UiO : Faculty of Law
University of Oslo

The Legality of Hacking Back under International Law

Candidate number: 9015

Submission deadline: 1 June 2020

Number of words: 15.612

Table of contents

1 Introduction

1.1 Opening remarks

hacking

(v.) *to gain illegal access to (a computer network, system, etc).*[1]

(v.) *to secretly find a way of looking at and/or changing information on somebody else's computer system without permission.*[2]

One of the first things of note when entering the field of cybersecurity is the lack of universal definitions for most of the key concepts. Part of this lack of universality stems from the overwhelming number of existing definitions.[3] While this may partly be the consequence of different interpreters not reaching consensus, the high pace at which technologies evolve is a good justification. After all, a less specific terminology can keep definitions more future proof by encompassing the basic ideas of information and communication technology (ICT) instead of only the currently available techniques and technologies.

Nevertheless, some handholds exist in literature. Cybersecurity from a Western perspective can be defined as protecting the confidentiality, integrity and availability of data and information systems.[4] China and Russia on the other hand argue for a broader definition that places the State-interest at its center. The disparity between the views on cybersecurity of the most powerful countries is a reason that no international instrument for cybersecurity can be agreed on.[5]

Further indication of the Wild West nature of cyberspace is the 'Declaration of the Independence of Cyber Space', a document written by John Perry Barlow.[6] While this Declaration is not a legal document, it properly illustrates the view that cyberspace is non-territorial and should fall outside of the scope of State sovereignty. Regardless, even though the flow of information through cyberspace is not always contained within specific jurisdictions its effects can be felt in the physical world.

Cross-border activities in the physical world are normally governed by distinct legal systems for public and private entities. The question is whether that distinction makes sense in a field where the distinction between public and private parties is not as clear. Some private compa-

---

[1] Merriam Webster Dictionary online version, <https://www.merriam-webster.com/dictionary/hack> last accessed 28-1-2020.

[2] Oxford Dictionary online version, <https://www.oxfordlearnersdictionaries.com/definition/english/hack_1?q=hack> last accessed: 28-1-2020.

[3] Worku Gedefa Urgessa, 'Multilateral Cybersecurity Governance: Divergent Conceptualizations and its Origin' (2020) 36 Computer Law & Security Review 2.

[4] Urgessa (n 3) 3.

[5] Urgessa (n 3) 6.

[6] J.P. Barlow, 'A Declaration of the Independence of Cyber Space' (*EFF* 1996) <https://www.eff.org/nl/cyberspace-independence> accessed 1 June 2020.

nies have more financial power[7] or technical capabilities[8] than States. A majority of the critical infrastructure in some States is owned by private entities, from ICANN to software and hardware developers.[9] Consequently, it is not always clear who is or should be responsible for the cyber defence of that critical infrastructure.[10] At the same time, the thresholds for conducting cyber-attacks are very low as they are not resource-intensive and it is easy to obscure ones identity.[11] Yet the networks on which cyberspace exists are omnipresent and as such provide vulnerabilities to critical infrastructure of States, with consequences reaching both their public and private organs.

Since there is no encompassing treaty on cyberspace, the question arises what the rules for protection of one's interests are. Is it legal to hack back under in international law?

To properly answer that question it has been divided in different chapters answering sub questions. The second chapter briefly explains the sources of international law that are relevant to discuss the legal status surrounding hacking back by State organs or their subsidiaries. The third chapter outlines general topics such as the qualification of hackers and illustrates the practical manifestations of cyber-attacks and counterattacks. The fourth chapter examines the grounds in international law that preclude wrongfulness of actions, both for States and non-State actors. That chapter is the core of this paper and should provide clarification on the main question of the legality of hacking back under international law.

With that in mind, this paper treats hacking back, counter hacking and active cyber defence as synonyms that mean technical measures aimed at not merely preventing unwanted access to data or systems, but actively engaging such unwanted attempts to access data or systems. The third chapter of this thesis will elaborate and clarify this definition by examining currently existing and employed measures to hack back.

1.2 Scope

Some opening remarks should be devoted to the scope of this thesis.

The research for this paper is comprised of a combination of legal scholarly articles, legal documents[12] and case law, articles with a more technical background and various in-depth news articles.

---

[7] Parag Khanna, 'These 25 Companies are More Powerful Than Many Countries' (*Foreign Policy*, 15 March 2016) <https://foreignpolicy.com/2016/03/15/these-25-companies-are-more-powerful-than-many-countries-multinational-corporate-wealth-power/> accessed 17 May 2020.

[8] Kristen E. Eichensehr, 'Public-Private Cybersecurity' (2017) 95 Tex L Rev 467, 497ff.

[9] Madeline Carr, 'Public-private Partnerships in National Cyber-security Strategies' (2016) 92 International Affairs 43, 52.

[10] Eichensehr (n 8) 499.

[11] Jay P Kesan and Carol M Hayes, 'Mitigative Counterstriking: Self-Defense and Deterrence in Cyberspace' (2012) 25 Harv J L & Tech 429, 438.

[12] *inter alia* the Tallinn Manual 2.0, the UN Charter and the ILC Articles on State Responsibility.

Only actions meeting the threshold of an internationally wrongful act are discussed.[13] Responses to cyber-attacks that can be considered unfriendly, but not unlawful are in principle beyond the scope of this article. Such responses could qualify as retorsion and do not carry the same weight and sense of urgency as cyber-attacks. Since retorsion concerns lawful responses, an analysis on that point could only dive into the question of which cyber activities pass the threshold of unlawful, but would not offer answers to the question on the legality of hacking back.[14]

The focus thus lies on the injured party and the legal regimes available in international law that can offer a basis to hack back. Instances where actions of the attacker are expanded on are limited to when it is necessary to illustrate them within a framework to clarify the situation of the response.

A number of sources of international law have been examined to assess whether there are legal grounds for hacking back by a State or non-State actor. Firstly, as the Tallinn Manual 2.0 is considered to be a respected and thorough collection of customary international law applicable to cyberspace,[15] it functions as the starting point for the analyses in this paper. While the Manual cannot qualify as *lex specialis* since it has no status of law, it is a worthy alternative because of its comprehensive nature and field of focus. The second layer of sources consists of case law and treaties, complemented with relevant general sources of public international law. The third layer is made up of scholarly writing to identify gaps, critique and possible expand on or augment the principles mentioned in the Tallinn Manual 2.0. Of note is that during the finalization of this paper, a publication with a similar scope and approach by Lahmann came out.[16] Time constraints have not permitted the current writer to do an in-depth analysis of Lahmann's latest work, but this paper has taken note of his points and where possible included them.

This paper's examination on the legality of hacking back under international law should not be read as to preclude other possible responses to cyber-attacks. The conclusions herein have no bearing on either practical measures qualified as retorsion, or legal avenues in the form of civil, criminal or international legal procedures.

2 Sources of International Law

"Law is that element which binds the members of the community together in their adherence to recognised values and standards".[17]

---

[13] The legal definition of an internationally wrongful act is elaborated on in Chapter 4.

[14] Michael Schmitt (ed), *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations* (CUP 2017) 174.

[15] Henning Lahmann, *Unilateral Remedies to Cyber Operations* (CUP 2020) 21.

[16] Lahmann (n 15).

[17] Malcolm Shaw, *International Law* (6th edn, CUP 2011) 1.

State sovereignty was once considered complete and States accepted no limitations to their conduct.[18] World War II and the failure of the League of Nations brought about some change in this attitude.[19] States realised that they needed to expand on trends started in the nineteenth century and accept an international legal order based on international agreements, regulations and organizations.[20] This is the context in which legality of actions on the international stage, such as hacking back, will be contemplated. Hacking back is a practical term describing a certain type of action and not a legal concept. The legality of responding to a cyber-attack can be tested under different regimes of international law.

Whether certain conduct is legal is a question that may be answered by looking at whether an action or omission is in accordance with the law. While international law lacks a clear system of governance with a legislative, executive and judiciary branch, it is widely accepted that the sources of international law are outlined in Article 38(1) of the Statute of the International Court of Justice.[21] In the absence of an overarching, authoritative Treaty on cyber conduct, this paper examines several different sources of law. These are the primary sources: a) treaties, such as the UN Charter; b) customary international law, which the Tallinn Manual 2.0 aims to present; and c) 'the general principles of law recognized by civilized nations'.[22]

An example of a rule of international law is how States have accepted limitations on when and how they can use force. The rules on the use of force by States are considered customary international law and laid down in the UN Charter.[23] Another relevant field of law is that on State responsibility, governed by the Articles on State Responsibility.[24] The relevance of the law on State responsibility for this paper lies in its broader focus than just conduct that constitutes a use of force. Most cyber-attacks will not reach the use of force threshold specified in the UN Charter, but can nevertheless impact State or private interests. The Articles on State responsibility offers circumstances precluding wrongfulness besides just self-defence.[25] In other words, situations in which hacking back or other responses that would normally be illegal, are justified.

Even though States are the original subjects of international law, there are specific areas of international law that can be applicable to non-State actors.[26] Some of these are human rights

[18] Shaw (n 17) 9.

[19] Shaw (n 17) 30.

[20] Shaw (n 17) 29, 30.

[21] Shaw (n 17) 70.

[22] Article 38(1) Statute of the International Court of Justice.

[23] Articles 2(4) io. 51 UN Charter; Eric Talbot Jensen, 'Computer Attacks on Critical National Infrastructure: A Use of Force Invoking the Right of Self-Defence' (2002) 38 STAN. J. INT'L L. 207, 217.

[24] International Law Commission ('ILC') Articles on the Responsibility of States for Internationally Wrongful Acts ('Articles on State Responsibility') (2001).

[25] Chapter V Articles on State Responsibility.

[26] Shaw (n 17) 261.

law (protection of individuals),[27] humanitarian law (obligations for non-State combatants)[28] and international criminal law (where individuals can be judged for *inter alia* committing war crimes and genocide)[29].[30] Another example is how the rights of individuals are protected under the GDPR and the right of individuals and non-profit organizations to enforce those protections through complaints and judicial remedies.[31] However, none of these areas can offer clarity on non-State conduct relating to hacking back. As witnessed by the NotPetya malware, a single attack can have detrimental consequences for private actors in numerous jurisdictions at the same time, making application of international private law more complicated than ever. Since private business are often victims of suspected international cyber-attacks it makes sense to determine if justifications for hacking back under international law can also apply to non-State actors. That is why non-State actors are included in the scope of this paper. Answering on the legality of hacking back for non-State actors seems to warrant examination beyond treaty and customary international law. As described above, those sources are either State-centric or specialised regimes not easily applicable to cybersecurity. Paragraph 4.3.2 on self-defence by non-State actors thus focuses on the third source of law: general principles of law recognised by nations.

3 Hacking

3.1 What is hacking?

3.1.1 Type of hacks

In order to get an understanding of the potential consequences of a hack and the practical context of this paper, the most common types of cyber-attacks are outlined below.[32]

- Malware: this is an umbrella term that includes the different types of malicious software used to impact the availability (ransomware), confidentiality (spyware) or integrity (virus, worms and other destructive types of malware) of data or networks.
- Phishing: this is the technique of using communication that falsely represents trusted contacts in order to get people to click a link or download a file. It can be used to gather information directly or as a means of installing malware.

