Åbo Akademi University
Social Sciences, Business and Economics, and Law

Otto Hykkönen

TALLINN MANUAL RESTRICTIONS ON CYBER WARFARE

Master's Thesis

Supervisor: Viljam Engström

Master's Degree Programme in International Law and Human Rights

Faculty of Social Sciences, Business and Economics, and Law

Åbo Akademi University 2024

E-mail: otto.hykkonen@abo.fi

ÅBO AKADEMI UNIVERSITY – Faculty of Social Sciences, Business and Economics, and Law
Abstract for Master's Thesis

Subject: Public International Law, Master's Degree Programme in International Human Rights Law	
Writer: Otto Hykkönen	
Title: Tallinn Manual Restriction on Cyber Warfare	
Supervisor: Viljam Engström	Supervisor:

Abstract: Digitization expands the whole gamut of human activities. Warfare is also experiencing digitization in the form of introduction of cyber warfare. Cyber warfare and cyber operations in general are an ever-expanding field of warfare. Is the new front of cyberspace a wild west outside the reach of international humanitarian law? Is International humanitarian law applicable to cyber warfare? If so, to what extent? Is data, the stuff of cyberspace, even an object? These are the questions that the current thesis seeks answers. The study at hand aims to answer them via scrutiny of the Tallinn Manual, a collection of international humanitarian law norms collated by NATO's Cyber Cooperative Cyber Defence Center of Excellence, and the writings of esteemed scholars in the field. The study finds that the cyberspace is not a free-for-all zone but that international humanitarian law does, indeed, apply to cyber warfare as well. The crux is the understanding of the word "attack", which the current law ties to its effects outside cyberspace. Cyber attacks thus understood, are restricted by principles of necessity, distinction, and proportionality. Where the current law is not futureproofed is in the case of objecthood of data. Currently, data is only protected as per its use by other protected entities. The study finds that the understanding of an object in the 1987 commentary to Additional Protocol I as something tangible and visible does not reflect the reality of information, the destruction of which releases energy implying physical existence in a similar manner to other objects that may also lack tangibility and visibility.

Keywords: Cyber, Warfare, Lege Lata, Hacking, Tallinn Manual

Date: 17 April 2024	Number of pages: 75

Table of Contents

1. Introduction

1.1 Intro

The axiom "War never changes" is a frequently invoked adage. In a broader context, this adage bears a semblance of truth. Warfare, regardless of its modalities, invariably emerges as an extension of national policy. Implicitly or explicitly, it is invariably underpinned by a calculus weighing human suffering against national interests or, in the case of non-state actors, collective gains. While this foundational calculus retains its fundamental constancy, the minutiae of its calculations undergo evolutionary transformations. Throughout history, warfare has witnessed transformative shifts, often concomitant with major technological advancements. The domestication of the horse, for instance, gave rise to cavalry; seafaring vessels ushered in naval warfare, while the advent of aircraft extended conflict into the skies. Given the trajectory of technological innovation, it was almost inevitable that the creation of cyberspace, facilitated by digital computers and computer networks in their embryonic stages of what would eventually evolve into the internet, would engender yet another arena for human belligerence. This sphere, aptly labeled "cyber warfare," materializes through the manipulation of lines of code that remain obscured from human observation, comprehensible solely to digital systems. The data these programs target and manipulate is similarly concealed within the binary realm of 0's and 1's, discernible solely through machines endowed with the requisite logic to decode and interpret it.

The paramount divergence between cyber warfare and more traditional forms of warfare hinges on the schism between "cyberspace," the digitized realm encapsulated within computers and computer networks, and "realspace," encompassing the corporeal world beyond. Cyberspace primarily comprises data, often organized in bytes or bits of binary code, serving as the substratum for all digital information, encompassing everything from personal photographs to financial records and governmental registries. This data is subject to interpretation and manipulation through software, which constitutes a set of executable instructions for computers. While software is also couched in binary code, its function diverges from that of data. In the context of cyber warfare, the immediate target of attacks predominantly pertains to data, with software frequently employed as a vector to facilitate these assaults.

The vectors through which cyber warfare is waged encompass a broad spectrum, ranging from distributed denial of service (DDoS) attacks, wherein adversarial networks are inundated with a barrage of service requests, crippling their functionality; to hacking, the unauthorized penetration of an adversary's computer systems and networks; to malware, malicious software crafted with the explicit intent of causing harm to an opponent's systems and data. The degree of automation and invasiveness inherent to these vectors varies considerably, mirroring the spectrum of their real-world counterparts. DDoS attacks, for instance, often exhibit a high degree of automation, leveraging multiple computers to collectively inundate a target system, thereby saturating its bandwidth and rendering it virtually inoperable for the duration of the attack. Analogous to real-world area denial operations, DDoS attacks impede an adversary's utilization of a specific digital space, though they elude the enduring physical ramifications typically associated with realspace area-denial tactics. Critically, DDoS attacks seldom inflict actual damage to the targeted data, as they exclusively focus on impeding access to the system.

Hacking, whether pursued through brute force to breach an adversary's protective measures or via social engineering to manipulate individuals with authorized access, parallels infiltration missions or raids in the physical realm. While the mere act of unauthorized system access may initially appear aligned with the domain of espionage rather than warfare, the crux of the matter rests upon the hacker's intent and conduct, which confer the character of warfare[1] to the endeavor. Depending on the extent of access secured by the hacker, they wield the capacity to influence the contents of the targeted system, including data manipulation, copying, defacement, or outright deletion. Naturally, the level of access obtained by the hacker significantly constrains their capabilities.

Finally, malware represents the cyber realm's closest analog to conventional weaponry. Various categories of malware exist, each tailored to specific purposes. Some function as hacking tools, while others are pre-installed within systems as logic bombs or latent vulnerabilities awaiting activation. The latter two classifications lay dormant within a system until specific conditions are met (in the case of logic bombs) or until activation triggers are invoked. Notably, these entities operate akin to realspace weapons, their effects spanning a gamut from physical destruction, exemplified by Stuxnet's sabotage of Iranian nuclear centrifuges via safety system deactivation, to data erasure, as illustrated by Petya, which poses as ransomware while effectively obliterating the afflicted system's data, disguising it as

1 “Warfare” is used as a general term of prosecuting an armed conflict rather than a term of art within international humanitarian law proper

encrypted but retrievable when not so. The peril posed by these malware entities to computer systems and their data is manifestly evident.

The absence of specific regulatory frameworks governing cyber warfare does not necessarily detract from the protection afforded to civilians under *jus in bello*. This is due to the inherent adaptability of the international humanitarian law framework, which is capable of accommodating contingencies, including the challenge posed by cyber warfare. A notable issue in this context pertains to the non-recognition of attacks devoid of real-world consequences as "armed attacks". Such attacks, relegated to the virtual realm, evade the regulatory purview of *jus in bello*. The rationale underlying this non-protection lies in the intangible nature of computer data, which presents a glaring disconnect between the protective ambit of *jus in bello* and the ever-expanding digital facets of contemporary society. Paradoxically, the status quo appears to favor states actively engaging in offensive cyber warfare, affording them greater latitude of action, bereft of obligations to safeguard civilian digital data.

