] Ransomware, or malware that locks one's computer and prevents access to data until a ransom is paid, has also come into vogue.[179] The WannaCry virus that paralyzed hospitals in the U.K. demonstrates the dangers of ransomware attacks.[180] In that attack, hackers denied doctors, nurses, and hospital staff access to their computer systems, which contained patient medical records, until a ransom was paid.[181]

The number of successful cyberattacks alone indicates that the cyber defense status quo is not working. The current approach to cybersecurity tends to overwhelmingly rely on static measures—that is, passive security measures intended to deny attackers access to systems without daily human involvement.[182] Examples of these measures include firewalls, antivirus software, and intrusion detection systems.[183] This inadequacy has moved both scholars and security

employee of a certain stature so as to gain access to an identity with better access privileges than the average employee. *See* Kim Zetter, *Hacker Lexicon: What is Phishing?*, WIRED (Apr. 7, 2015), http://perma.cc/739D-KWSG.

[175] Watering hole attacks "compromise a website commonly visited by targets to hack victims' computers." Andy Greenberg, *Hackers Gain Direct Access to U.S. Power Grid Controls*, WIRED (Sept. 6, 2017), http://perma.cc/6BUT-5AYX.

[176] Although hackers can be government-sponsored or members of organized crime syndicates, the most serious challenges to cybersecurity are posed by "private criminals interested in private gain." *See* Mary Ellen O'Connell, *Cyber Security Without Cyber War*, 17 J. CONFLICT & SEC. L. 187, 191 (2012).

[177] Alison DeNisco Rayome, *2017 was 'worst year ever' in data breaches and cyberattacks, thanks to ransomware*, TECHREPUBLIC (Jan. 25, 2018), http://perma.cc/EZ48-AMRG.

[178] *See*, *for example*, Josh Horwitz & Cate Cadell, *Chinese chipmakers ambitions come unstuck with US Indictment*, REUTERS (Nov. 2, 2018), http://perma.cc/2GBB-EZQ6; *See also*, Koerner, *supra* note 8; Newman, *supra* note 7.

[179] *See* Newman, *supra* note 7.

[180] *Id.*

[181] *Id.*

[182] Robert M. Lee, *The Sliding Scale of Cybersecurity*, SANS INSTITUTE, 1, 8 (Aug. 2015), http://perma.cc/TU3K-XEFU.

[183] *Id.*

professionals to begin considering the utility of more active defense measures in order to change the dynamic between attackers and defenders.[184] Indeed, some have referred to the "scan, firewall, and patch" tradition of passive defense as the "duck and cover"[185] of modern cybersecurity.[186]

B. Constructing a Definition of Active Cyber Defense

This Comment defines "active defense" as "[t]he synchronized, real-time capability to discover, detect, analyze, and mitigate threats."[187] This definition was initially constructed by Paul Rosenzweig, a professorial lecturer in law at the George Washington University School of Law. Under this definition, defenses "operat[e] at network speed using sensors, software and intelligence to detect and stop malicious activity ideally before it can affect networks and systems."[188] As will be discussed later, this definition captures a broad variety of measures. Such breadth is particularly important because this Comment seeks to draw lines and set boundaries for what active cyber defense measures are permissible under the Budapest Convention.

That being said, a multitude of definitions of "active cyber defense" have been proposed in various spheres—from government, to the technology sector, to the military, to the legal community. These definitions include:

- "[E]lectronic countermeasures designed to strike attacking computer systems and shut down cyber attacks midstream;"[189]
- "[A]n approach to achieving cyber security predicated upon the deployment of measures to detect, analyse, identify and mitigate threats to and from communications systems and networks in real-time, combined with the capability and resources to take proactive or offensive action

[184] Kesan & Hayes, *supra* note 11, at 474; Rosenzweig, *supra* note 17, at 103–04.