---

[27] Shaw (n 17) 257ff.

[28] Shaw (n 17) 1167ff.

[29] Shaw (n 17) 397ff.

[30] See also Hessbruegge (n 165) 68.

[31] Articles 77-80 Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC ('General Data Protection Regulation').

[32] 'What are the Most Common Cyber Attacks' <https://www.cisco.com/c/en/us/products/security/common-cyberattacks.html#~how-cyber-attacks-work> accessed on 1 June 2020; 'What is Cyber Security' <https://www.kaspersky.com/resource-center/definitions/what-is-cyber-security> accessed on 1 June 2020.

- Man-in-the-middle attack: unsecure (public) networks can be abused by intercepting information send over that network by users. Man-in-the-middle attacks can also follow from installed malware that allows the hacker to process a victim's information.
- Denial-of-Service attack (DoS): causing a system or network to be unavailable by flooding it with information traffic. DoS attacks are often conducted through 'botnets' (a network of compromised devices) and then referred to as Distributed DoS or DDoS attacks.
- Zero-day exploit: these are vulnerabilities in networks, software or hardware that have not been patched yet and can be used to install malware or gain access.
- Structured Language Query ('SQL') Injection: the use of malicious SQL statements in data-driven applications to gain access to a database and steal the information therein.

These different methods of attack should not be seen as mutually exclusive or detached since they are often used together.[33] A couple of cases are mentioned hereafter to get a better grasp of the practical application and consequences of these hacking tools.

- NotPetya: a group of presumably Russian hackers had digitally infiltrated a Ukrainian company that provided updates for *inter alia* administrative software. They uploaded a malware called NotPetya into one of the updates. As the update was downloaded by users around the world, the malware was spread. It was built combining two know exploits and imitated the ransomware Petya. Despite pretending to be ransomware, NotPetya destroyed systems, there was no way to restore files by paying a sum and receiving an encryption key. The consequences: over ten billion US dollar in damages by shutting down shipping operations worldwide and affecting more than 300 companies in Ukraine alone. [34] While hospitals were also affected, there is no mention in the media of lives being lost because of the NotPetya attack. As will be explained further in chapter 4, the effects of a hack play a role in the legal concepts that could justify hacking back. Especially if a worm similar to NotPetya causes more destruction or the loss of lives, this is the kind of cyber-attack that could invoke the right to self-defence.[35]
- Stuxnet: This virus is considered to be the first offensive cyber weapon.[36] It was built based on information of software and hardware used in an Iranian nuclear facility with the aim of destroying the centrifuges in that facility. An alleged combined effort of

[33] Kesan and Hayes (n 11) 433.

[34] Andy Greenberg, 'The Untold Story of NotPetya, the Most Devastating Cyberattack in History' (*Wired*, 22 August 2018) <https://www.wired.com/story/notpetya-cyberattack-ukraine-russia-code-crashed-the-world/> accessed 1 June 2020.

[35] See an analysis on the right to self-defence in chapter 4.3.

[36] Kim Zetter, 'An Unprecedented Look at Stuxnet, the World's First Digital Weapon' (*Wired*, 11 March 2014) <https://www.wired.com/2014/11/countdown-to-zero-day-stuxnet/> accessed on 1 June 2020.

American-Israelian programmers and Dutch human intelligence operations created a virus and smuggled it into the closed network of the Natanz facility.[37] Because of the physical damage Stuxnet caused, the virus is often used as one that might meet the threshold of an 'armed attack', if the scale of the damage had larger.[38]

- Cozy Bear: A group of Russian hackers referred to as Cozy Bear was monitored by a Dutch intelligence agency for over three years from 2014 to 2017. During this period the agency witnessed how Cozy Bear used phishing emails to collect passwords and gain entry to the United States' Ministry of Foreign Affairs, the White House and servers of the Democratic Party.[39] While not causing direct damage, the above illustrates how phishing operations can provide access to some of the most prestigious government institutions.
- Zero-day exploits can be both used to gain access to systems or as offensive means. The Stuxnet virus described above contained five zero-day exploits.[40] There is a big market for such exploits in which both hackers, private companies and governments are active. Sometimes the zero-day exploits are sold to companies in order to make them aware of security flaws, but in other cases parties purchase them in order to create malware to benefit from the exploits. This is even true for governments who can either exploit them or make sure they patch vulnerabilities before opposing parties find them.[41]

The examples above show the danger of cyber-attacks, but hacking tools can also be used in combination with traditional means of attack. As Russia and the United States have shown, crippling defence systems through cyber means before launching conventional weapons can be an effective strategy.[42]

3.1.2 Methods of active cyber defence

Active cyber defence is comprised of three distinct steps. Firstly, the detection of a hack. Secondly, tracing that attack to its source. Thirdly, conducting a counter hack.[43] Another way to grasp the concept of active cyber defence is to realise that active defence is 'direct action against specific threats' whereas passive cyber defence are the measures taken to strengthen

[37] Huib Modderkolk, *Het is Oorlog maar Niemand die het Ziet* (Podium 2019) 62.
[38] Tallinn Manual 2.0 (n 14) 342.
[39] Modderkolk (n 37) 186ff.
[40] Eichensehr (n 8) 475.
[41] Eichensehr (n 8) 484ff.
[42] Dinniss (n 59) 281.
[43] Kesan and Hayes (n 11) 442.

the security of a network against attacks.[44] As Kesan and Hayes explain, the counter hack has the goal of deterring an attacker and making sure that the injured party is able to execute their inherent right of self-defence[45]. Deterrence generally consists of a mix of a punitive and a thwarting element.[46] The first is to scare an attacker away or punish them and the second is to prevent an attack or its effects from succeeding. As explained in chapter 4.3, a counter hack purely aimed at punishing an attacker will likely not meet self-defence standards for failing the conditions of necessity and immediacy.

Without delving into the technical and practical intricacies of hacking back, it is good to briefly illustrate the concept with an example of counter hacking actions.

Botnets have been successfully taken down in different ways. Microsoft in 2010 took down a botnet by transferring the internet addresses that were used to control infected computers.[47] In 2011, the United States government took down another botnet by replacing the servers that were used and issuing a command to infected computers to stop using running the malware they were infected with.[48] Both of these takedown operations followed court ordered restraining orders that provided a legal basis for the takedown and the execution seemingly caused no damage. These are thus examples of active cyber defence that were in the fortunate position of being able to follow court procedures and have a clean operation. As mentioned, this paper focuses on the hacking back operations that do not have the time to follow those procedures or are at risk of causing damage to systems, making their legality less obvious. The technologies used for hacking back are often the same that are used by attackers.[49] Which means that the more extensive list in the previous paragraph can also be considered the arsenal for active cyber defence.

3.2 Who committed an attack? (Attribution)

3.2.1 Identifying an actor

One of the main difficulties in analysing, condemning or even preventing a hack is that of attribution.[50] In a system where VPNs and bots are accessible means to obfuscate the true source of a cyber-attack, the problems are numerous. Evaluating a system's passive defence is harder when it is not immediately clear what to defend against; a State-sponsored cyber team

---

[44] Dorothy E Denning, 'Framework and Principles for Active Cyber Defense' (2014) 40 Computers & Security 108, 109. Denning claims that active cyber defence and hacking back are different things, but within the context of this paper the current writer does not share that view.

[45] This is in line with the inherent right to self-defence against harm to oneself or one's property described in parapgraph 4.3.2.

[46] Kesan and Hayes (n 11) 434.

[47] Eichensehr (n 8) 480.

[48] Eichensehr (n 8) 480.

[49] Kesan and Hayes (n 11) 434.

[50] Kesan and Hayes (n 11) 438; Lahmann (n 15) 65, 80.

or a kid in their basement. Filing charges is complicated if you cannot tell who to file a claim against. Lastly, mounting a counter-hack while there is uncertainty whether the system the response is aimed at is the actual perpetrator may leave the original victim open for charges in civil or criminal lawsuits.

Generally, three standards exist in international law through which acts can be attributed to a State.[51] Firstly, acts of State organs are attributed to that State. Secondly, acts of entities or State agents that are under instruction, direction or control of the State can be attributed to that State. Finally acts that are recognised and tolerated by a State can be attributed to that State.[52]

While those standards provide a framework for possible attribution, it is not always clear how to attribute an attack. It is well established that attribution under the law on State responsibility does not entail a burden of proof or a certain evidence standard as is common in criminal cases, but follows a standard of context-dependent reasonableness.[53] The 2015 Report of the United Nations Group of Governmental Experts (UN GGE) emphasized that accusations of wrongful acts "should be substantiated", but they could not agree on the kind of proof that would meet that standard.[54] They did conclude that the fact that a hack originates from a State's governmental infrastructure or reports back to such infrastructure is no more than an indication that that State is involved with the hack and thus additional evidence would be required to attribute an attack to a State.[55] Simultaneously, the International Court of Justice has recognised the difficulty of finding evidence on another State's territory and allowed 'a more liberal recourse to inferences of fact and circumstantial evidence'.[56]

A couple of factors influence the level of evidence needed to rightfully attribute a wrongful act to a State.[57] A reasonability test has to be applied before a State attributes a wrongful act to another State and decides to respond. What is reasonable in a certain situation is always context-dependent and includes, but is not limited to, factors as: reliability, specificity, nature and directness of the information. Another factor is the severity of both the obligation breached by the attacker and of the potential response.[58] How much these factors influence the legality of a response is further discussed in the chapters detailing Countermeasures, Necessity and Self-Defence.

---

[51] Nicholas Tsagourias, 'Cyber Attacks, Self-defence and the Problem of Attribution' (2012) 17 Journal of Conflict & Security Law 229, 236.

[52] These standards are incorporated in Rules 15-17 Tallinn Manual 2.0 (n 14).

[53] Tallinn Manual 2.0 (n 14) 81, 83.

[54] William Banks, 'State Responsibility and Attribution of Cyber Intrusions after Tallinn 2.0' (2017) 95 Tex L Rev 1487, 1505.

[55] Tallinn Manual 2.0 (n 14) 91.

[56] *The Corfu Channel* (*United Kingdom v. Albania*) (Merits) [1949] ICJ Rep 4, 18.

[57] Lahmann (n 15) 70ff.

[58] Tallinn Manual 2.0 (n 14) 82.

There are some circumstances that make attribution easier. Take the case of a hacker leaving a signature in his code, a cyber-attack that is supplemented by an attack with conventional means or hacking activities that can be traced back based on language, time zone activity and other location identifiers. However, these are all indications that can be falsified or evaded by spoofing. IP spoofing is the tactic through which a hacker hides behind an IP address other than his own. Using this method one can hide their activities or deliberately implicate another party by pretending to be them.[59]

3.2.2 Qualifying the actor

Once the hack can be attributed to an individual or group, new issues arise.