The research endeavor aims to ascertain the extent to which cyber warfare is presently regulated within the framework of international humanitarian law, as interpreted in the Tallinn Manual of 2017. The Tallinn Manual is a collection of rules applicable to cyber warfare, collated by a group of experts at NATO's Cooperative Cyber Defense Center of Excellence. The Manual itself is not a legally binding document but is supposed to reflect the current state of international humanitarian law as laid out elsewhere. However, as a reflection of *lege lata*, the Manual is considered authoritative and, subsequently, continues to be used as a reference by academics. For the current research, this status as an authoritative collection of rules is the main reason it will be used as the main source of *lege lata*. The significance of this research is underscored by the rapid proliferation of digital technologies within society and the inherent challenges posed by warfare that is both clandestine and potentially destructive. It is important to clarify that the objective here is not to advocate for the creation of new legislative measures but to establish a foundational understanding of the limitations delineated by a panel of experts responsible for the Tallinn Manual. In instances where dissenting viewpoints are salient, these will be explored to present a comprehensive perspective on the subject matter.

1.2 Research Question

The central research question addressed in this study is as follows: To what extent do the prevailing norms of international humanitarian law curtail the states' right to engage in cyber war and their

choices as per methods of cyber warfare? How do established norms, originally conceived for conventional warfare, accommodate a form of conflict characterized by its near-invisibility and potential for inconspicuous consequences? Furthermore, even if binding norms exist, to what extent do states adhere to them, given their sovereignty and propensity to assert autonomy in the realm of cyber warfare? This overarching question engenders several subquestions, notably those pertaining to the legal status of computer data as an object within the purview of international humanitarian law. The examination of object status is deemed pivotal, as it constitutes a threshold that profoundly influences the protective measures afforded by international humanitarian law. A determination that computer data qualifies as an object under this legal framework carries profound implications, triggering the application of principles such as distinction, proportionality, and precaution, thereby placing constraints on belligerent states engaged in cyber warfare. These constraints not only take into account potential harm inflicted within the digital realm but also extend to consider repercussions beyond the confines of cyberspace. This inquiry is situated within the broader context of the ongoing digitization of society, which continues to permeate an ever-expanding array of human activities, including ownership and control over critical data. Consequently, the affirmative resolution of the research questions regarding the protection of computer data under international humanitarian law would serve to establish a durable framework for the evolving landscape of digital warfare. Moreover, such a framework would strike a balance between military exigencies and humanitarian imperatives, thereby reaffirming the primacy of the latter within the context of armed conflicts.

1.3 Limitations

The thesis at hand does come with limitations on some pertinent questions. The issue of artificial intelligence is not discussed. This is, in part, due to the novelty of artificial intelligence at the time of writing and to the rapid development of the technology itself. The analysis written here would most likely be obsolete upon release. Furthermore, the capabilities of artificial intelligence are not well defined as to its capability to upend the way wars are fought.

Another limitation the thesis holds is the ambit of state practice it covers. Limiting state practice outside the thesis is done more with realities at the time of writing in mind. The process of collating an update to one of the main sources of the thesis, the Tallinn Manual, focused on state practice, is ongoing at the time of this writing.

1.4 Method and Materials

The primary source of international humanitarian law applicable to cyber warfare is the Tallinn Manual of 2014, subsequently updated in 2017. It is imperative to note that the Tallinn Manual does not possess binding legal force; rather, it constitutes a comprehensive analysis compiled by a panel of experts. Nevertheless, the manual holds eminence as a widely recognized and esteemed guide for interpreting the norms of international humanitarian law as they pertain to cyberspace. In the present thesis, the Tallinn Manual serves as the chief source for elucidating the *lege lata* pertinent to the subject matter. As an official publication of the NATO Cooperative Cyber Defense Center of Excellence (CCDCOE), the manual's authority derives added weight, bolstered by its endorsement from the world's most influential military alliance.

Methodologically, this study adheres to a qualitative approach, commencing with a comprehensive exploration of the *lege lata* within and beyond the realm of cyber warfare. The study draws upon a corpus of published academic articles and pertinent literature, eschewing quantitative methods such as questionnaires. Moreover, this thesis embraces a *de lege ferenda* approach, acknowledging the inherent fluidity of the subject matter. At the time of writing, the Tallinn Manual is in the process of assembling its 3.0 version, potentially signifying alterations to its interpretive guidelines. These prospective developments could exert a significant impact on state attitudes toward military operations in cyberspace. Consequently, this study not only assesses the current applicability of international humanitarian law to cyber operations but also engages in critical analysis of the requisite changes needed to ensure the continued relevance and efficacy of humanitarian safeguards for non-combatants in the ever-evolving realm of cyberspace.

In terms of primary source material, this study primarily draws upon the Geneva Conventions and their Additional Protocols, which constitute the primary sources of positive international humanitarian law. Furthermore, the updated Tallinn Manual of 2017 serves as a central reference, notwithstanding its non-binding character. The Manual's prominence lies in its wide acceptance as an authoritative interpretation, frequently referenced by academics in the field of international humanitarian law within the context of cyberspace. Additionally, this study leverages the scholarly contributions of other academics and scholars, as well as the publicly available national positions of states relevant to the research inquiry. These supplementary sources serve to either reinforce or challenge the positions

articulated within the Tallinn Manual, shedding light on ongoing debates and potential shifts in legal interpretations.

To address the central research question and its associated subquestions, this study adopts a qualitative approach. The tone of the thesis will be critical to the current law and its understanding of issues such as the definition of object, which are based on an analysis decades out of sync with ruling realities. The thesis begins by defining a list of core concepts, including a short history of cyber warfare. The thesis continues with an examination of the prevailing norms within international humanitarian law, encompassing a comprehensive review of the principal sources, namely the Geneva Conventions of 1949 and their Additional Protocols I and II of 1977. This foundational inquiry is further buttressed by an exploration of authoritative, though non-binding, interpretive sources of law, with a particular focus on the Tallinn Manual 2.0. Given the manifold complexities inherent to the subject matter, this initial segment of the research is dedicated to elucidating the fundamental norms that govern all forms of warfare, both cyber and conventional. Core concepts such as "armed attack" and "protected object" undergo conceptual analysis within the context of international humanitarian law interpretations, laying the groundwork upon which subsequent analyses of various means and methods of cyber warfare and computer data will be conducted.

As articulated above, the approach to collating relevant norms predominantly adopts a linguistic framework. The analysis commences with the examination of the phrase "objects protected from an attack" and its constituent elements. This approach has been selected due to its suitability for dissecting the subject matter into discrete components, which can then be examined individually.

Subsequently, the study pivots its focus towards an appraisal of computer data and its potential classification as an object under international humanitarian law. This segment endeavors to determine whether computer data can attain an inherent status as an object within this legal framework and whether such classification can fluctuate. This inquiry spans across the conventional definitions of "object" as stipulated in the primary sources of international humanitarian law, juxtaposed against viewpoints that may expand beyond these established definitions. Where pertinent, the national perspectives of Finland, the United Kingdom, and the United States will be used to add nascent state practice to the issue at hand.