[185] "Duck and cover" refers to the Cold War-era drills conducted in schools in which students were instructed to duck under their desks for cover in the event of a nuclear attack. Sarah Pruitt, *How 'Duck-and-Cover' Drills Channeled America's Cold War Anxiety*, HISTORY (Mar. 26, 2019), http://perma.cc/92RH-C9YZ. As one might infer, this tactic would not be terribly helpful in the event of a nuclear attack.

[186] Kesan & Hayes, *supra* note 11, at 474. For more on "scan, firewall, and patch," *see* Mark Ward, *Tips to Help You Stay Safe Online*, BBC NEWS (Oct. 7, 2006), http://perma.cc/RUV5-TPR4 (suggesting that readers scan their systems regularly for viruses and malware, erect and maintain firewalls to prevent unwanted intrusions on their systems, and ensure that their operating system and software are updated with the latest security patches).

[187] Rosenzweig, *supra* note 17, at 105.

[188] *Id.*

[189] Erik M. Mudrinich, Cyber 3.0: The Department of Defense Strategy for Operating in Cyberspace and the Attribution Problem, 68 A.F. L. REV. 167, 180 n.70 (2012).

against threats and threat entities including action in those entities' home networks;"[190]

- "[A] collection of synchronized, real-time capabilities to discover, define, analyze and mitigate cyber threats and vulnerabilities . . . [which] would enable cyber defenders to more readily disrupt and neutralize cyberattacks as they happen . . . [and which are] solely defensive in nature;"[191] and
- "[A] . . . category of response to cyberattacks [that] enable[s] attacked parties to detect, trace, and then actively respond to a threat by, for example, interrupting an attack in progress to mitigate damage to the system."[192]

This list is non-exhaustive. The lack of agreement on the precise contours of what qualifies as an active cyber defense measure has created extensive difficulties in categorizing and characterizing different options as lawful or unlawful under international law.

This Comment purposefully uses a broad definition of active defense in order to more specifically define what conduct is and is not permissible under international law. Accordingly, it will utilize the definition from Paul Rosenzweig as set forth above:

> [T]he synchronized, real-time capability to discover, detect, analyze, and mitigate threats. It operates at network speed using sensors, software and intelligence to detect and stop malicious activity ideally before it can affect networks and systems. While intrusions may not always be stopped at network boundary, an entity may operate and improve upon its advanced sensors to detect, discover, map, and mitigate malicious activity on an entity's network.[193]

This comprehensive definition of the term permits a thorough examination of all measures that could conceivably be considered "active defense," even those that seem facially "passive."

Understanding this definition requires fleshing out a few finer distinctions. First, the process of pursuing an "active defense" can be broken down into three steps: 1) detecting an intrusion, 2) identifying its origin, and 3) responding in some form.[194] Second, active defense tactics can be divided into "internal" and "external" measures.[195] Internal measures are those taken on one's own network. Examples include monitoring network traffic for irregularities, blocking incoming

[190] Hoffman & Levite, *supra* note 1, at 7–8.

[191] *Active Cyber Defense (ACD)*, DEFENSE ADVANCED RESEARCH PROJECTS AGENCY, http://perma.cc/PRC9-HKFM.

[192] Kesan & Hayes, *supra* note 11, at 475.

[193] Rosenzweig, *supra* note 147, at 105.

[194] Kesan & Hayes, *supra* note 11, at 475.

[195] Rosenzweig, *supra* note 147, at 105–06.

traffic selectively based on its source, and constructing traps for would-be hackers.[196] External measures are those undertaken outside of one's network—whether they be on an adversary's network or one belonging to a neutral third party. [197] These tactics could include identifying the sites or servers from which suspicious network activity originated, modifying that originating server in some way to halt its activities in relation to one's network, accessing data from an adversary on his or her home turf, or even outright attacking the adversary's servers to cause damage.

In order to separate legal conduct from illegal conduct, these activities must be categorized along some sort of spectrum. Some authors have undertaken this task in the past.[198] This Comment will focus less on the legality of individual, discrete measures and more on crafting a set of categories for evaluating the legality of active cyber defense measures.