If the actor is (part of) a State organ, their actions are attributed to that State regardless of whether those actions were part of the State's policy. This is only if they were acting in an official capacity. To illustrate; when an individual who is part of a government agency would hack from their government's computer network, they can be assumed to working in an official capacity. On the other hand, if a person instigates a cyber-attack and intelligence shows that this person works for a government, they mere fact that they are employed by the government is not enough to conclude that they were acting in an official capacity.[60]

Where the actor is not (part of) a State organ, but adjudicated by domestic law to exercise elements of government authority, their acts can still be attributed to a State. This relies on the actions being conducted in the capacity of 'exercising government authority'. Whether conduct should be qualified as a governmental function depends on a contextual analysis. Elements to consider are the activities in question and the characteristics of the specific State.[61] If an entity is deemed to exercise government authority, their actions can be attributed to the State if they generally fall within the scope of their duty.[62] It can be difficult to assess the exact scope of a private actor's governmental authority. Private contractors that are charged with military tasks abroad are sometimes more common than actual military personnel,[63] but how should their acts be qualified. A couple of factors make such qualifications especially difficult to establish in the cyber world. Firstly, an attack can originate from a company charged by a State to provide cyber defence. This is not definitive proof that the attack was actually conducted by that company since another actor might have used the company's network. Secondly, increasingly blurred lines between public and private entities and the privatisation of cer-

---

[59] Heather Harrison Dinniss, *Cyber Warfare and the Laws of War* (CUP 2012) 100.

[60] Tallinn Manual 2.0 (n 14) 89.

[61] Tallinn Manual 2.0 (n 14) 89.

[62] Tallinn Manual 2.0 (n 14) 90.

[63] Micah Zenko, 'The New Unknown Soldiers of Afghanistan and Iraq' (*Foreign Policy*, 29 May 2015) < https://foreignpolicy.com/2015/05/29/the-new-unknown-soldiers-of-afghanistan-and-iraq/> accessed 17 May 2020.

tain sectors makes it difficult to distinguish when acts are part of the 'government function'. Especially in the field of cybersecurity traditional roles seem to lose their meaning and private parties have increasing decision making powers on the functions they perform.[64] In extreme circumstances, where a State is not capable of fulfilling its governmental functions, a private entity may step in to fulfil that role. Subsequent actions by that entity in the exercise of that role would then also be considered attributable to the State. Regardless of whether the State has requested the private party to take on that role.[65] Eichensehr's analysis on private parties creating their own role in cybersecurity due to government absence may be an example of such a situation.[66]

Even the actions of non-State actors can be attributed to a State under certain circumstances. The non-State actor must be acting on the instructions of, or under the direction or control of a State. Acting under the instructions of a State generally can be a situation where a State has asked a private entity for help without legally empowering to fulfil certain government tasks.[67] The cyber forensics done by private companies in the United States and subsequent public attribution of hacks to other States may be an example of this.[68] 'Under the direction or control' has been translated by the International Court of Justice to a test on 'effective control'. [69] Effective control under this doctrine is established when a State determines the execution and course of an operation. Furthermore, the non-State actor's conduct must be an integral part of said operation. Part of having effective control is being able to both cause the operation to start and cease. There have been disagreements between international courts about whether a lower threshold of 'overall control' should apply to organised groups and the 'effective control' threshold should be limited to acts of individuals and unorganised groups.[70] Acts conducted outside of the instructions or effective control of the State are not attributable to that State unless they were 'incidental' to the operation.[71]

Even if a State was not involved with a private entity's actions, they can still be attributable to that State. This is the case when it acknowledges and adopts actions as its own. Acknowledgement can happen through identifying that a particular factual situation exists, through conduct or statements. Adoption then means that the State treats it as an action that was done

[64] See for an analysis on the relationship between public en private entities: Eichensehr (n 8).

[65] Tallinn Manual 2.0 (n 14) 15.

[66] Eichensehr (n 8).

[67] Tallinn Manual 2.0 (n 14) 95.

[68] Eichensehr (n 8) 489.

[69] Tallinn Manual 2.0 (n 14) 12; *Military and Paramilitary Activities in and Against Nicaragua* (*Nicaragua v United States*) (Merits Judgment) [1986] ICJ Rep 14 para 212.

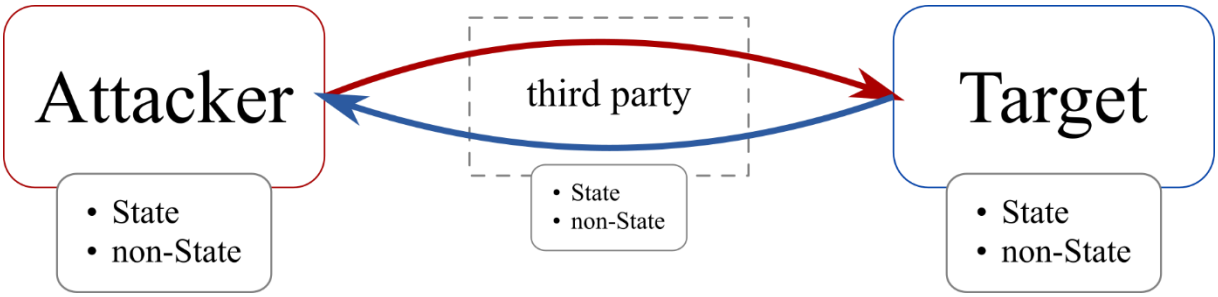
[70] Tsagourias (n 51) 238.

[71] Tallinn Manual 2.0 (n 14) 98.

on its behalf. The State decides which specific acts it adopts, doing so does not make all of that non-State actor's operations attributable to the State.[72]

As will be outlined below in Chapter 4 the attribution of cyber-attacks has consequences for assessing the legality of active cyber defence since some justifications can only be invoked by States or against States.

3.2.3 Affected parties



As the figure illustrates, it is important to realise the parties involved in a cyber-attack and possible hacking back action. While this paper cannot include an analysis on the exact position of third parties, it has to be pointed out that third parties might be affected. Take for instance malware employed as a means of active cyber defence that spreads.[73] Such a situation will create legal obligations regardless of those between the original hacker and injured party. Examining the content of such consequences and role of third parties is beyond the scope of this paper.

International law is generally aimed at regulating nation States whereas conduct of people and companies are subject to domestic legal systems.[74] In cyberspace it can be difficult to distinguish between actions of States and non-State actors. This means that to determine the legality of hacking back one cannot rely on the traditional distinction between those legal systems. In order to comprehensively answer the question of the legality of hacking back under international law, the way non-State actors are affected must be included. To answer this systematically, the applicability of circumstances precluding wrongfulness are specified per scenario of actors involved.

Non-State actors are especially caught in a peculiar situation. Without State-issued cyber protection in the form of technology and an acknowledgement of an inherent right to active defence, companies have little options to survive large cyber-attacks. Not only can they face financial penalties for a data-breach, it seems to be problematic to insure against damages caused by cyber-attacks. A recent example is the case of food producer Mondelez that got a

[72] Tallinn Manual 2.0 (n 14) 99.

[73] Lahmann (n 15) 131

[74] Shaw (n 17) 1.

$100 million claim denied by their insurer Zurich Insurance after the NotPetya attack. According to the insurer the attack could be attributed to the Russian government and thus fell under the 'war exclusion' provision of their policy.[75]

Another reason this paper involves non-State actors in the context of international law is the fading of roles and responsibilities. Kristen Eichensehr has written an extensive examination on the public-private cybersecurity system in the United States.[76] One of the things she notes is the difference between the public-private relationship in the cybersecurity field compared to privatisation of tasks in other areas. Whereas normally a government would contract out certain tasks to private companies and have checks and balances, in cybersecurity it is often the private companies who take it upon themselves to decide what function they should perform, outside of contractual frameworks.[77] In some cases this is caused by a government's failure to act.[78]

4 Precluding the wrongfulness of hacking back

The definition of hacking back used in this paper describes acts that in themselves would be violations of international law. That means that the following paragraphs presume that the hacking back breaches an international obligation. If this breach of an international obligation can be attributed to a State, it is considered an 'internationally wrongful act' under the law of State responsibility.[79] This concept is defined in the ILC Articles on State Responsibility as an act that is 'attributable to the State under international law' and 'constitutes a breach of an international obligation'.[80] In situations where the counter hacker is a State, the first criterium is easily met. Examples of customary international obligations of States in this context are to respect sovereignty of other States, a prohibition of intervention in other States and a prohibition to use force against other States.[81] Hacking back that can be attributed to a State that directly affects the confidentiality, integrity or availability of the attacker's program, computer network or system may in the act or its consequences breach one of these obligations. Whether the hacking back is aimed at another State or a non-State actor it will most likely violate the principle of sovereignty and, depending on the scale and effects of the counter hack, the relat-

[75] Justine Ferland, 'Cyber insurance – What coverage in case of an alleged act of War? Questions raised by the Mondelez v. Zurich case' (2019) 35 Computer Law & Security Review 369.

[76] Eichensehr (n 8).

[77] Eichensehr (n 8) 472.

[78] Eichensehr (n 8) 494ff.

[79] Article 3 ILC Articles on State Responsibility.

[80] International Law Commission, *Draft Articles on Responsibility of States for Internationally Wrongful Acts* November 2001 Supplement No. 10 (A/56/10).

[81] Tallinn Manual 2.0 (n 14) 85.

ed principles of non-intervention and prohibition on the use of force.[82] Since this paper has to limit its scope, the following analyses of the applicability of circumstances precluding wrongfulness assumes that the hacking back response of States is always one meeting the standard of an 'internationally wrongful act'. This allows a focus on exploring if circumstances precluding wrongfulness can be applied to cyber actions. Situations where both the attacker and the initial victim are within one State are beyond the scope of this paper since their conduct is governed by the State's domestic legal system. Situations in which a non-State actor responds to a cyber-attack are slightly different as they generally have no international obligations towards States. Consequently, they cannot commit an 'internationally wrongful act' as described in the Articles on State Responsibility, but the notion of concepts precluding the wrongfulness of actions are not limited to these Articles. In the following paragraphs it will be discussed whether there are reasons to assume the same or similar justifications could exist for non-State actors if they decide to hack back.

4.1 Countermeasures

Countermeasures are means that a State can employ to combat a wrongful act.[83] The conditions under which countermeasures are allowed in international law are stipulated in the ILC Articles on State Responsibility. According to the Tallinn Manual the 'customary international law of State responsibility undeniably extends to cyber activities'.[84] Even though the Articles of State Responsibility are widely accepted as iterations of customary international law, the precise boundaries of for example countermeasures are not unilaterally agreed on.[85]

The Tallinn Manual 2.0 describes countermeasures in Rule 20:

- A State may be entitled to take countermeasures, whether cyber in nature or not, in response to a breach of an international legal obligation that it is owed by another State.

A set of additional requirements apply to countermeasures in order to preclude their wrongfulness. They may;

---

[82] *Nicaragua* case (n 69) para 115; *Application of the Convention on the Prevention and Punishment of the Crime of Genocide* (*Bosnia and Herzegovina v. Serbia and Montenegro*) (Judgment) [2007] ICJ Rep 43, para 400.