The third section of the study delves into the known vectors of cyber attack, scrutinizing their compatibility with the overarching framework elucidated in the first section. The objective here is to

assess the various cyber operations and software suites employed in cyber warfare, with a particular focus on their conformity with the principles of *jus in bello*. This evaluation relies on a thorough examination of reports published by cybersecurity labs and organizations, providing insights into the functionalities and potential harm posed by these software and operational techniques to computer systems and the associated data. Furthermore, this section scrutinizes the possibility of these operations or software spreading in contravention of the limitations prescribed by international humanitarian law. It also explores known instances of cyber warfare operations, offering insights into the extent to which they adhered to the constraints of international humanitarian law. Within this context, the study explores the potential categorization of certain cyber operations as "armed attacks" under the lens of *jus ad bellum*, shedding light on the divergence between the layman's perception of a "cyber attack" and the precise legal criteria defining such actions. Notably, this section also revisits the national positions of the aforementioned states to glean additional insights into emerging state practices and perspectives on cyber warfare.

In the subsequent phase of the study, the findings amassed throughout the research are synthesized, culminating in the formulation of conclusions. These conclusions encompass reflections on prospective developments in the field and considerations regarding the future trajectory of cyber warfare.

2. Cyber Warfare and History Thereof

2.1 Cyber Warfare

The understanding of the concept of cyber warfare begins with examining the meanings of individual words 'cyber' and 'warfare'. 'Cyber' traces its origins to 'cybernetics', a word from the 1940's denoting the study of humans interfacing with machines. The word evolved closer to its current form as a prefix in the 1984 novel Neuromancer by William Gibson. In the novel, 'cyberspace' is a globe-expanding network used to visualize digital data. With the invention and proliferation of such a network as the Advanced Research Projects Agency Network (ARPANET) and, later, the internet, the word would separate itself from fiction and begin to have a meaning in the real world. Since the network is the space in 'cyberspace', the cyber would have to be a denotation of the nature of the network. The network is made of computers and connections between them. Ergo, the definition of 'cyber' is anything to do with computers and connections between them.

'Warfare' has been given multiple different definitions throughout history; however, it can be tied to the definition of 'war' as it is the means and methods thereof. Sun Tzu defines 'war' as "a matter of life and death [to the state], a road to safety or ruin".[2] On a more practical level, 'warfare' is "to take the enemy's country whole and intact"[3] and, in its supreme form, "breaking the enemy's resistance without fighting".[4] However, when the supreme form of warfare is not reached, destruction and death follow. For von Clausewitz, war is "nothing but a duel on a larger scale"[5] and "an act of force to compel our enemy to do our will."[6] The crux of both Sun Tzu and von Clausewitz's definitions is that war is an event where one or both parties must apply force to another to attain a goal and the means of applying that force. Ergo, the definition of war is an event, and warfare is the means of violence between parties.

Combined, the definition of 'cyber warfare' becomes the following: means of violence connected with, or involving, computers and computer networks. 'Cyber' is an all-encompassing prefix in cyber

2 Sun Tzu, Art of War (Allandale Online Publishing c2000) 1
3 *Ibid*. 8
4 *Ibid*.
5 Carl von Clausewitz, On War (Princeton University Press c1976) 75
6 *Ibid.*

warfare. The field of battle is ‘cyberspace’; the weapons of choice are ‘cyber weapons’; armed forces of the world hire and train dedicated ‘cyber operatives’[7] to conduct ‘cyber operations’, etc.

Stiennon’s definition of ‘cyber warfare’ is as follows and mirrors the definition above:

> Cyber warfare is an extension of policy by actions taken in cyberspace by state actors (or by non-state actors with significant state direction or support) that constitute a serious threat to another state’s security, or an action of the same nature taken in response to a serious threat to a state’s security (actual or perceived).[8]

As was illustrated above, the term ‘cyberspace’ began its life as a term in speculative fiction. Later on, with the invention and proliferation of similar digital data sharing networks as described in Neuromancer, such as ARPANET and the internet, the term would become to denote a computer network. Today, the internet is the most well-known example of cyberspace, but the term is not limited to globe-spanning mega-networks of millions of computers, servers, or sub-networks. In fact, computers and networks of computers exist outside the internet, and cyberspace exists within them. Cyberspace is all the space within which digital data moves or is used.

Using similar heuristics as above, the definition of ‘cyber weapon’ can be understood in its most basic sense. ‘Weapon’, in its most basic sense, means an instrument of violence. Weapons come in many different forms and use various different principles to cause violence, but the basic function of all weapons is to cause violence. Non-weapons can become weapons in use; one can take a tool such as a hammer and use it to project violence on another. At this moment, the hammer loses its character as a tool and becomes a weapon, albeit an improvised one. Ergo, the definition of ‘cyber weapon’ is tools of violence within cyberspace.

On a less theoretical level, cyber weapons are software meant to cause harm or provide illicit access to a computer or a computer network. These can be further divided into autonomous software that proliferates itself and executes itself without an outside command and operator-controlled software that may proliferate and hide itself within the infected system autonomously but does not execute itself without confirmation from the operator.

7 Royal Navy, 'Cyber Operator' (Job Role, 22 September 2020) <https://www.royalnavy.mod.uk/careers/roles/cyber-operative> accessed 30 January 2024

8 Richard Stiennon, A Short History of Cyber Warfare. in James A. Green (ed), Cyber Warfare: A Multidisciplinary Analysis (Routledge 2015) 8

Cyber weapons are similar to malware, such as viruses and trojans, in function. The distinction comes from the fact that cyber weapons are often purposefully created to perform a specific task in a specific environment, where common viruses are often less sophisticated and less clandestine in their function. Additionally, the users of cyber weapons are often state organs, such as cyber warfare units of the military or intelligence agencies, or state-affiliated hacker groups, whereas the users of more traditional malware are less organized, sometimes singular people.

Cyber operations can be either offensive, such as the deployment and use of cyber weapons, or defensive, such as counter-hacking or IP-blocking. Due to their nature as more active, most cyber operations are offensive. Cyber defense is deployed in anticipation of the enemy's offensive operations.

Offensive operations are not limited to the use of cyber weapons. An offensive operation can overwhelm the target with incoming traffic, thus rendering the target temporarily useless. Hacking into a system can happen via a software suite that is installed on the system surreptitiously or via the use of pre-installed backdoors in the hardware or vulnerabilities in the software. Getting credentials can happen via a bad link in an email or via installing a keylogger on the machine.