C. Defining a Spectrum of Active Cyber Defense

To assist in understanding the point at which legal conduct becomes illegal conduct, this Comment outlines a spectrum of cyber defense activities divided into three categories: passive, active-passive, and active. Section IV of this Comment analyzes each category's permissibility under international law.

Passive defenses are used entirely within the boundaries of one's own network and never involve reaching beyond it. Such defenses include installing and upgrading antivirus software, constructing firewalls, segmenting certain critical servers in a way that prevents connection to the internet, and engaging in basic "cyber hygiene" practices.[199] This category also encompasses blocking incoming traffic to one's network, whether selectively or universally, and employing notification beacons that alert system administrators to attempts to remove files or otherwise tamper with the network.[200]

Active defenses mark the opposite end of the spectrum. Carol Hayes, a research fellow at the University of Illinois College of Law, and Jay Kesan, the H. Ross and Helen Workman Research Scholar at the University of Illinois College of Law, offer an apt characterization for tactics in this category. These types of

[196] *Id.*

[197] *Id.*

[198] *Cf.* Center for Cyber & Homeland Security, *supra* note 2, at 10–11; Hoffman & Levite, *supra* note 1, at 9.

[199] Cyber hygiene is defined as thinking proactively "to resist cyber threats and online security issues." *Good Cyber Hygiene*, NORTON SECURITY CENTER, http://perma.cc/GMB9-VGZW (giving examples of cyber hygiene).

[200] *Cf.* Center for Cyber & Homeland Security, *supra* note 2, at 10–11; Hoffman & Levite, *supra* note 1, at 9.

defenses tend to be "offensive actions undertaken with the goal of neutralizing an immediate threat rather than retaliating."[201] These measures are used outside of the network and impact external systems, whether belonging to an attacker or to a neutral third party. Examples of active tactics include hacking an adversary in order to retrieve stolen information,[202] disrupt its network, or damage its network.[203]

Active-passive defenses lie somewhere between these endpoints. These measures encompass those like digital "dye-packs" or other devices that enable defenders to track data taken from their networks,[204] hunt and expel intruders on the network, funnel potential adversaries to decoy networks, and other actions taken to investigate and attribute intrusions.[205] This category will require the most intensive, case-by-case analysis in order to establish permissibility under international law. That analysis must be guided by the principles set forth in the Budapest Convention, as it is the only legally binding international instrument addressing the question of when an act in cyberspace becomes criminal.[206]

It is entirely possible that the same tactic could appear in all three categories depending on how a given tactic is built and operated. Take, for example, the honeypot from the opening scenario. If that honeypot functioned solely to permit defenders to observe network traffic and has zero effect on any other system, it is likely considered passive. If that honeypot infected intruders with a tracking beacon that allows cyber defenders to determine where a particular intruder is based, it would be characterized as active-passive. Finally, if that honeypot attached a virus that would delete all data on the intruder's home system, it would qualify as an active defense.

Clearly, the term "active cyber defense" is very broad and cannot be characterized as "legal" or "illegal" on its face. Rather, these finer distinctions permit a more nuanced understanding of when and why a particular tactic might rise to the level of an international crime.

[201] Kesan & Hayes, *supra* note 11, at 475.

[202] Center for Cyber & Homeland Security, *supra* note 2, at 11.

[203] Hoffman & Levite, *supra* note 1, at 8–9.

[204] *Id.*

[205] Center for Cyber & Homeland Security, *supra* note 2, at 10–11; Hoffman & Levite, *supra* note 1, at 8.

[206] *See* Section II, *supra.*

IV. Identifying Active Defense Strategies Permissible under International Law

Section II undertook a comprehensive interpretation of the Budapest Convention and relevant explanatory documents. That interpretation highlighted several offenses related to computer data, forbidding any creation of false or otherwise "inauthentic" data, and outlawing the actions that lead to data loss via "input, alteration, deletion, or suppression of computer data" or "any interference with the functioning of a computer system."[207] Further, the Convention requires States Party to criminalize the creation of tools that could be used to unlawfully access or interfere with systems or data.[208]