[83] Their legality has been recognized by international courts and tribunals, see; *Nicaragua* case (n 69) para 249; *Case concerning the Gabčíkovo-Nagymaros project* (*Hungary v. Slovakia*) (Judgment) [1997] ICJ Rep 7, paras 82–83. See also *Naulilaa* (*Portugal v. Germany*) (Arbitral Award) [1928] Permanent Court of Arbitration para 1025–1026; *Responsabilité de l'Allemagne en raison des actes commis postérieurement au 31 juillet 1914 et avant que le Portugal ne participât à la guerre ('Cysne') (Portugal v. Germany)* [1930] 2 RIAA 1035, at 1052; *Air Services Agreement Case* (*France v. United States of America*) (Arbitral Award) [1978] 18 RIAA 416, paras 81–96.

[84] Tallinn Manual 2.0 (n 14) 80.

[85] For instance, prior to adoption of the Articles by the International Law Commission, the United States stated, '[w]hile we welcome the recognition that countermeasures play an important role in the regime of State responsibility, we believe that the draft articles contain unsupported restrictions on their use'. US 1997 Comments, at 1. See also US 2001 Comments, at 1.

- Only be taken by an injured State with the goal of inducing a responsible State to comply with their legal obligations (Rule 21);
- Not include actions affecting human rights, violate a peremptory norm, constitute prohibited belligerent reprisals, or infringe on diplomatic or consular inviolability (Rule 22);
- Not be disproportional (Rule 23);
- Only be taken by an injured State (Rule 24), and;
- Not violate legal obligations owed to third States (Rule 25).

4.1.1 State v. State Hacker

Countermeasures are generally considered to be actions or omissions by a State that do not meet the use of force threshold because such measures would sooner fall within the self-defence concept of the UN Charter.[86]

In a State v. State situation, the original attack can be attributed to another State. That means that the attack can be traced back to a State organ. State organs are all persons and entities that are considered organs of the State under its domestic law.[87] Apart from empowered in domestic laws, 'persons, groups or entities acting in complete dependence on the State' can also be considered State organs and thus invoke State responsibility. [88] Important to note is that it is the actor responsible for the initial wrongful act that must be a State organ under this definition to establish State responsibility. This reasoning fits in with the Group of Expert's remark that the fact that an attack can be traced to a State's network cannot constitute attribution in itself.[89] Firstly, a network cannot conduct an action and secondly, it is common for attackers to launch an attack using another's network. Views on the consequences of mistaken attribution differ. A majority of the International Group of Experts and the commentary of the Articles on State Responsibility are of the opinion that mistakenly attributing activities to a State will render the subsequent countermeasures illegitimate.[90] This opinion seems in conflict with the reasonability threshold generally employed for attribution of an attack.[91] Only time and inevitable conflict will tell what should be considered opinio iuris and State practice on reasonable, but in hindsight mistaken attribution. For the moment, the nature of cyber-attacks makes it unfeasible to use a standard of 'beyond any doubt' to be able to execute countermeasures. This view is strengthened by the fact that countermeasures seem to have an upper limit in severity.

---

[86] Tallinn Manual 2.0 (n 14) 112.

[87] Art. 4 Articles on State Responsibility; Rule 15 Tallinn Manual 2.0 (n 14).

[88] *Genocide* judgment (n 82) para 392.

[89] Tallinn Manual 2.0 (n 14) 91.

[90] Tallinn Manual 2.0 (n 14) 116; See also Commentary to Article 49 of the Articles on State Responsibility.

[91] See this paper under 3.2

Once attribution has been established, there are a number of rules the countermeasure must follow. Note, that countermeasures can only be employed in response to a breach by another State, so no anticipatory countermeasures exist.[92] The condition that a countermeasure 'may only be taken […] with the goal of inducing a responsible State to comply with their legal obligations' is mostly a formal one. Except for clear retaliatory actions, it will often be difficult to determine if the goal of a countermeasure did not have an element of punishment besides pressuring to restore the breach of obligations. The requirement to notify the other State before carrying out the countermeasure strengthens this idea by offering the breaching State the opportunity to fix the situation without becoming a victim to measures. A side-effect is that notification to a State and their response to the accusation of a breach can provide extra intelligence to strengthen the attribution.

There are some clearly defined upper limits for what a countermeasure may impact. They may not violate a peremptory norm,[93] breach the law of armed conflict[94] or infringe on consular or diplomatic inviolability[95]. A more complicated limitation is that they may not include actions that affect fundamental human rights.[96] The complication lies in the fact that there is no consensus on the scope of fundamental human rights.[97] A right to privacy can be argued to be fundamental, but especially within areas with strict definitions of privacy and personal data as the EU, nearly any cyber countermeasure would be rendered unfeasible as violating privacy. Furthermore, there is no consensus on the extra-territorial applicability of human rights. Such a question is suitable for research in itself, but beyond the scope of the current paper. It suffices to say that a reasonable approach is to respect human rights to the greatest possible extent, but that a breach of certain rights cannot be avoided in every instance of a countermeasure. In such cases a State will have to assess the possible infringement caused by the proposed measure. If there is no other measure possible and the action is proportional there should be no issue. A parallel can be seen with the margin of appreciation that the European Court of Human Rights has established for States to be able to exercise some discretion in interpreting the law by finding a balance between human rights and national interests.[98]

Another limitation is less crystalized. Both among scholars and within the International Group of Experts there is uncertainty about how the use of force and armed attack paradigms relate

---

[92] *Gabčíkovo-Nagymaros* case (n 83) para 83; See also Tallinn Manual 2.0 (n 14) 118.

[93] Article 53 Vienna Convention on the Law of Treaties jo. Art. 50(1)(d) Articles on State Responsibility.

[94] Art. 50(1)(c) Articles on State Responsibility.

[95] Articles on State Responsibility, Art. 50(2)(b). See also *United States Diplomatic and Consular Staff in Tehran* (*United States of America v. Iran*) (Judgment) [1980] ICJ Rep 3 paras 83, 86; Vienna Convention on Consular Relations, Arts. 33, 35.

[96] Rule 22 Tallinn Manual 2.0; See also Art. 50(1)(b) Articles on State Responsibility.

[97] Tallinn Manual 2.0 (n 14) 123.

[98] ECHR *Case of Handyside v The United Kingdom* (Judgment) (7 December 1976).

to each other and countermeasures.[99] Halberstam explains how there are some scholars who argue that there is a gap between the definitions of "use of force" from Article 2(4) UN Charter and "armed attack" from Article 51 UN Charter. This distinction leads to three thresholds and respective theories on what countermeasures may entail.[100] In short, they can be summarized as follows:

- Countermeasures may not constitute a use of force.
- Countermeasures may constitute a use of force, but not an armed attack.
- Countermeasures may constitute an armed attack.

Important to note is that the countermeasures according to the latter theories may only meet the respective thresholds if the initial wrongful act met that same threshold. Halberstam proposes a more practical model to deal with those different thresholds by, counterintuitively, ignoring them. He refers to this model as the "proportionate counter CNA (Computer Network Attack) rule". The rule puts more emphasis on the principle of proportionality. If the three conditions of attribution to a State, ordering that State to cease their CNA and necessity of the countermeasure to stop the CNA have been met, then a State could respond with a proportionate counter CNA. The sole focus on proportionality means that the initial attack does not have to be weighed against different thresholds. This also bypasses the discussion on whether a countermeasure may constitute a use of force or armed attack, because proportionality requires that the countermeasure can have the same severity as the initial attack.[101] Note that a countermeasure under this rule that constitutes an armed attack may not be that different to an act of self-defence. While Halberstam's wording implies that he is suggesting a new rule, this is not necessarily true. Comparing the conditions of the rule to the conditions laid down in the two benchmark cases about countermeasures, they seem to pose the same requirements.[102]

In the (more recent) Tallinn Manual, the Experts unanimously denied the possibility of countermeasures constituting an armed attack, referring to the self-defence regime, and only a minority accepted that the gap theory and a possibility for measures using force but not constituting an armed attack.[103] Frankly, this view is posed in a dismissive manner without really being substantiated. The fact that countermeasures and self-defence might overlap is not a

---

[99] Manny Halberstam, 'Hacking Back: Reevaluating the Legality of Retaliatory Cyberattacks' (2013) 46 Geo Wash Intl L Rev 199ff; Stephen Petkis, 'Rethinking Proportionality in the Cyber Context' (2016) 47 Georgia J Intl L 1431; Tallinn Manual 2.0 (n 14) 125.

[100] Halberstam (n 99) 212ff

[101] Halberstam (n 99) 224ff.

[102] cf *Gabčíkovo-Nagymaros* case (n 83) para 83ff: 1. Wrongful act by another State 2. Call to State to cease wrongful conduct, notify of intention to employ countermeasures and offer negotiation 3. Countermeasure must be proportional 4. Countermeasures must have the goal to end the violations of international law by the first State; cf *Naulilaa* arbitration (n 83) 1025: 1. Preceded by a violation of international law by another State 2. Preceded by unfulfilled demand for cessation and redress 3. Proportionate to the original injury.

[103] Tallinn Manual 2.0 (n 14) 125.

reason to dismiss their application to certain circumstances. Furthermore, limiting countermeasures this way is not obvious based on the conditions laid down in case law. Finally, this limitation is counterintuitive to the idea that proportionality should be applied. The elements of a proportionality assessment in the context of countermeasures are after all 'the extent of the harm, significance of the breached rule, rights of the injured and responsible State and the need to effectively cause the responsible State to comply with its obligations.'[104] This is different from the proportionality assessment used for self-defence. Countermeasures take the harm suffered into account, were self-defence takes the threat as baseline for proportionality.[105] This should not be read as a justification for an 'eye for an eye' approach, but it would also be illogical to exclude forcible measures based on these criteria. Only time will likely provide more clarity on the exact thresholds and their application, as the Experts recognise 'What this approach might mean in the cyber context will remain an open question until uncertainty as to the use of force and armed attack thresholds is resolved.'[106]

The final Rule on countermeasures might be the most problematic. A countermeasure may affect third parties, but it may never violate a legal obligation owed to a third State or other party.[107] In cyberspace, networks are so connected globally that excluding effects and possibly breaches towards third parties seems far-fetched. Testament to this are the Stuxnet and NotPetya viruses that wreaked havoc far beyond their intended targets.[108]

4.1.2 Non-State v. Hacker

The concept of countermeasures is one that exclusively falls within the limits of public international law. After all, it is laid down in the ILC Articles on *State* Responsibility as a circumstance precluding wrongful acts by *States*. The only way a non-State actor's conduct could be classified as a countermeasure is if it can be attributed to a State because the State has empowered a particular actor to take the countermeasure.[109] When that is the case, the situation fits within the paradigm of a State conducting a countermeasure discussed in paragraphs 4.1.1 and 4.1.3. The fact that countermeasures are not a concept applicable to non-State actors does not necessarily mean that their options for hacking back are more limited. This is further discussed in the paragraph on Self-defence.

---

[104] Tallinn Manual 2.0 (n 14) 128.

[105] Dinniss (n 59) 107.

[106] Tallinn Manual 2.0 (n 14) 126.