Cyber defense, like defensive works in realspace, is created in anticipation of an attack. Cyber defense itself does not meaningfully differ from normal cyber security measures taken by companies. The combination of measures that create a comprehensive cyber defense includes installing and updating anti-virus software and operating systems both in the computers themselves and network infrastructure; shutting down ports that are not used for data transfer; using localized networks not connected to the wider internet; and practicing overall cyber hygiene, which can take the form of using encryption, not connecting foreign machines or mass storage devices to a critical network; using complex and changing passwords, etc. Cyber defense can be active, such as in cases of attack-back mechanisms, the most common form of which is counter-hacking. Counter-hacking consists of identifying the attacker's IP address and then launching an attack against it. However, the attacker can use a spoofed IP address other than their own, which would direct the counter-hack in the direction of someone other than them.

Armed forces around the world hire and train specialists in cyber warfare. However, a more common form of cyber operator is an independent group of hackers sponsored by the state. The reason states prefer the use of independent groups is the problem of attribution. Using an independent group makes it harder for the victim state or states to point at the state as the originator of the attack. Should the IP address of the attack be traced back to the origin state, the attacker can disavow all connections to the

group. This creates confusion on the nature of the attack; if the attacker is pointed to be the state, the possibility of a military option can be considered should other qualifications for an armed attack be fulfilled; if the attacker is an organization with tenuous or no provable links to the state, the attack is closer in nature to cyber crime, where the choice of response is limited to the sphere of law enforcement and international cooperation on that level. The previous analysis changes if the hacker group can be linked to a non-state armed group where the possibility of a forceful response can be considered.

Some of the famous hacker groups with suspected state sponsorship are Double Dragon and Red Apollo (China); Lazarus Group (North Korea); Fancy Bear, Venomous Bear/Turla, and Sandworm (Russia); and Equation Group (United States). The presented groups have been involved in cyber crime like the 2014 Sony hacks; cyber espionage such as the hacking of the Democratic National Committee of the United States Democratic Party in 2016; and cyber warfare such as the development of cyber weapons like NotPetya and Stuxnet. All of the above are suspected to have links to, or be an organ of, their origin states; however, none can be conclusively proven to be such.

2.2 Brief History of Cyber Warfare

2.2.1 Moonlight Maze (1996-2016)

The history of cyber warfare runs parallel with the history of cyber espionage. The link between the two is through similar means and methods. Software used for hacking into a system can be used both for data exfiltration and data destruction, all dependent on the fashion in which a given suite is made, since gaining access and different ways to manipulate data within the system can require different levels of access to the system. The maker of the software suite needs to choose what level of access it tries to gain within the hacked system. Often, the level of access wanted is root access, which gives the operator the ability to access all data contained with full access to manipulate it. However, as the malware used in cyber warfare and cyber espionage is often modular and made for the task, gaining root access can be disregarded if the aim of the operation so allows.

With the similarities between cyber warfare and cyber espionage in mind, the history of cyber warfare starts with Moonlight Maze. Moonlight Maze was a string of attacks on multiple U.S. governmental

departments such as the Pentagon, the Department of Energy, and NASA, in the process exfiltrating a quantity of data that, if printed on paper, would reach the height of three times the Washington Monument. The attacks are thought to have started in 1996. Moonlight Maze is one of the first classifications of Advanced Persistent Threat (APT), which is given to threat actors with direct links to, or close ties to, a state. Initially, Moonlight Maze was thought to be an isolated attack, but forensic analysis shows that similar attacks had been perpetrated with similar malware suites at least until 2016.[9] The analysis of the means and methods used suggests[10] that the originator of the attacks is Russian, more specifically a known APT called Turla that has been connected to Russia's intelligence agency FSB. Should such a link exist, Moonlight Maze would be one of the first instances of a state using cyber means against another.

2.2.2 Bronze Soldier Cyber Attacks (2007)

The idea of cyber warfare became known to the public in 2007 with the Bronze Soldier incident and subsequent cyber attacks against the Estonian government and companies. The incident began with Estonian plans to relocate a Bronze Soldier monument commemorating the Soviet soldiers killed in the liberation of Tallinn and the remains of Soviet soldiers buried therein from downtown Tallinn to the Tallinn Military Cemetery. The monument was the focal point of the already existing tension between the Estonian and Russian-speaking strata of Estonian society. The decision to relocate the monument led to protests from the Russian government and a Russian disinformation campaign both in traditional media and the Russian parts of the internet.

The decision to move the statue led to a wave of cyber attacks on Estonian governmental and private actors such as the parliament and several banks, newspapers, and telecommunication service providers. The cyber attacks varied in volume and method of attack and lasted for three to four weeks in total.[11] Most of the attacks were different forms of denial-of-service attacks, the goal of which was to make web resources unavailable by overwhelming them with queries. Some hacking via SQL injection was involved, and on the Russian internet, various instructions for hacking and other forms of attack were posted. Most of these hacks were not successful, and those that found success were against non-critical

9 Juan Andres Guerrero-Saade and others, Penquin's Moonlit Maze: The Dawn of Nation-State Digital Espionage (Kaspersky Lab 2015) 13-14

10 *Ibid.* 14-15

11 Nato Cooperative Cyber Defence Centre of Excellence, 'Cyber Attacks Against Estonia (2007)' (Cyber Law, 15 October 2018) <https://cyberlaw.ccdcoe.org/wiki/Cyber_attacks_against_Estonia_(2007)> accessed 13 February 2024

resources.[12] The traffic aimed at causing the denial of service was overwhelmingly outside Estonia and was politically motivated as per the use of language in queries flooding the servers.[13]

The result of these cyber attacks was a failure of server infrastructure that rendered the sites inoperable for the duration of the attacks. As a countermeasure Estonian actors under attack blocked all foreign traffic to their servers, effectively cutting them off from the rest of the internet outside Estonia. Not only did this blocking of outside traffic render web services unavailable, but many Estonian banks also stopped accepting transfers to foreign accounts as a preventative countermeasure.[14] Also, some web pages were defaced with anti-Estonian slogans and had edits to their content. The estimated financial loss caused by the attacks is 1 million U.S. dollars.[15] However, material losses were limited to ancillary internet infrastructure, causing damage to routers.

Due to the nature of the attacks, it is impossible to definitively say who perpetrated them. However, the scale of the attacks suggests that an organized group must have been involved. Furthermore, many attackers' IP addresses point to Russia as their origin. Estonia never publicly accused Russia of these attacks. In the wake of the attacks, NATO established its Cooperative Cyber Defence Center of Excellence (CCDCOE) in Tallinn.

2.2.3 Stuxnet (2010)

The Stuxnet attack is significant in the history of cyber warfare for two reasons: it is the first time an autonomous cyber weapon has been deployed, and it is the first cyber warfare operation with destructive consequences in realspace. As will be seen below, the latter fact is the definitive starting point of cyber warfare proper.