In order to rise to a criminal level, the Convention makes clear that these offenses must be committed intentionally and "without right"—that is, without authorization or outside the parameters of legal defenses acceptable within a State Party's domestic legal system.[209] The Convention also provides some clarity on the jurisdictional ambiguities existing in cyberspace. It requires States Party to establish, via domestic legislation or other measures, that each State Party has jurisdiction over offenses committed in its territory, onboard a ship flying its flag, or onboard an aircraft registered under its laws.[210] States Party must further establish jurisdiction over offenses committed by one of a State Party's nationals if the offense was committed somewhere that criminalizes the underlying conduct or somewhere outside any State Party's territorial jurisdiction.[211] This is done via the passage and implementation of domestic legislation in that State Party.[212]

Applying the law as set forth in the Budapest Convention to the spectrum of cyber defenses laid out in Section III of this Comment clarifies when active cyber defense measures are considered cybercrimes under international law. This Comment proposes using a "stoplight framework" to categorize various actions. That is, defenders should freely implement certain measures (green light), should use caution when considering more ambiguous ones (yellow light), and should never undertake others (red light).

Understanding where the line is drawn between legal and illegal conduct in cyberspace permits cyber defenders to have a general sense of what is and is not permissible when putting together defense strategies. The ultimate test, however, as to whether a particular tactic is illegal under international law is whether it rises to the level of an offense that the Budapest Convention seeks to criminalize.

[207] Budapest Convention, *supra* note 14, at art. 7–8.

[208] *Id.* at art. 6.

[209] Explanatory Report, *supra* note 20, at ¶ 38.

[210] Budapest Convention, *supra* note 14, at art. 22.

[211] *Id.*

[212] *See generally* Explanatory Report, *supra* note 20, at ¶ 235–236.

The stoplight framework proposed in this Comment easily maps on to the three categories of actions—passive, active-passive, and active—based upon the number of Budapest Convention offenses potentially implicated in each category. As is explained in greater detail below, defenders can confidently employ passive measures, ought to approach active-passive measures with caution, and should refrain from using active measures in order to avoid engaging in conduct that is illegal under international law.

A. Passive Cyber Defense Measures: Proceed with Confidence

In general, defensive measures in the passive category will be permissible. Defenses in the passive category could run afoul of only a few Convention offenses: illegal interception of data, data interference, or misuse of devices.[213] Measures that fall into this category include the use of antivirus software, the construction of firewalls, server segmentation and air-gapping,[214] and engaging in basic cyber awareness activities like password protection and wariness in opening emails.[215] They may also include, for example, blocking incoming traffic to one's network or utilizing notification beacons that alert administrators to any attempts to tamper with or remove files.[216]

Considering that such defenses involve little to no ongoing engagement by cyber defenders and generally are only deployed within the defender's own network, the likelihood that they would constitute offenses defined by the Budapest Convention is minimal. If they were to rise to the level of a potential violation, the only offenses that would likely be implicated are the illegal interception of data, data interference, or misuse of devices.

One can imagine a scenario in which passively surveilling intruder communications *could* constitute an illegal interception of data, because this behavior is roughly analogous to the "cyber wiretaps" the Convention's drafters sought to prevent.[217] However, this scenario seems highly unlikely because, under the Budapest Convention, the illegal interception offense appears to require the offender to have accessed another computer system in order to be guilty of committing that crime.[218] Monitoring one's own system and the communications

[213] *See* Budapest Convention, *supra* note 14, at arts. 3, 4, 6.

[214] An "air gap" "refers to computers or networks that are not connected directly to the internet or to any other computers that are connected to the internet." Kim Zetter, *Hacker Lexicon: What Is an Air Gap?*, WIRED (Dec. 8, 2014), http://perma.cc/ZH9P-YQXL.

[215] *See* Lee, *supra* note 182, at 7–8.

[216] *Cf.* Center for Cyber & Homeland Security, *supra* note 2, at 10–11; Hoffman & Levite, *supra* note 1, at 9.

[217] Explanatory Report, *supra* note 20, at ¶ 51.