[107] Rule 25 Tallinn Manual (n 14).

[108] Andy Greenberg, 'The Untold Story of NotPetya, the Most Devastating Cyberattack in History' (Wired, 22 August 2018) <https://www.wired.com/story/notpetya-cyberattack-ukraine-russia-code-crashed-the-world/> accessed 17 May 2020; Kim Zetter, 'An Unprecedented Look at Stuxnet, the World's First Digital Weapon' (*Wired*, 11 March 2014) <https://www.wired.com/2014/11/countdown-to-zero-day-stuxnet/> accessed on 1 June 2020.

[109] Rules 15, 17, 33 Tallinn Manual 2.0 (n 14).

Persons that would normally be considered as non-State actors can nevertheless legally conduct countermeasures if they do so on behalf of the State, for example if a private company is contracted by a State to perform certain, normally governmental, functions. In such a case the non-State actor's actions would be attributable to a State and in fact be situation as discussed in paragraphs 4.1.1 and 4.1.3.[110]

An exception is made for international organizations as they are often entities with legal personality under international law. The International Law Commission has drafted specific Articles on the Responsibility of International Organizations, but has stated that they partly represent a 'progressive development of the law' and not existing customary law. Partly for this reason, the International Group of Experts who wrote the Tallinn Manual 2.0 expresses the view that international organizations may take and be subject to countermeasures, but the conditions are not the same as for States, nor clear and uncontested.[111]

4.1.3 State v. non-State Hacker

In case an attack can be retraced to a non-State actor, the role of that actor must be assessed. If the non-State actor can be considered a State organ, then the action can be attributed to the State it is an organ of. If it is not a State organ, its action can still be attributed to a State if that State instructed, directed or controlled the entities actions or if the State acknowledged and adopted the actions as its own.[112] The necessary scope of direction or control exercised by a State has been clarified by the ICJ as being "effective control".[113] In short, effective control is established if a State determines the course of a specific operation and has the ability to both cause and stop activities. Even if the hack by a non-State actor cannot be attributed to a State under these rules, there might be acts that are related to the hack that can induce State responsibility and thus grounds for countermeasures. Think for example of a State supplying malware to an insurgent group in another State or a State that violates its due diligence obligations by not halting an ongoing cyber-attack by entities on its territory.[114]

A loophole might exist here. Consider a situation where a hack by a non-State attacker cannot be attributed to the State they are in. One could argue that part of a State's due diligence obligation is to make sure that it knows the cyber activity on its digital infrastructure. Breach of due diligence can constitute an international wrongful act and thus may grant a right for the victim State to adopt countermeasures. The proportionality of such a countermeasure will

[110] Tallinn Manual 2.0 (n 14) 90ff.

[111] Tallinn Manual 2.0 (n 14) 166.

[112] Rule 17 Tallinn Manual 2.0 (n 14).

[113] ICJ *Nicaragua* (n 69) para. 115; *Genocide* judgment (n 82) para. 400.

[114] Tallinn Manual 2.0 (n 14) 100.

however be judged against the severity of the State's omission to control their network and not against the severity of the cyber-attack.[115]

Due to specificity of the legal regime of countermeasures and its connection to international obligations by States owed to each other, countermeasures are not an appropriate instrument to a non-attributable attack of a non-State actor. This does not necessarily mean that there are no options for a State to respond to non-State actors in certain circumstances. Firstly, if the attack meets the use of force threshold and requires an immediate response, the State could legally hack back under the regimes of necessity or self-defence. Secondly, if the attack meets a lower threshold, the proper action would be to notify the State it originates from and if that State then does not intervene or give consent[116] to the injured State to stop the activities it would be in breach of its due diligence obligation. One could even argue that States have an obligation to provide consent to allow an injured State to stop hacks coming from their territory. After all, if a State exercises its due diligence, but is unable to put the nefarious activities to a halt, not allowing the victim State to take measures can be interpreted as unwillingness to stop the activities. In cases where a State acknowledges and adopts operations as its own, those operations can also be attributed to that State.[117] While this standard should be narrowly applied, a case could be made that willingly allowing cyber activities by a non-State actor from your territory harming another State, and not taking measures yourself nor providing the injured State with options to put a stop to them, may make the conduct attributable to you. As a consequence, countermeasures are back in the picture for the injured State.

4.2 Necessity

The Tallinn Manual 2.0 describes necessity in Rule 26:

- A State may act pursuant to the plea of necessity in response to acts that present a *grave and imminent peril*, whether cyber in nature or not, to an *essential interest* when doing so is the sole means of safeguarding it.

While in a first impression Rule 26 seems to have a lower threshold for a plea of necessity than the original Article 25 of the Articles on State Responsibility, the commentary mentions two additional conditions that it considers integral to the Rule, reaffirming the same threshold as Article 25 held.

It cannot be emphasized enough that a plea of necessity has a high threshold to be accepted. The International Law Commission as well as the Group of Experts and the International Court of Justice hold the view that it may only be invoked in exceptional cases.[118]

---

[115] Tallinn Manual 2.0 (n 14) 100.

[116] Articles on State Responsibility, Art. 20. See also *Case Concerning Armed Activities on the Territory of the Congo* (*Democratic Republic of the Congo v. Uganda*) (Judgment) [2005] ICJ Rep 168, paras 45–46.

[117] Rule 17(b) Tallinn Manual 2.0 (n 14) See also *Tehran Hostages* case (n 95) para. 74.

[118] Tallinn Manual 2.0 (n 14) 135.

> Countermeasures must also be distinguished from actions taken based on a plea of necessity (Rule 26). The former differs from the plea [of necessity] in two main ways. First, there must be an underlying internationally wrongful act to justify countermeasures, whereas necessity has no such condition precedent. In other words, the act that precipitates a countermeasure must be attributable to a State, while acts pursuant to the plea of necessity may be taken in response to the cyber operations of non-State actors (or even when the author of the act is unidentified). Second, mere international wrongfulness suffices to trigger the right to take countermeasures; action based on necessity is only permissible when the situation amounts to a grave and imminent peril to an essential interest of the acting State.[119]

The Group of Experts clearly emphasizes that non-State actors cannot invoke the responses that are available under the law on State responsibility.[120] Consequently, this chapter focuses only on the possibility of a State for a plea of necessity. Furthermore, the argued right for non-State actors to self-defence in paragraph 4.3.2 covers roughly similar justifications for active cyber defence.

4.2.1 State v. Hacker

Because a plea of necessity under the law of State responsibility is focused on the State faced with the grave and imminent peril to its interests, it does not discriminate at who the 'necessary' response is directed. Therefore, this chapter has no distinguished paragraphs for State v. State and State v. non-State pleas.

Firstly, there is no clear definition of what constitutes an essential interest in the context of necessity.[121] The Group of Experts has observed a tendency of States to classify certain sectors as 'critical infrastructure' and sees this as an indication, but not determinative to qualify something as essential.[122] Usage of the term 'critical infrastructure' is increasingly seen in legal instruments with a focus on cybersecurity.[123] Since the Group of Experts has undoubtedly done a thorough examination of customary international law, we may assume that there is no opinion iuris or State practice to interpret what essential is. The trend to use the label of 'critical infrastructure' seems to provide some clues. As Sarah Heathcote proposed to look at

---

[119] Tallinn Manual 2.0 (n 14) 114.

[120] Tallinn Manual 2.0 (n 14) 175.

[121] Tallinn Manual 2.0 (n 14) 135.

[122] Tallinn Manual 2.0 (n 14) 135.

[123] Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security, 'Developments in the field of information and telecommunications in the context of international security' (2015) UN GA A/70/174; Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology cybersecurity certification and repealing Regulation (EU) No 526/2013 (Cybersecurity Act).

the international community as a whole to see if there is some consensus on what is to be considered consensual.[124] Some States have started initiatives to create a common concept of critical infrastructure.[125] One initiative concerned a limited group of like States (Australia, Canada, New-Zealand, the United Kingdom and the United States). [126] Their concepts will likely not fully align with States like China and Russia, but it is a reasonable assumption that the importance of sectors like energy, healthcare, transportation systems and water are widely recognised. As always when new situations arise under law, there are uncertainties about interpretations, thresholds and limitations. A clearer picture on the legal importance of a 'critical infrastructure' designation will emerge once there have been pleas for necessity linked to critical infrastructure or cyber incidents. The subsequent response by the international community and judicial procedures will then further determine its status under international law. Nevertheless, 'critical infrastructure' is not the only possible indicator of essential interests. Other factors linked to a State's sovereignty such as territorial integrity, public security and political independence will likely also be considered essential interests under the right circumstances.[127]

Secondly, the threat to an essential interest must be 'grave'. In the Tallinn Manual this is described as 'It involves interfering with an interest in a fundamental way, like destroying the interest or rendering it largely dysfunctional.'.[128] This is most likely the case where a State's security, economy, public health, safety or environment face severe negative impact. Because of the exceptionality of the circumstances, and as the term 'necessity' implies, possibly impacted rights of non-responsible States or non-State actors are subsidiary to the measures needed to diminish the threat. Nevertheless, according to the Group of Experts, a plea of necessity is no justification for actions that seriously impair another State's essential interests.[129] However, this view of balancing interests is not supported by State practice, nor uncontested in literature[130] and case law[131]. It seems to stem from domestic criminal law systems, but the analogous inclusion in the concept of necessity in international law is problematic.[132]

---

[124] Sarah Heathcote, 'Circumstances Precluding Wrongfulness in the ILC Articles on State Responsibility: Necessity', in James Crawford, Alain Pellet & Simon Olleson (eds), *The Law of International Responsibility* (2010) 491, 497.

[125] Christian Schaller, 'Beyond Self-Defence and Countermeasures: A Critical Assessment of the Tallinn Manual's Concept of Necessity' (2017) 95 Tex L Rev 1619, 1631.

[126] See at <https://perma.cc/HWU3-342S> accessed on 11 May 2020.

[127] Schaller (n 125) 1631.

[128] Tallinn Manual 2.0 (n 14) 136.

[129] Art. 25(1)(b) Articles on State Responsibility; Tallinn Manual 2.0 (n 14) 137.

[130] Schaller (n 125) 1629.

[131] *Rainbow Warrior* (*New Zealand v France*) (Arbitration Tribunal) (1990) 20 RIAA 215, 254.

[132] Robert Sloane, 'On the Use and Abuse of Necessity in the Law of State Responsibility' (2012) 106 AJIL 447, 458-459, 478-481.