Stuxnet itself was a worm, a type of self-propagating malware, designed to attack industrial centrifuges for the Iranian nuclear program. The development of Stuxnet is speculated to have started in 2005; however, it was first detected on June 17, 2010.[16] The network of the targeted centrifuges was a closed one, so it was not connected to the outside internet. This has led to the hypothesis that it was introduced

12 Rain Ottis, Analysis of the 2007 Cyber Attacks Against Estonia from the Information Warfare Perspective (NATO Cooperative Cyber Defence Command of Excellence 2018) 2
13 *Ibid.*
14 *Ibid*.
15 Nato Cooperative Cyber Defence Centre of Excellence (n 10)
16 Nato Cooperative Cyber Defence Centre of Excellence, 'Stuxnet (2010)' (Cyber Law, 18 December 2018) <https://cyberlaw.ccdcoe.org/wiki/Stuxnet_(2010)> accessed 14 February 2024

to the system via a USB flash drive. Once inside the system, Stuxnet would use several zero-day vulnerabilities, weak points in device security that are part of the system and remain unpatched, and hide itself in the machine. Next, it would scan the system for the specific programmable logic controls (PLC) within the system. Should a specified PLC be found, it would then take control of it and start raising the centrifuges' speed, causing mechanical damage while reporting no discrepancies with their operation, thus further hiding itself. Should the infected system be without the specified PLC, it would remain hidden on the system, only propagating itself to other machines in the network. The propagation of Stuxnet is extraordinary in the fact that it managed to break from the closed network and spread to machines outside, demonstrating the ability of an autonomous cyber weapon to cause potentially widespread damage and run amok.

The origin of Stuxnet remains a mystery since no actor has taken blame for it, and attempts to trace its development have not been successful. However, it has been speculated to be a joint creation of Israeli and American origins. Analysis in 2019 has managed to discover a conglomerate of threat actors, named Gossip Girl, having links to the operation of Stuxnet.[17] One of the threat actors included in Gossip Girl is the American hacker group, Equation Group, which has been speculated to have links to the National Security Agency (NSA) of the United States.[18]

2.2.4 Cyber Warfare in the Russo-Ukrainian War (2014-)

The Russo-Ukrainian conflict is filled with examples of cyber warfare, with many firsts when it comes to cyber warfare methods affecting a state. In 2014-2016, APT Fancy Bear developed and distributed a malicious application (X-Agent) meant to look like an application used by Ukrainian artillery forces with D-30 howitzers.[19] X-Agent was meant to collect and relay data, including geo-location data from Ukrainian artillery. During the time X-Agent was in use, some 25-50% of Ukrainian D-30 howitzers

17 *Ibid.*

18 Nato Cooperative Cyber Defence Centre of Excellence, 'The Shadow Brokers publishing the NSA vulnerabilities (2016)' (Cyber Law, 20 December 2018) <https://cyberlaw.ccdcoe.org/wiki/The_Shadow_Brokers_publishing_the_NSA_vulnerabilities_(2016)> accessed 14 February 2024

19 Nato Cooperative Cyber Defence Centre of Excellence, 'Use of malware to track and target Ukrainian artillery units (2014-2016)' (Cyber Law, 29 August 2023) <https://cyberlaw.ccdcoe.org/wiki/Use_of_malware_to_track_and_target_Ukrainian_artillery_units_(2014-2016)> accessed 14 February 2024

were lost, though the extent to which X-Agent was the cause is not known. Ukrainian officials deny that it led to any rise in the loss rates of the artillery pieces.[20]

In 2015, the Russian APT Sandworm used malware dubbed BlackEnergy against the Ukrainian power grid. Some 225,000 people suffered from the power outage caused by it. This was the first time a state's critical infrastructure had been targeted and meaningfully crippled by cyber warfare.[21] The power grid was ultimately restored to its prior operability. Speculations about the motive of the attack point to it being a testrun on how a cyber weapon could be used against a state's critical infrastructure.

In 2017, NotPetya, a wiper software meant to delete data from infected machines, was used against the Ukrainian financial sector. It masqueraded itself as ransomware, a malware that encrypts data and demands payment for decryption; however, it lacked the capability of restoring the encrypted data. The attack has been officially attributed to Russia by Ukraine, the United Kingdom, and the United States.[22] NotPetya caused economic losses estimated to reach 10 billion USD.[23]

In 2022, two wipers, malware designed to destroy data on a system, spread on Ukrainian systems: Whispergate and HermeticWiper/FoxBlade. Both were meant to cause damage via the deletion of data within the infected system, their targets being Ukrainian government agencies[24] and enterprises, with HermeticWiper/FoxBlade also found in systems in Lithuania and Latvia.[25] Whispergate has not been connected to any specific actor; however, the scale of its propagation hints at an organized, possibly state-sponsored actor. HermeticWiper/FoxBlade has been connected to a Russian group called IRIDIUM with suspected links to Russian military intelligence (GRU). Both wipers were successful in making the infected systems inoperable.

20 *Ibid.*

21 Nato Cooperative Cyber Defence Centre of Excellence, 'Power grid cyberattack in Ukraine (2015)' (Cyber Law, 20 December 2018) <https://cyberlaw.ccdcoe.org/wiki/Power_grid_cyberattack_in_Ukraine_(2015)> accessed 14 February 2024

22 Nato Cooperative Cyber Defence Centre of Excellence, 'NotPetya (2017)' (Cyber Law, 15 October 2018) <https://cyberlaw.ccdcoe.org/wiki/NotPetya_(2017)> accessed 14 February 2024

23 *Ibid.*

24 Nato Cooperative Cyber Defence Centre of Excellence, 'Cyber operations against government systems in Ukraine (January 2022)' (Cyber Law, 28 February 2022) <https://cyberlaw.ccdcoe.org/wiki/Cyber_operations_against_government_systems_in_Ukraine_(January_2022)> accessed 14 February 2024

25 Nato Cooperative Cyber Defence Centre of Excellence, 'HermeticWiper malware attack (2022)' (Cyber Law, 27 May 2022) <https://cyberlaw.ccdcoe.org/wiki/HermeticWiper_malware_attack_(2022)> accessed 14 February 2024

3. Norms of International Humanitarian Law

Before delving into the specific inquiry regarding cyber warfare and the application of international humanitarian law norms to it, it is essential to establish a foundational understanding of the general norms within international humanitarian law. These norms find their primary embodiment in the Geneva Conventions of 1949, complemented by the Additional Protocols of 1978 and 2005. The Geneva Conventions are accompanied by extensive commentary, elucidating various aspects, such as the definitions of terminology used in the conventions, the intentions of the conventions' drafters, and more. These conventions, alongside their accompanying commentary, serve as the initial reference points for this analysis. Additionally, authoritative interpretive guides like the Tallinn Manual will be consulted where applicable. While the Tallinn Manual primarily focuses on international humanitarian law in the context of cyberspace and cyber warfare, it has addressed analogous issues, the insights and conclusions of which may provide valuable guidance for this thesis.

The elements under scrutiny encompass the following: the interpretation of "objects" as understood within the purview of international humanitarian law, the conceptualization of "attack," and the nature of the protection afforded to these objects. This logical progression of examination will be adhered to throughout the chapter. Notably, it is essential to clarify that in the context of this thesis, the term "objects" pertains specifically to non-person entities rather than objects in a general sense.