[218] *Id.* at ¶ 53.

on it—even if those communications take place between intruders—would seem not to meet that threshold.

Similarly, although these measures might result in the suppression of computer data, thereby implicating the data interference offense, they likely would not constitute a true violation. The Explanatory Report clarified that the suppression offense is meant to target actions that prevent people who "ha[ve] access" to the computer containing the data at issue.[219] While a hacker might technically have "access," such access likely is not legitimate—suggesting that using passive defense actions does not actually violate this provision. Finally, the misuse-of-devicesdevice-misuse offense is mainly implicated only insofar as such devices are used to achieve illegal interception or interference with data.[220] If neither of those offenses is committed, a device-misuse offense is likely not committed either.

Accordingly, actions in the passive category get a green light. Because of the extremely low likelihood that they would be considered "offenses" as defined in the Budapest Convention, defenders can generally employ them without concerns about their illegality under international criminal law.

B. Active-Passive Cyber Defense Measures: Proceed with Caution

By contrast, tactics in the active-passive category are more likely to implicate a greater number of Budapest Convention offenses—nearly all of the offenses discussed in this Comment, in fact. Although active-passive measures are unlikely to result in computer-related fraud, they may well lead to illegal access to computer systems, illegal interception of transmissions, data interference, system interference, misuse of devices, or computer-related forgery.

This category includes measures that track data taken from networks, [221] identify and remove network intruders and lead adversaries to decoy networks, as well as efforts to investigate and attribute intrusions.[222]

On its face, the Convention seems very clear: unauthorized access to or interference with systems or data is strictly prohibited.[223] Under such a reading of the Convention, any access to an external system without the prior agreement of its owner appears to constitute a crime under international law. Consider, for example, a honeypot that deployed a virus back to visitors' networks. Infecting

[219] *Id.* at ¶ 61.

[220] Budapest Convention, *supra* note 14, at art. 6.

[221] Hoffman & Levite, *supra* note 1, at 8–9.

[222] Center for Cyber & Homeland Security, *supra* note 2, at 10–11; Hoffman & Levite, *supra* note 1, at 8–9.

[223] Budapest Convention, *supra* note 14, at arts. 4–5.

another network with malware would constitute illegal access to that system, at the very least. More violations could ensue, depending on how the malware was built, what virus or code it delivered to the new system, and how the virus or code functioned in the new environment.

The Explanatory Report, however, indicates three potential caveats to this conclusion. The first is the Report's intimation that actions falling within the ambit of a domestic justification for committing a crime—like necessity, self-defense, or consent—are permissible.[224] The second is the Report's clarification that the Convention did not intend to criminalize common commercial practices, and indeed considered such practices to occur "with right."[225] Finally, at least with regard to the system interference offense, the Report suggested that States Party are permitted to define some minimum amount of damage that must take place before an intrusion is considered an "interference" under the Convention.[226]

These caveats open three possible avenues for the use of active-passive defenses. First, an entity may lawfully access the intruder's system under the domestic legal conceptions of self-defense or necessity in that entity's jurisdiction.[227] To be clear, the ability to invoke either defense would depend on meeting the standards required by the domestic law of the State Party prosecuting the violation. This is a function of the Budapest Convention's structure—the Convention requires States Party to criminalize certain conduct within their respective domestic laws.[228] Generally, these defenses are raised in response to criminal prosecution. Because the Convention depends on domestic law as an enforcement mechanism, there will almost certainly be some variation in when and how these principles may be invoked. However, they may be an available defense, depending on an entity's circumstances.

The second possible avenue for implementing active-passive measures is within existing commercial and industry practice. One could argue that, as governments prove less and less willing to assume defense responsibilities for private companies, an "industry practice" of active-passive cyber defense is in the early stages of emerging.[229] Indeed, the concept of active cyber defense has been under discussion in the industry for nearly two decades. In 1999, a web service company hosting the World Trade Organization's servers defended itself against

[224] Explanatory Report, *supra* note 20, at ¶ 38.

[225] *Id.* at ¶ 62.

[226] *Id.* at ¶ 67.