Thirdly, the peril must be 'imminent'. This implies that preventive action may be taken on a plea of necessity. The standard used to determine the imminence is one of reason; '[A] State may only act when a reasonable State in the same or similar circumstances would act prior to the advent of the harm that will be caused by the cyber operation.'[133] 'Imminent' in the context of necessity goes beyond anticipating a threat and includes ongoing or even finished operations, as long as harm is still being done to a State's essential interests.[134] Furthermore, peril can still be considered even if the manifestation of the harm lies a long time ahead. The important consideration is whether the harm is certain and inevitable.[135] Consequently, a hunch that something awful will happen is not enough. A threat must be objectively established and the decision that hacking back is required to prevent harm must be based on evidence 'reasonably available at the time'.[136]

4.3 Self-defence

Sources: Articles 2(1), UN Charter

International law includes a prohibition on the use of force, recorded in Article 2(4) of the UN Charter for Member States, but this is recognised generally as a rule of customary international law.[137] One of the only exceptions to this prohibition is the use of force by a State to defend itself against an armed attack. Since the legal framework of the UN Charter addresses only the rights and duties of States, it will be used to assess the legality of hacking back activities for States only. Unlike concepts as countermeasures and, to a lesser degree, necessity, self-defence is a term that can be observed in legal systems everywhere. Consequently, this chapter includes a more elaborate examination of the possibility for non-State actors to legally hack back. After all, general principles of law are considered a primary source of international law.[138] While the extent of general principles of law are not clearly defined, research has been done on the merits of qualifying self-defence as such a principle.

4.3.1 State v. State Hacker

The Tallinn Manual 2.0 describes the right to self-defence in Rule 71:

- A State that is the target of a cyber operation that rises to the level of an *armed attack* may exercise its inherent right of self-defence. Whether a cyber operation constitutes an armed attack depends on its *scale and effects*.

The following additional conditions apply to a State contemplating a self-defence action;

---

[133] Tallinn Manual 2.0 (n 14) 138.

[134] Tallinn Manual 2.0 (n 14) 139.

[135] *Gabčíkovo-Nagymaros* case (n 83) para 54.

[136] Art. 25 Articles on State Responsibility paras 15, 16 of the commentary.

[137] Tallinn Manual 2.0 (n 14) 329; ICJ *Nicaragua* (n 69) paras 188-190.

[138] Art. 38(1) Statute of the International Court of Justice (1945); Shaw (n 17) 70.

- Exercising the right of self-defence must be necessary and proportionate (Rule 72);
- The right to self-defence arises when an attack is occurring or imminent and the defensive action must also meet a standard of immediacy in order to distinguish defence from retaliation (Rule 73).

A response by a State that qualifies as legitimate self-defence precludes the wrongfulness of an act that otherwise would constitute a violation of the use of force and sovereignty principles.[139] The Group of Experts bases this formulation of self-defence for States on a combination of Article 51 of the UN Charter, Article 21 Articles on State Responsibility and its status of customary international law.

The first condition is that a State must be victim of a cyber-attack that meets the threshold of an 'armed attack'. In the Tallinn Manual the theory that an armed attack requires the use of weapons is acknowledged but disregarded.[140] One could argue that even if that theory had merit, there is no reason to assume that computers or computer programs used to hack would not fall within the common definition of a weapon.[141] This is in line with the International Court of Justice's position that 'force' under the UN Charter is not limited to specific weapons.[142] Nevertheless, the consequences and surrounding circumstances determine whether the threshold is met, not the instruments used for a cyber-attack.[143] One requirement of an armed attack is that it must have a trans-border or international element. A response to an attack that is solely conducted from within a State against that State will be governed by the State's domestic law and potential human rights obligations. As long as the hack that meets the threshold of armed attack can be attributed to a State it does not matter whether it was executed by government entities, individuals, groups or organisations.[144] What it means for the right to self-defence if an attack can be attributed to a non-State actor, but not a State is slightly controversial and highlighted below in paragraph 4.3.3.

The questions on when a certain action exactly qualifies as a use of force or an armed attack under the UN Charter have not been definitively answered. Amongst other things this has led to discussion on whether there is a gap between the two concepts. The Tallinn Manual and the International Court of Justice claim that there is and that an armed attack has a higher threshold than a use of force.[145] While this view is generally considered to be current international law, there have been dissenting opinions pointing out that a gap between the thresholds has potential negative consequences. If countermeasures may not reach the use of force and a

---

[139] Tallinn Manual 2.0 (n 14) 107.

[140] Tallinn Manual 2.0 (n 14) 340.

[141] See Merriam Webster Online Dictionary's definition of 'weapon': 1. Something used to injure, defeat or destroy <https://www.merriam-webster.com/dictionary/weapon> accessed on 13 May 2020.

[142] *Legality of the Use or Threat of Nuclear Weapons* (Advisory Opinion) (1996) ICJ Rep 226 para 39.

[143] Tallinn Manual 2.0 (n 14) 328.

[144] See 3.2 for the rules on attribution.

[145] Tallinn Manual 2.0 (n 14) 341; ICJ *Nicaragua* (n 69) para 191.

right to self-defence only exists after an armed attack, there would be no proportional answer for an injured State to actions falling in that gap.[146]

The accepted test to judge an action against the threshold of an armed attack is to assess its scale and effects. These criteria come from the *Nicaragua* case and have been reaffirmed in the *Oil Platforms* case.[147] What was missing from the judgments were specific parameters that could be used to assess incidents. The only indication is that the scale and effects need to be 'grave'. Even though necessity and self-defence are different constructs, perhaps the classification of grave used for a plea of necessity can offer further guidance. Even when accepting the gap theory exists, Dinniss argues that because of the undesirable possible effects, that gap cannot be too great.[148] What is at least accepted is that not every form of pressure on a State's territorial integrity or political independence are considered force under the UN Charter's prohibition.[149] Economic measures are an example of pressure that can be applied by States, but that will never amount to a use of force, regardless of the effects. Following from this, the current notion is that injury, death, damage and/or destruction are necessary effects if something is to be considered an armed attack.[150] The Stuxnet virus that caused the destruction of centrifuges in an Iranian nuclear facility is widely considered to constitute a use of force, precisely because the cyber-attack resulted in physical damage.[151] Whether the damage was grave enough to meet the threshold of an armed attack is debatable. It follows from the Tallinn Manual and a logical interpretation of the case law that multiple smaller attacks that would not meet the threshold could cumulatively be considered an armed attack. This depends on whether the smaller attacks come from the same party and together are of such a scale and effects to meet the threshold.[152]

The requirements of necessity and proportionality in a self-defence context are not included in the UN Charter. They have been emphasized by the International Court of Justice in cases judging on self-defence.[153] Acts of self-defence must be necessary to repel an imminent attack or defeat an ongoing one and not for the furtherment of other goals. In the case of a defensive action amounting to a use of force, it is not a requirement that other means to stop the attack are unavailable, it is merely required that they are insufficient to stop the attack.[154] However,

---

[146] *Case Concerning Oil Platforms* (*Iran v. United States*) (Judgment) (2003) ICJ Separate Opinion of Judge Simma para 12.

[147] ICJ *Nicaragua* (n 69) para 195; ICJ *Oil Platforms* (n 146) para 64.

[148] Dinniss (n 59) 79.

[149] Tallinn Manual 2.0 (n 14) 331.

[150] Tallinn Manual 2.0 (n 14) 342.

[151] Dinniss (n 59) 112.

[152] Tallinn Manual 2.0 (n 14) 342.

[153] *Nicaragua* case (n 69) paras 176 and 194; *Nuclear Weapons* opinion (n 142) para. 41; *Oil Platforms* case (n 146) para. 74; *Armed Activities* case (n 116) para. 147.

[154] Tallinn Manual 2.0 (n 14) 348.

this also means that if passive cyber defence is enough to repel an attack, active cyber defence would not be deemed necessary. The assessment on the necessity of self-defence is to be made by the victim. Its considerations must be whether self-defence is 'reasonable in the attendant circumstances'.[155]

Proportionality in the context of self-defence is based on the amount of force that is necessary to successfully thwart the attack. Factors to be considered for this determination are the scale, scope, duration and intensity of the defensive action. The action taken in self-defence does not have to use the same means as the original attack. [156] This means that conventional weapons can be used in response to a cyber-attack meeting the armed force threshold and the other way around.[157]

The final consideration is that of imminence and immediacy. After all, the right to self-defence can be invoked 'if an armed attack occurs' according to Article 51 UN Charter. The Group of Experts defines this 'occurring' as a situation that has caused or is in the process of causing damage or injury.[158] Beyond that, anticipatory self-defence is widely recognised though its boundaries are heavily debated. The standard used to determine imminence comes from correspondence between the United States and the United Kingdom concerning the *Caroline* incident. It was claimed that self-defence is allowed when its necessity is 'instant, overwhelming, leaving no choice of means and no moment for deliberation'. [159] Take for example the placement of a backdoor in a system. The creation of a backdoor is often the first step to ensure access for nefarious activities at a later time. Since the nature and scope of such actions often cannot be determined just by seeing that a backdoor exists, noticing such a backdoor will not meet the 'imminent' criteria.[160] However, it may give right to implement countermeasures if the backdoor can be qualified as a wrongful act by another State. The majority of the Experts working on the Tallinn Manual 2.0 held the opinion that anticipatory self-defence adopted a 'last feasible window of opportunity' standard. This standard dismisses a pure temporal application of imminence. Its focus lies in the last opportunity to mount a successful defensive operation, which under certain circumstances can be long before the attack actually occurs. Just as with necessity, the victim State must make a reasonable assessment whether a last window of opportunity before an attack exists. Mere preparations or capabilities of another State are not enough to satisfy this test. There should be at least implied or explicit indications that the other State will launch an attack.

---

[155] Tallinn Manual 2.0 (n 14) 349.

[156] Tallinn Manual 2.0 (n 14) 349.

[157] Dinniss (n 59) 104.

[158] Tallinn Manual 2.0 (n 14) 350.

[159] Letter from Daniel Webster to Lord Ashburton (6 August 1842), reprinted in 2 Intl L Dig 412 (John Bassett Moore ed, 1906).

[160] Dinniss (n 59) 90.

Where imminence has its focus on the (anticipated) attack, immediacy relates to the other side of the self-defence. Factors as the time between the attack and the response, including preparation time for a defensive operation and the time necessary to identify the attacker will distinguish legitimate self-defence from illegal retaliation. If that period of immediacy has passed, self-defence would no longer be allowed.[161]

4.3.2 Non-State v. Hacker

'Non-State actors are not entitled to engage in the responses that States may conduct under the law of State responsibility when facing hostile cyber operations by or attributable to other States.'[162]

As mentioned at the outset of this chapter, this paragraph will argue that international law also recognises a right to self-defence for non-State actors. Following from the previous paragraphs that non-State actors' self-defence would not fall within the scope of Art 51 UN Charter, this argument is based on another source. The source referred to is as a 'general principle of law recognised by civilized nations'.[163] Other primary sources seem to hint at a right to self-defence, but have not explicitly recognised it. In the context of treaties, it is mentioned in Article 31(1)(3) of the Rome Statute of the International Criminal Court and Article 2(2)(a) of the European Convention on Human Rights. To be considered a norm of customary law there has to be State practice and an opinion iuris. Whereas State practice might be established through the widespread recognition in domestic legal systems and cases, opinion iuris would be harder to prove.[164] Opinio iuris requires the conviction of States that the State conduct is required by international law. Hessbruege argues that because the concept of self-defence predates international law, it would be hard to establish that the existence of it in their domestic law comes from the belief that it is required by international law.[165] The inclusion of general principles of law in Statute of the ICJ is precisely to fill any gaps between treaty and customary law.[166]

A principle can be recognised as a general principle of law if it is common to, at least the major, domestic legal systems of the world. The execution and wording do not have to be the same in all legal systems, as long as a same basic principle can be distinguished.[167] If such a

---

[161] Tallinn Manual 2.0 (n 14) 353.