3.1 Objects under International Humanitarian Law

In order to understand the norms of international humanitarian law as they pertain to cyber warfare, it is crucial to begin by examining the objects granted protection under this body of law. These objects are the linchpin of the entire system, as they define the scope of constraints imposed on the conduct of hostilities. As previously stated, this thesis focuses specifically on non-person objects within the context of international humanitarian law. The common denominator among all objects under international humanitarian law is their status as non-combatants.

The protection of objects under international humanitarian law is established in the Geneva Conventions and their Additional Protocols. To illustrate the protection afforded to non-person objects, several key points from these instruments are highlighted below:

1. Medical Units and Materiel (Convention I): Convention I grants non-combatant status to medical units and their associated materiel.[26] This would reasonably encompass equipment such as computers, systems, and databases containing medical data, including medical records. The rationale is that these items are essential for medical units to fulfill their humanitarian tasks.

2. Personal Belongings of Prisoners of War (Convention III): Convention III provides protection for the personal belongings of prisoners of war.[27] However, the commentaries from 2020 suggest that items like mobile phones or personal data storage devices, such as thumb drives, may be considered comparable to military equipment and military documents, which are liable for confiscation.[28] Nevertheless, there is a provision allowing for the withdrawal of valuables on security grounds, and objects capable of data transfer, like mobile phones, are prime candidates for such security-related actions.[29]

3. Valuables of Interned Civilians (Convention IV): Convention IV has similar provisions regarding the valuables of interned civilians. However, it lacks the explicit possibility to confiscate valuables on security grounds, implying that mobile phones and personal data drives of civilians cannot be confiscated outright. A lack of such possibility is present because the commentaries to Convention IV are from a time before mobile computing.

The primary source that defines civilian objects within the Conventions and Protocols is Additional Protocol I. The Protocol starts by distinguishing between civilian populations and objects on the one hand and military objectives on the other.[30] Article 52 of the Protocol provides a negative definition of civilian objects, stating that they are "all objects which are not military objectives." The Protocol then elaborates on what constitutes military objectives, considering factors like the nature, location, purpose,

26 Geneva Convention for the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field (adopted 12 August 1949 entered into force 21 October 1950) 75 UNTS 970 (Convention I) art. 33

27 Geneva Convention Relative to the Treatment of Prisoners of War (adopted 12 August 1949 entered into force 21 October 1950) 75 UNTS 972 (Convention III) art. 18

28 International committee of the red cross, 'Property of Prisoners' (Commentary of 2020, 14 December 2022) <https://ihl-databases.icrc.org/en/ihl-treaties/gciii-1949/article-18/commentary/2020?activeTab=undefined> accessed 15 April 2024

29 *Ibid*.

30 Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts (adopted 8 June 1977 entered into force 7 December 1978) 1125 UNTS 17512 (Additional Protocol I) art. 48

and use of the object, as well as the definite military advantage its destruction, neutralization, or capture would provide at the time of the analysis. This definition implies that the classification of an object as a military objective can change over time and depends on the circumstances. In cases of ambiguity as to the status of an object, it shall be considered a civilian object.

While this approach may seem to give substantial discretion to individual military commanders, it strikes a balance between humanitarian protection and military necessity. Stricter definitions might risk non-compliance by parties, potentially leading to more significant harm to non-combatants.

Sassoli introduces a distinction between military objectives and civilian objects based on the use of an object by enemy forces at the time of targeting. According to his perspective, every material and tangible object can be either a military objective or a civilian object protected by international humanitarian law norms.[31] The criteria for an object to be a military objective include its current use by the enemy and the requirement that its destruction or neutralization would yield a definite military advantage to the attacker. The rationale behind this analysis is that for military victory, only the military capacity of the adversary must be depleted.[32]

Articles 53 and 54 of Additional Protocol I provide specific protections for cultural objects, objects of worship, and objects indispensable to the survival of the civilian population. Article 53 prohibits attacks or reprisals against objects of worship and cultural significance. Article 54 prohibits the use of starvation as a method of warfare and extends this prohibition to any acts that would *de facto* lead to the starvation of the civilian population. However, if objects indispensable to survival are in the sole use of the armed forces of an adverse party or are used directly in support of military action, starvation and means of warfare leading to it may be permitted under specific conditions. An example of such specific conditions would be the use of scorched earth tactics by a retreating defender.

The effects-based approach of Article 54 is particularly relevant to the thesis as it pertains to modern industrial farming, computers, software, and data that play critical roles in food security. This introduces additional vulnerabilities in the food security system that must be protected against during the conduct of cyber warfare operations. Similarly, installations containing dangerous forces may have kindred vulnerabilities and should not be made the target of an attack.[33]

31 Marco Sassòli, Legitimate Targets of Attack under International Humanitarian Law (International Humanitarian Law Research Initiative 2003) 2
32 *Ibid.* 3
33 Additional Protocol I art. 56

Additional Protocol II extends similar protections to objects indispensable to survival, installations containing dangerous forces, and cultural objects, but in the context of non-international armed conflicts.[34] However, it does not provide the same latitude for military necessity to derogate from the protections as found in Additional Protocol I. This distinction arises from the typically uneven nature of non-international armed conflicts, where civilian populations are more likely to be caught in the crossfire.

It's important to note a significant discrepancy between treaty law in international and non-international armed conflicts. While the ICRC had initially extended similar protections to civilian objects in both types of conflicts, Additional Protocol II offers narrower protection than Additional Protocol I.[35] This discrepancy is attributed to concerns about legitimizing non-state armed groups' actions and nebulous references to state sovereignty.[36] However, these concerns have been challenged on the basis that international humanitarian law already permits non-state armed groups to target state soldiers, and thus, affording protection to civilian objects would not further legitimize their actions.[37] Furthermore, nothing prohibits states from using their domestic criminal codes to punish members of non-state armed groups.

Of particular relevance to the thesis is Article 60 of Additional Protocol I, which addresses demilitarized zones. In the context of cyber warfare, where computer networks and data are often decentralized, the demarcation and protection of such zones can be challenging. Accidental loss of protection due to the dual-use nature of systems or data banks, as well as the unintended spread of cyber warfare effects, is a concern. Similar considerations apply to Geneva Convention I and hospital zones and Convention II regarding hospital ships. These aspects underscore the complexities of applying international humanitarian law to cyber warfare scenarios.

34 Protocol Additional to the Geneva Conventions of 12 August 1949 and Relating to the Protection of Victims of Non-International Armed Conflicts (adopted 8 June 1977 entered into force 7 December 1978) 1125 UNTS 17513 (Addtional Protocol II) art. 14-16

35 Noam Zamir, 'Distinction Matters: Rethinking the Protection of Civilian Objects in Non-International Armed Conflicts' (2015) 48 Isr L Rev 111 113

36 *Ibid*.