[227] *Id.*, at ¶ 38. Although technically consent to entry is a justification, it seems unlikely that an intruder would consent to its own target's investigation.

[228] Perloff-Giles, *supra* note 4, at 217–18.

[229] *See* Hoffman & Levite, *supra* note 1, at 1.

a denial of service attack by reflecting the incoming traffic back at its source.[230] This would be an active defense measure under *any* definition. Later, in 2004, Symbiot Inc. made a product that could "execute appropriate countermeasures" against a cyber threat and would even provide a graded range of response levels that could be matched to the level of the attack.[231]

More recently, private sector entities around the world have been responding to the uptick in cyberattacks by implementing some forms of active cyber defense, and many entrepreneurs have been very willing to assist.[232] Increasing numbers of cybersecurity professionals advertise active-passive defense measures, like honeypots.[233] Some companies even outsource their active cyber defenses when such measures would not be legal under the domestic law of their own country.[234]

The financial sector, in particular, is motivated to innovate, as it faces "the most severe and persistent threats."[235] For example, an entire "stealth market" in the Netherlands exists to enable banks and other financial services companies to hire others to target their attackers' servers.[236] Some governments even seem open to the idea of private sector active defense; for example, the United Kingdom included in its *National Cyber Security Strategy 2016–2021* a statement that it "will draw on its capabilities and those of industry to develop and apply active cyber defense measures to significantly enhance the levels of cyber security across UK networks."[237]

Clearly, a "gray market"—not quite a legitimate market, but not fully a black market, either—for active cyber defense measures is growing, aided and abetted by the legal ambiguities in this area.[238] This could satisfy the requirement of an existing "industry practice" of active self-help measures that could potentially justify limited active-passive actions taken to investigate and attribute a cyber intrusion.[239]

Finally, the Convention leaves the door open for States Party to establish some minimum amount of damage that must occur before criminal liability will attach for a system interference.[240] Depending on the State Party from whence the cyber intrusion originated, this could create space for some minimal investigation

[230] Center for Cyber & Homeland Security, *supra* note 2, at 8.

[231] *Id.*

[232] Hoffman & Levite, *supra* note 1, at 4.

[233] *Id.* at 15.

[234] *Id.*

[235] *Id.*

[236] *Id.*

[237] Her Majesty's Government, National Cyber Security Strategy 2016–2021, 2016, ¶ 1.8 (UK).

[238] Hoffman & Levite, *supra* note 1, at 4.

[239] Explanatory Report, *supra* note 20, at ¶ 62.

[240] *Id.* at ¶ 67.

on the part of the targeted entity for the purposes of determining the intrusion's source. However, it would be difficult to undertake such an investigation with a particularly high level of confidence, as the constraints on activities would vary by State Party. This freedom for States Party to establish threshold damage levels could be a useful tool for signaling their attitude towards domestic cyber actors to the international community. Those states wishing to disincentivize entities within their jurisdiction from undertaking cyberattacks could set a very high bar for the minimum amount of damage that must occur before criminal liability will attach—thereby permitting more extensive investigation and activity on the part of targeted entities. On the other hand, states wishing to protect the ability of entities within their jurisdiction to operate unfettered in cyberspace could dramatically lower the bar.

The idea that any minimal efforts to investigate cyber intrusions would constitute cybercrimes under international law seems ill-considered from an efficiency standpoint. Although such strict interpretation may have made sense when the Convention was drafted in the 1990s, it hardly seems in step with the current status of cyberspace. With tens of thousands of cyberattacks and intrusions targeting businesses every year,[241] the idea that only national law enforcement can conduct any level of investigation in order to attribute the activities seems unwieldy at best and unworkable at worst. In light of the revolution the internet has undergone since the Convention's drafting, it makes sense to interpret the Convention as permitting private entities to undertake some minimal level of investigation that does not cause damage to external servers for the purposes of attribution. This interpretation can be achieved via any of these three caveats.