[162] Tallinn Manual 2.0 (n 14) 175.

[163] Art. 38(1) Statute of the International Court of Justice (1945)

[164] *Certain German Interests in Polish Upper Silesia* (*Germany v. Poland*), 1926 PCIJ (ser A) No. 7 (25 May 1926) at 19; Wolff Heintschel von Heinegg, 'Die weiteren Quellen des Völkerrechts' in Knut Ipsen (ed), *Völkerrecht* (6 edn, CH Beck 2014) 470–510 at 478.

[165] Jan Arno Hessbruegge, 'The Right to Personal Self-Defense as a General Principle of Law' in: *Human Rights and Personal Self-Defense in International Law* (OUP 2017) 18.

[166] James Crawford, *Brownlie's Principles of Public International La*w (8th edn, OUP 2012) 25.

[167] Hessbruegge (n 165) 23.

principle can be distinguished, it will not automatically also constitute international law. The principle has to be suitable for use in an international environment.[168] Application of the principle of a right to self-defence in the context of hacking back is especially suitable due to the international nature of cyber-attacks. Hessbruege extensively demonstrates how self-defence meets both the requirements of natural and positivist schools of thought and how the principle is suitable in international law. The arguments he makes are briefly outlined below.

The naturalist justification is built on a comparative argument and a rationalist argument. Firstly, a right to personal self-defence has been recognised by different cultural, religious and philosophical traditions across centuries.[169] Demosthenes, Cicero, Thomas Aquinas, Grotius, de Vattel, Hobbes and Locke are a number of the most influential legal scholars in Western history, all of who recognised a right to self-defence.[170] A well written moral justification for why only trusting in officers of the law is not enough is given by Bentham:

> This right of [self-]defense is absolutely necessary. The vigilance of magistrates can never make up for the vigilance of each individual on his own behalf. The fear of the law can never restrain bad men so effectually as the fear of the sum total of individual resistance. Take away this right, and you become, in so doing, the accomplice of all bad men.[171]

The view that one must be able to defend oneself is not limit to Western culture; inter alia Islam[172], Taoism[173], Judaism[174] and Hinduism[175] recognise this right.

The rationalist argument comes in the form of multiple theories that will not be examined in turn here.[176] Combined they make a compelling argument that can be described as follows. A right to self-defence serves to protect autonomy, individual rights and the socio-legal order. Furthermore, acts of self-defence offer an exception to the normal prohibition on violence because they arise in a situation where someone is forced to choose by an aggressor between his own rights and that of the aggressor. The outcome of a situation where self-defence is necessary would lose its morality if the initial victim would be culpable for the aggressor's actions or if the resulting harm would be disproportionally greater than what was prevented.[177]

---

[168] Hessbruegge (n 165) 24.

[169] Hessbruegge (n 165) 29.

[170] Hessbruegge (n 165) 30ff.

[171] Jeremy Bentham, *Theory of Legislation Vol. II* (Richard Hildreth ed, Weeks, Jordan & Company: Boston 1840) 38.

[172] Rudolph Peters, *Crime and Punishment in Islamic Law* (CUP 2005) 25.

[173] George W. Wolfe, *The Spiritual Power of Nonviolence* (JOMAR Press 2012) 104–105.

[174] Babylonian Talmud Sanhedrin 72a; Boaz Sangero, *Self-Defence in Criminal Law* (Hart 2006) 31-33.

[175] *The Books of Manu* (originally ca 500 BCE, Georg Bühler ed, Library of Alexandria 1964) 118.

[176] Hessbruegge (n 165) 48ff; Respectively they can be described as the punishment, self-preservation, intent, individual autonomy and freedom, rights forfeiture, right against the State, balance of harms, defence of the socio-legal order and forced choice theories.

[177] Hessbruegge (n 165) 58.

One may start to see how the theories will culminate in the parameters with which can be decided when self-defence is justifiable. Those will be explicitly outlined after the positivist basis to recognise the right.

The positivist justification of a right to self-defence as a general principle of law hinges on the question whether the principle can be derived from domestic law, that is whether it is common in different domestic legal systems.[178] Different, continent-wide, comparative law studies have not been able to find a single domestic legal order that did not have a provision on self-defence.[179] This does not mean that every legal order has the same implementation of self-defence. In order to establish whether there are enough commonalities to establish a general principle, both the legal consequences and requirements in the different national iterations are examined.[180]

It is important to realise that self-defence is a right instead of a liberty. A right can be considered as a justification, taking away the illegality of otherwise illegal actions. A liberty would be considered an excuse, where conduct is still illegal but under circumstances does not invoke punishment.[181] This distinction prevents escalation of violence, at least legally. Since self-defence as a right connected to situation of defending oneself from initial aggression does not give the aggressor a right to invoke self-defence in response. A liberty or excuse on the other hand create no obligations for others. Meaning that it could escalate by allowing the aggressor to respond to self-defence with more violence. In practice an aggressor would not wilfully respects another's right and let the defensive force come over him or her. Legally the aggressor has to respect the victim's right and actions. In other words; there is no accepted plea of self-defence by an aggressor because he or she was confronted with defensive violence.[182]

The right to self-defence is universally based around the same three parameters. Those are immediacy, necessity and proportionality.[183]

Applying this construct to the sphere of cyber-attacks works in the following way. Even though cyber-attacks to a hospital could cost lives, generally cyber aggression will result in damage or loss of property. Besides from life, physical integrity and physical liberty, property

---

[178] Hessbruegge (n 165) 58.

[179] Jean Pradel, *Droit Pénal Comparé* (3rd ed, Dalloz 2008) 138; Eduardo José Pintore, *Das Präventive Selbstverteidigungsrecht im Lichte Allgemeiner Rechtsgrundsätze* (Peter Lang 2011) 133ff; Stanley Yeo, 'Anglo-African Perspectives on Self-Defense' (2009) 17 African JICL 118ff.

[180] Hessbruegge (n 165) 59.

[181] Hessbruegge (n 165) 60ff.

[182] See Weizsaecker et al ["The Ministries Case"], *Trials of War Criminals before the Nuremberg Military Tribunals Vol XIV* (1946-1949) 308, 329; *Wharton's Criminal Law* (vol I, 12th edn, Lawyer's Cooperative Publishing Company 1932) 180.

[183] Hessbruegge (n 165) 30.

is another interest that may be defended under all major legal systems.[184] Firstly, a hack in a computer system constitutes an act of aggression. With the problems surrounding attribution it will not always be clear if the hack was done by a culpable aggressor. It could be a virus that spread beyond its intended goals or objectives. The consequences of culpability will return later in this argument. Secondly, both the threat to a non-State actor's interest and the response must be immediate. This is relevant because defence against an attack that is not certain to happen will just be an act of aggression itself. Similarly, hacking back after an attack is over will no longer constitute defence but rather revenge or vigilantism and infringe on the State monopoly to enforce law. Thirdly, the act of self-defence must be necessary. If it is possible to negate the threat or damage of a hack in a way that is not damaging to the aggressor or the victim, that option must be taken. After all, if avoiding damage is possible but one still chooses to engage in further aggression the idea of self-defence would lose it's meaning of defence and shift more to an active excuse to use force. Finally, the hacking back must be proportional. Take an example where an e-mail server is compromised. If it turns out that there are no internal ways to stop the attack, a right to self-defence might be invoked. Nevertheless, it would lose its meaning and justification if the counter hack would take out an entire Internet Service Provider's operation. Pufendorf provides an accurate description on application of the necessity and proportionality criteria that can form an adequate parallel for an attack by a hacker:

> We must first try the more harmless remedies; for instance, we must endeavour to keep out the invader by cutting off his access to us; to withdraw into strong places; and to admonish him to desist from his outrageous fury. And it is also the duty of a prudent man to put up a slight wrong, if it may conveniently be done, and to remit somewhat of his right, … especially if that thing or concern of ours upon which the attempt is made, be such as may easily be amends for or repaired.[185]

Proportionality in the context of self-defence does not mean that the defence is equal or less than the attack. It just cannot be 'grossly disproportional'.[186]

The suitable of the right to self-defence as a general principle in international law is easily proven through practice. It is recognised in international humanitarian, international criminal law, the law of the sea, the law of diplomatic relations and applied in case law.[187] Especially the law of the sea is interesting in the comparison to hacking back since it recognises a right

---

[184] Hessbruegge (n 165) 63; Sangero (n 174) 123; Pradel (n 179) 138.

[185] Samuel von Pufendorf, *The Whole Duty of Man According to the Law of Nature* (Andrew Tooke tr, Ian Hunter/David Saunders eds, Liberty Fund 2003) 83.

[186] Hessbruegge (n 165) 66.

[187] *Naulilaa* arbitration (n 83) para 1025; *Prosecutor v. Tadic* (Judgment) ICTY-94-1 (7 May 1997) para 640; *Prosecutor v. Bagosora* (Judgment) ICTR-98-41 (18 December 2008) para 2238; *Prosecutor v. Abu Garda* (Decision on the Confirmation of Charges) ICC-02/05-02/09 (8 February 2010) para 83.

to self-defence for private vessels against piracy.[188] Even in situations involving diplomatic representatives, the right of self-defence has been recognised and considered to overrule inviolability of diplomatic personnel.[189] This provides a reference in cases where active cyber defence is considered, but the suspected aggressor is a State official.

Would active cyber defence be used to protect a non-State actor's interests at the expense of another party's? Almost certainly. However, this is not problematic when the factors above are taken into account. The right only exists when the hack by an aggressor forces the victim to make a choice between its own interests and that of the aggressor. Under such circumstances it would be irrational, immoral and contrary to the principles of law to not allow defensive actions. As 'there would be no condition of happiness for the world, nay, its condition would be one of utter misery, if oppressors and robbers and plunderers could with impunity commit their crimes and oppress the good and innocent, and these latter could not in turn retaliate on them.'[190] A fear that recognising a right to hack back would create an online wild west and vigilantism mirrors arguments against recognising traditional self-defence as a right. As Hessbruegge points out, that argument fails to consider that self-defence can deter attacks and reaffirms the validity of the social legal order. Furthermore, it is highly unlikely that if people were no longer allowed to defend themselves or others, unlawful violence by aggressors would decline.[191]

4.3.3 State v. non-State Hacker

The extent to which a State can invoke self-defence against a non-State attack will not reiterate all the criteria outlined above. Nevertheless, it deserves some specific attention.

The use of force paradigm of Article 2(4) UN Charter clearly addresses States, whereas the self-defence exception from Article 51 does not. Article 51 solely relies on an attack that meets the threshold of an 'armed attack'. However, as argued in the previous paragraph and supported by other authors, Article 51 does not establish a right but outlines an inherent one.[192] Complemented by the ICJ's admission that the UN Charter does not reflect the whole body of law governing the use of force indicates room for interpretations not limited to the scope of the UN Charter.[193]

Specific problems arise when it comes to proportionality. Acts that were claimed out of self-defence against non-State actors that targeted the infrastructure of the State were considered

---

[188] Alice Priddy, 'The Use of Weapons in Counterpiracy Operations' in Stuart Casey-Maslen ed, *Weapons under International Human Rights Law* (CUP 2014) 134.