37 *Ibid*. 116-117

3.2 Attack

3.2.1 Attack *Jus in Bello*

The subsequent aspect under scrutiny is the concept of an "attack." This investigative step is substantiated by the necessity to establish what activities are encompassed by this term, as this delineation is pivotal in elucidating the ensuing discussion on how these entities are safeguarded from the effects of it. The primary source employed to address this issue is Additional Protocol I. This protocol offers a succinct and unequivocal definition of an attack as "acts of violence against the adversary, whether in offense or defense."[38] The 1987 Commentary on this protocol provides a rationale for specifying that the term 'attack' encompasses both offensive and defensive acts of violence, grounded in an effects-based approach. This perspective emphasizes that both attacks and counterattacks can have comparable consequences for civilians and non-combatants alike.[39]

To enhance clarity, the Commentary suggests an alternative term, namely "combat action," to denote the instruction given to armed forces.[40] This alternative terminology lends greater intelligibility to the concept, as the term 'attack' sometimes implies actions solely undertaken by the aggressor, even when the actions of both parties may be functionally indistinguishable. The Commentary underscores this distinction and further expounds upon it.[41] Moreover, the chosen word in the Commentary underscores that 'attack' as understood in *jus in bello* differs from 'armed attack' as understood in *jus ad bellum*. These two notions are not interchangeable, and this distinction is reiterated in the Commentary.[42] Additionally, the Commentary elaborates on the fact that the latter concept primarily concerns issues of responsibility for the conflict, whereas the former is concerned solely with the utilization of weapons in an aggressive manner.

A vital criterion that 'combat action' must satisfy to be classified as an attack under international humanitarian law is that it must be directed against an adversary.[43] This requirement is grounded, in part, in the principle of military necessity, as certain forms of 'combat action' necessitate such differentiation. The Commentary exemplifies this with the notion of a scorched earth policy. While an

38 Additional Protocol I art. 49
39 International Committee of the Red Cross, 'Definition of attacks and scope of application' (Commentary of 1987, 14 March 2023) <https://ihl-databases.icrc.org/en/ihl-treaties/api-1977/article-49/commentary/1987?activeTab=undefined> accessed 15 April 2024
40 *Ibid*.
41 *Ibid*.
42 *Ibid.*
43 *Ibid.*

invading army is categorically prohibited from targeting objects indispensable for the survival of the civilian population under any circumstances, the prohibition may not be as absolute when applied to a defending army operating within its own territory. In cases of "imperative military necessity,"[44] the defending army may resort to a campaign of scorched earth, as the destruction of its own country's resources may not be construed as an attack per se.

Schmitt's analysis underscores a pivotal factor in the definition of an attack under international humanitarian law: it is the violent nature of acts that constitutes an attack, rather than the specific targets of these acts.[45] This perspective shifts the focus away from the common understanding that attacks solely involve violent acts against an adversary, thus recognizing that attacks encompass acts of violence regardless of the recipient. This perspective is grounded in the premise that acts of violence or 'combat actions' against non-combatants should unequivocally be categorized as attacks. When evaluating the violent acts themselves, the determining factor in ascertaining whether they qualify as violent acts under international humanitarian law is the resultant consequences. This consequence-based approach aligns with the core objective of international humanitarian law, which is to protect non-combatants from the deleterious consequences of armed conflict.[46]

Some analysts initiate their examination of the term 'attack' by distinguishing between the Geneva and Hague tracts of international humanitarian law. They contend that the Conventions and Additional Protocols employ language that bridges both tracts, implying that the understanding of the term 'attack' should be similarly comprehensive.[47] For instance, it is argued that since Convention I employs language suggesting that the seizure of a medical facility constitutes an offense against the protection afforded to military medical units and their provisions, the scope of 'attack' must extend beyond the general understanding in Hague law, as delineated in Additional Protocol I, to encompass acts of violence. Notably, this argument does not consider the specificity of the language regarding medical units and materiel but rather emphasizes the broader analysis.[48] Another example presented to support this broad understanding of 'attack' pertains to cultural objects and the protection accorded to them. The

44 *Ibid*.

45 Michael N. Schmitt, “Attack” as a Term of Art in International Law: The Cyber Operations Context. in Czosseck and others (eds), 4th International Conference on Cyber Conflict (NATO Cooperative Cyber Defence Command of Excellence 2012) 289

46 *Ibid*. 290-291

47 West Point Lieber Institute, 'The Definition of an “Attack” under the Law of Armed Conflict' (Articles of War, 3 November 2020) <https://lieber.westpoint.edu/definition-attack-law-of-armed-conflict-protection/> accessed 8 January 2023

48 *Ibid*.

ICRC's study on customary international humanitarian law uses the term "military operations" when referring to cultural objects and their protection. Consequently, this broadened perspective would provide a more extensive and precise comprehension of the protection conferred than the analysis presented in Additional Protocol I's commentary.[49] This interpretation is further expanded through functional analysis, asserting that the ultimate impact on medical units, materiel, or cultural objects—regardless of whether they are destroyed, pillaged, or rendered unusable—remains the same, as these losses are often irreversible. Therefore, a functional understanding of 'attack,' encompassing all military operations, could yield more comprehensive protection without taking into account the *de facto* control of the objects or territory.[50]

While this broad understanding of 'attack' may have merit in analyzing the nuances of the term, particularly with regard to medical and cultural objects, it is essential to underscore that international humanitarian law must strike a delicate balance between humanitarian imperatives and military necessity. Whether this broader interpretation of 'attack' and the constraints it may impose on warfare are adopted or adhered to by armed forces in practice remains a subject of conjecture.

Of the two frameworks for understanding 'attack'—the narrow definition utilized in Article 49 of Additional Protocol I and the broad interpretation discussed above—the thesis adopts the former. While arguments in favor of the broad understanding present compelling reasons for its adoption, the narrow definition aligns more closely with the delicate equilibrium between military necessity and humanitarian considerations. The terminology employed by the two frameworks offers an illustrative contrast: the narrow definition employs "combat actions," whereas the broad interpretation extends the term to encompass all military operations. Linguistically, this extension appears unwarranted, as it fails to account for the myriad military operations that cannot reasonably be classified as attacks. Furthermore, adopting the narrow understanding does not compromise the protection already afforded to medical and cultural objects under the existing framework of the Conventions and Additional Protocols, rendering an extension of the definition unnecessary.

3.2.2 Armed Attack *Jus ad Bellum*

According to Article 51 of the UN Charter, an armed attack is the sole condition on which a state may resort to armed force against another. The right to self-defense lasts as long as the armed attack lasts or

49 *Ibid*.
50 *Ibid*.

until the Security Council takes the necessary measures to maintain international peace and security. The Charter itself does not go into the definition of what constitutes an armed attack.

Ruys points out that Article 51 does not exist in a vacuum but is a constituent part of a more general regime of articles and must be read in that context.[51] The other articles of this regime are Articles 2(4), 39, 42, and 53. Taken together, these articles establish an international order that puts a ban on unilateral use of force and creates an institution in the Security Council enshrined with powers to uphold and enforce international peace. In this context, Ruys argues that Article 51 is more descriptive of an extraordinary state of things than a settled customary law norm.[52] Consequently, states' right to unilateral self-defense would become not a right states would have naturally, but a right enshrined for a specific, abnormal situation, that being the use of force by another state.