Although measures falling into the active-passive category are more likely to constitute an offense under the Budapest Convention, it is possible that those employing them would have some kind of legal defense or exception to justify their actions. Accordingly, such measures should be implemented with caution, and fall into the "yellow light" category of this Comment's suggested stoplight framework.

C. Active Cyber Defense Measures: Do Not Proceed

Active cyber defenses are the most difficult to justify under the strictures of the Budapest Convention, as they can implicate every single Convention offense discussed in this Comment.

Measures falling under this category can be generally summarized as "offensive actions undertaken with the goal of neutralizing an immediate threat

[241] Rayome, *supra* note 177.

rather than retaliating."[242] These tactics are used on and impact external systems—whether belonging to an attacker or a neutral third party—with the specific purpose of stopping a particular intrusion or attack. This means that they will almost certainly access and interfere with systems in direct violation of the Budapest Convention[243] and are much more likely to actually damage networks and systems falling under another State Party's jurisdiction than, for example, purely investigative measures would be.

As a result, these measures have a hard time fitting into any of the three available justifications for the employment of measures external to one's own network. The self-defense and necessity justifications might remain available to an entity in the case of an exceptionally serious cyberattack; however, an attack of that magnitude is precisely when national law enforcement authorities would likely get involved.[244] The "common commercial practice" justification is likely not sufficiently strong to excuse outright external attacks; indeed, a 2012 survey conducted at the Black Hat USA cybersecurity conference "found that 36% of respondents claimed to have engaged in retaliatory hacking."[245] This is probably not deeply rooted enough to constitute a "common commercial practice" for the purposes of purely active measures. Finally, unless a Party has set an incredibly high bar domestically for the damage required to establish system interference, an outright active measure is far more likely to exceed this threshold. Its purpose is to bring an end to an intrusion or an attack; therefore, it is, in a sense, created to cause damage.

Thus, with active measures, the number of potential offenses implicated is not offset by the availability of legal defenses. Accordingly, they belong in the "red light" category—that is, cyber defenders ought to refrain from using them. These actions are very likely to constitute illegal conduct under international law.

V. CONCLUSION

The Budapest Convention is the only binding international law defining which actions are permissible in cyberspace and which are not. It is imperfect, and suffers from the failures of imagination that characterize late twentieth century attempts to regulate the internet. However, it is the international community's only definition of when behavior in cyberspace becomes criminal and what justifications might be relied upon to excuse certain actions. Essentially any

[242] Kesan & Hayes, *supra* note 11, at 475.

[243] *See* Budapest Convention, *supra* note 14, at arts. 2, 5.

[244] *See*, *for example*, Perlroth, *supra* note 9 (discussing the hacking operation undertaken against the Wolf Creek Nuclear Operating Corporation in Kansas in the United States and the subsequent joint investigation conducted by the U.S. Federal Bureau of Investigation and the U.S. Department of Homeland Security).

[245] Hoffman & Levite, *supra* note 1, at 15.

unauthorized access to, or interference with, computer systems or data is criminalized under the Convention if it does not fall into three categories of exceptions: a legal defense recognized under domestic law, a common commercial practice, or an action that falls below the threshold set by individual States Party.

Purely passive measures are highly unlikely to implicate any offense listed in the Convention, as they never venture outside the confines of an entity's own network. Active-passive measures, or those that are internal to a network with possible external repercussions, can fall under the umbrella of one of the Convention's justifications. Finally, although it is possible that an attack would be so egregious that a purely active measure specifically targeting external networks would be justified, this scenario is highly unlikely. Therefore, active measures are nearly always unlawful under international law and should be avoided.

It is difficult to blame private companies for wanting to innovate when it comes to defending their assets in cyberspace. They face an unprecedented threat environment, with tens of thousands of cyberattacks directed at businesses each year.[246] As they scramble to defend themselves, intruders claim significant victories, whether by stealing personal data, humiliating a company's employees by releasing their emails, or even gaining a foothold into nuclear power systems critical for national security. The stakes of cyber defense are extremely high, and the Budapest Convention ought to be interpreted accordingly.

[246] Rayome, *supra* note 177.