[189] *United States v. Brenner* (Case No 14,568) 24 Fed Cas 1084 (1830) 1085.

[190] Francisco de Vitoria, *De Indis et de Ivre Belli Relectiones* (John Pawley Bate tr, Ernest Nys ed, Carnegie Institution 1917) 163ff.

[191] Hessbruegge (n 165) 63.

[192] See Dinniss (n 59) 96.

[193] *Nicaragua* case (n 69) para 176.

to be disproportional when they were not aimed at the exact area where the attack originated from.[194] For computer network attacks it is often challenging to find where they originate from and when employing a counter hack, to not also target a State's infrastructure.
Regardless, the ICJ has claimed that a State's right to self-defence could only be invoked towards other States.[195] This position has been met with a criticism and as going against the customary law principle on self-defence originating from the *Caroline* case.[196]

5 Conclusion

It is an often-heard misconception that recognition of a right to hack back would lead to chaos and vigilantism. As illustrated above, the contrary is true. The acceptance of a legal right to defend oneself or one's property has never been without limits. Through the evolution of legal thought, a right to self-defence has been recognised in situations where both the threat and the defensive action were immediate and the defender's actions were proportionate.
The whole concept of acting to defend life or property in legal systems revolves around actions that would normally be illegal if it were not for the initial attack justifying the response. This line of argument creates a mostly self-governing order to behaviour. As hacking back would be illegal if no prior attack and perpetrator can be proven to exist, one would just be guilty of an illegal act and liable to bear the consequences. Moreover, the idea of 'circumstances precluding wrongfulness' in international law can only excuse certain breaches of international obligations for as long as the condition that justifies this breach exists.[197] According to this, parties that have the technical capabilities to execute a counter hack will only do so under circumstances where they have ensured their justification to the largest possible extent. Furthermore, a recognition of permissible defensive action does not take away the distinction between defence and punishment, the latter which is a State prerogative.
Under established public international law, the appropriate response to a situation caused by a wrongful act attributed to another State is dependent on different circumstances. Because international law has been mainly focussed on State conduct so far, the thresholds have been thoroughly discussed in literature and decided on by international courts. Self-defence is reliant on an (imminent) armed attack, countermeasures on the breach of an international obligation owed by another State and necessity can only be invoked when essential interests are at stake and the reaction does not seriously impair the essential interests of the attacking State.
The three possible justifications for hacking back discussed here each have specific gaps for certain situations that can be complimented by one of the other concepts. Countermeasures

[194] *Armed Activities* case (n 116) para 147.

[195] *Consequences of the Construction of a Wall in the Occupied Palestinian Territory* (Advisory Opinion) [2004] ICJ Rep 136 para 139.

[196] Tsagourias (n 51) 241.

[197] *Gabčíkovo-Nagymaros* case (n 83) paras. 47–48. See also *Rainbow Warrior* arbitral award (n 131) para. 79.

can only be conducted by States, but they cannot be used against non-State actors if that non-State actor's conduct cannot be attributed to a State. A plea of necessity can justify actions taken against States or non-State actors, but its availability for non-State actors is very questionable. Finally, the principle of self-defence is a broad enough for all types of victims to rely on, but similar to necessity it is a last resort.

Regarding the gap theory something has to give. Either the traditional theory is followed that self-defence by States may not be invoked when they are the victim of a use of force not amounting to an armed attack. In that case the gap has to be filled by the possibility for States to resort to countermeasures that could constitute a use of force. Or the notion that countermeasures may never reach the level of force is upheld, in which case the gap theory should be discarded and Judge Simma's opinion in the *Oil Platforms* case should be followed.[198] Necessity is left out of the equation here since as 'armed attack' is a graver form of use of force, use of force will most likely never constitute a grave and imminent peril without also meeting the 'armed attack' threshold.

Cyber-attacks can, in some cases, be considered to constitute a use of force. Even if they do not meet that threshold, they can still have tremendous impact on individual lives. The main function and legitimisation of statehood is to have a monopoly on the use of force to protect citizens in a territory from internal and external threats of aggressive force.[199] It follows naturally that if a State fails to do so, citizens or more generally non-State actors, take this protection upon themselves. Private companies have taken this right into their own hands in some instances, but uncertainty about the legality of hacking back has forced those actions to happen under the radar.[200] This article hopes to provide a bit more clarity in that regard and outline the situations in which international law offer a legal basis to hack back.

---

[198] 'it makes no logical sense to prohibit a state from forcibly defending itself or permitting its allies to come to the state's defence where it is the subject of an unlawful use of force.' As Dinniss (n 59) 79 paraphrases Judge Simma (n 146) in para 12.

[199] Max Weber, 'Politics as a Vocation' in FROM MAX WEBER: ESSAYS IN SOCIOLOGY 77, 78 (HH Gerth & C Wright Mills eds., 1946) ("[A] state is a human community that (successfully) claims the monopoly of the legitimate use of physical force within a given territory.... Specifically, . . . the right to use physical force is ascribed to other institutions or to individuals only to the extent to which the state permits it.").

[200] Eichensehr (n 8) 499.

Table of reference

Literature

- Banks W, 'State Responsibility and Attribution of Cyber Intrusions after Tallinn 2.0' (2017) 95 Tex L Rev 1487
- Carr M, 'Public-private Partnerships in National Cyber-security Strategies' (2016) 92 International Affairs
- Crawford J, *Brownlie's Principles of Public International La*w (8th edn, OUP 2012)
- Dinniss HH, *Cyber Warfare and the Laws of War* (CUP 2012)
- Eichensehr KE, 'Public-Private Cybersecurity' (2017) 95 Tex L Rev 467
- Ferland J, 'Cyber insurance – What coverage in case of an alleged act of War? Questions raised by the Mondelez v. Zurich case' (2019) 35 Computer Law & Security Review 369
- Greenberg A, 'The Untold Story of NotPetya, the Most Devastating Cyberattack in History' (*Wired*, 22 August 2018) <https://www.wired.com/story/notpetya-cyberattack-ukraine-russia-code-crashed-the-world/>
- Halberstam M, 'Hacking Back: Reevaluating the Legality of Retaliatory Cyberattacks' (2013) 46 Geo Wash Intl L Rev
- Heathcote S, 'Circumstances Precluding Wrongfulness in the ILC Articles on State Responsibility: Necessity', in James Crawford, Alain Pellet & Simon Olleson (eds), *The Law of International Responsibility* (2010) 491
- Hessbruegge JA, 'The Right to Personal Self-Defense as a General Principle of Law' in: *Human Rights and Personal Self-Defense in International Law* (OUP 2017)
- Jensen ET, 'Computer Attacks on Critical National Infrastructure: A Use of Force Invoking the Right of Self-Defence' (2002) 38 STAN. J. INT'L L. 207
- Kesan JP and Hayes CM, 'Mitigative Counterstriking: Self-Defense and Deterrence in Cyberspace' (2012) 25 Harv J L & Tech 429
- Khanna P, 'These 25 Companies are More Powerful Than Many Countries' (*Foreign Policy*, 15 March 2016) <https://foreignpolicy.com/2016/03/15/these-25-companies-are-more-powerful-than-many-countries-multinational-corporate-wealth-power/>
- Lahmann H, *Unilateral Remedies to Cyber Operations* (CUP 2020)
- Modderkolk H, *Het is Oorlog maar Niemand die het Ziet* (Podium 2019)
- Petkis S, 'Rethinking Proportionality in the Cyber Context' (2016) 47 Georgia J Intl L
- Priddy A, 'The Use of Weapons in Counterpiracy Operations' in Stuart Casey-Maslen ed, *Weapons under International Human Rights Law* (CUP 2014)
- von Pufendorf S, *The Whole Duty of Man According to the Law of Nature* (Andrew Tooke tr, Ian Hunter/David Saunders eds, Liberty Fund 2003)
- Schaller C, 'Beyond Self-Defence and Countermeasures: A Critical Assessment of the Tallinn Manual's Concept of Necessity' (2017) 95 Tex L Rev 1619

- Schmitt M, *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations* (CUP 2017)
- Shaw M, *International Law* (6th edn, CUP 2011)
- Sloane R, 'On the Use and Abuse of Necessity in the Law of State Responsibility' (2012) 106 AJIL 447
- Tsagourias N, 'Cyber Attacks, Self-defence and the Problem of Attribution' (2012) 17 Journal of Conflict & Security Law 229
- Urgessa WG, 'Multilateral Cybersecurity Governance: Divergent Conceptualizations and its Origin' (2020) 36 Computer Law & Security Review
- de Vitoria F, *De Indis et de Ivre Belli Relectiones* (John Pawley Bate tr, Ernest Nys ed, Carnegie Institution 1917)
- Zenko M, 'The New Unknown Soldiers of Afghanistan and Iraq' (*Foreign Policy*, 29 May 2015) < https://foreignpolicy.com/2015/05/29/the-new-unknown-soldiers-of-afghanistan-and-iraq/>
- Zetter K, 'An Unprecedented Look at Stuxnet, the World's First Digital Weapon' (*Wired*, 11 March 2014) <https://www.wired.com/2014/11/countdown-to-zero-day-stuxnet/>

Case law

- *The Corfu Channel* (*United Kingdom v. Albania*) (Merits) [1949] ICJ Rep 4
- *Military and Paramilitary Activities in and Against Nicaragua* (*Nicaragua v United States*) (Merits Judgment) [1986] ICJ Rep 14
- *Case concerning the Gabčíkovo-Nagymaros project* (*Hungary v. Slovakia*) (Judgment) [1997] ICJ Rep 7
- *Naulilaa* (*Portugal v. Germany*) (Arbitral Award) [1928] Permanent Court of Arbitration
- *Responsabilité de l'Allemagne en raison des actes commis postérieurement au 31 juillet 1914 et avant que le Portugal ne participât à la guerre ('Cysne') (Portugal v. Germany)* [1930] 2 RIAA 1035
- ECHR *Case of Handyside v The United Kingdom* (Judgment) (7 December 1976)
- *Rainbow Warrior* (*New Zealand v France*) (Arbitration Tribunal) (1990) 20 RIAA 215
- *Legality of the Use or Threat of Nuclear Weapons* (Advisory Opinion) (1996) ICJ
- *Case Concerning Oil Platforms* (*Iran v. United States*) (Judgment) (2003) ICJ
- *Certain German Interests in Polish Upper Silesia* (*Germany v. Poland*), 1926 PCIJ (ser A) No. 7 (25 May 1926)
- *Trials of War Criminals before the Nuremberg Military Tribunals Vol XIV* (1946-1949) 308
- *Prosecutor v. Tadic* (Judgment) ICTY-94-1 (7 May 1997)

- *Consequences of the Construction of a Wall in the Occupied Palestinian Territory* (Advisory Opinion) [2004] ICJ Rep 136