Armed attack can be viewed as having three aspects: *ratione materiae*, or what constitutes an armed attack; *ratione temporis*, or when does an armed attack happen; and *ratione personae*, or from whom does the armed attack emanate.[53] When r*atione materiae* is examined through the lens of the jurisprudence of the ICJ in the *Nicaragua* case, a twin regime can be observed. Within this regime, the use of force between states is divided into "the most grave forms of use of force" and "other less grave forms."[54] The ICJ reaffirms its stance on the matter by distinguishing the sending of armed bands of fighters into the territory of another state from 'mere frontier incidents'. The distinguishing factor is the scale and effects of the operation.[55] The Court reaffirmed its ruling with direct reference to the framework established in *Nicaragua* in *Oil Platforms.*[56] The ICJ's decision has been widely criticized for setting a threshold for military activity states' having the extraordinary right to answer with force.[57]

The ICJ is not alone in advocating for a threshold of intensity that must be crossed for mere border incidents to become armed attacks. The UN General Assembly reaffirmed this in Resolution 3314 on the definition of aggression by stating that "[acts of aggression] must be considered in the light of all

51 Tom Ruys, Armed Attack and Article 51 of the UN Charter: Evolutions in Customary Law and Practice (Cambridge University Press 2013) 59
52 *Ibid*.
53 *Ibid.* 126
54 *Case Concerning Military and Paramilitary Activities in and Against Nicaragua (Nicaragua v. United States of America)*, (Merits, Judgment, 27 June 1986), ICJ Reports 1986 (*Nicaragua*) 101
55 *Ibid*. 103
56 *Case Concerning Oil Platforms (Islamic Republic of Iran v. United States of America)*, (Merits, Judgment, 6 November 2003), ICJ Reports 2003 (*Oil Platforms*) 186
57 Tom Ruys (n 50) 147

circumstances of each particular case"[58] and that "the Security Council may...conclude...that the acts concerned, or their consequences, are not of sufficient gravity."[59] Nevertheless, the Resolution does offer a non-exhaustive list of *ratione materiae* that qualify *prima facie* as aggression, including attacks by armed forces, blockades of ports and coasts, and letting an aggressor use the state's territory to attack a third state.[60] Through the official statements of states, there is no clear indication where the threshold of an armed attack lies; however, it is clear that the gravity measurement is contingent more on the effects of the act than the act itself.[61]

It is clear from the wording of the UN Charter that an armed attack is a specific form of illegal use of force and that not all uses of force entitle the victim state to use coercive countermeasures. Through customary practice, it is clear that even limited use of military power can be counted as an armed attack, provided that the attack results in or is liable to result in death and destruction.[62] Possible overreaches in retaliating are accounted for by the necessary and proportionate qualifiers it must stay within. The needle-prick theory, which some scholars subscribe to, posits that even smaller scale incursions can collectively trigger the right to self-defense.[63] Furthermore, the military actions of a state must have a certain *animus aggressionis*, or hostile intent, driving them.[64] This intent is always context-sensitive. The requirement for *animus aggressionis* is very sensible since that diminishes the possibility of an accident spiraling out of hand into a wider war.

Armed attack *ratione temporis*, or the question of when does an armed attack occur, has a *prima facie* simple answer. An attack starts with the first pull of the trigger or similar commencement of violent military operations and ends when these operations cease, and the right to self-defense subsists for a similar period. Arguments wholly against such a reading are not many but warrant analysis. The crux of the arguments for so-called 'anticipatory self-defense' rely on the existence of nuclear weapons and the potentiality of a first strike so devastating that waiting for it would pragmatically speaking deny the target nation's right to defend itself. Similarly, one can argue for anticipatory self-defense in cases where an armed attack has not yet commenced but is in its preparatory stages, and the only fighting

58 UNGA Res 3314 (XXIX) (14 December 1974) Annex
59 *Ibid*. art. 2
60 *Ibid.* art. 3
61 Laurie R. Blank, 'Irreconcilable Differences: The Thresholds for Armed Attack and International Armed Conflict' (2020) 96 Notre Dame L Rev 255
62 Tom Ruys (n 50) 155
63 *Ibid*. 168
64 *Ibid.* 177

chance the target state has is if it strikes first, not dissimilar to the Six-Day War between Israel and the Arab Coalition. Anticipatory self-defense can be further divided into pre-emptive and preventive self-defense, depending on whether or not an armed attack is thought imminent or not.[65]

An argument most often cited by those advocating for the right to anticipatory self-defense is the Webster formula. It states that in cases where the necessity for self-defense is "instant, overwhelming, leaving no choice of means, and no moment for deliberation."[66] Should the Webster formula be accepted as reflecting customary law, one cannot draw other conclusions than that the states' right to defend themselves is not limited to *a posteriori* the connection of the possible attack that a strict reading of the UN Charter would imply. Those arguing against such a reading argue that using the Webster formula is anachronistic and misguided based on the facts of the *Caroline* case.[67] Furthermore, whatever customary law norm might have preceded the UN Charter on the matter of self-defense got changed with the global accession to it, thus becoming reflective of the new *status quo,* viz. anticipatory self-defense. However, the limits this puts on armed responses to imminent armed attacks are not ironclad.

During the run-up to the Iraq War, the doctrine of anticipatory self-defense suffered yet more blowbacks. The so-called Bush doctrine offered by the United States to justify the invasion had its basis in Saddam's Iraq having the capabilities to produce weapons of mass destruction and its possible willingness to use them at some indeterminate point in the future. This was firmly rebutted as having no basis in international law by multiple academics, with Australian academics calling it a 'contradiction of the prohibition of the unilateral use of force.[68] Analyses by academics conducted after Operation Iraqi Freedom nearly unilaterally concluded that the Bush doctrine was overbroad and self-defense could not be resorted to against non-imminent threats.[69] The overarching consensus points to the situation where the imminence of the threat is a key factor in determining whether or not a state can resort to force and legitimately claim self-defense. How imminence is calculated has many forms, one of which is the Bethlehem principles. The Bethlehem principles underline that the imminence calculation must include all relevant circumstances surrounding the armed activity leading to the possible armed attack. The circumstances include the nature and immediacy of the possible armed

65 *Ibid*. 253-254
66 Yale Law School, 'The Caroline' (Yale Avalon Project, 19 April 2009) <https://avalon.law.yale.edu/19th_century/br-1842d.asp#ash1> accessed 20 November 2023
67 Tom Ruys (n 50) 258-289
68 *Ibid*. 322-323
69 *Ibid*.

attack; how probable the armed attack is; whether or not there exists a pattern of previous armed activity; the scale of the armed attack and the damage it could cause without an armed answer; and the possibility of mitigating the resulting damage by seizing the opportunity and launching an attack of one's own.[70] No matter what heuristic a state uses to assess the immediacy of an armed attack warranting anticipatory self-defense, it is clear that the assessment of the facts must have its basis in good faith and sound evidence.[71]

Armed attack *ratione personae*, or who can launch armed attacks, is also a sphere of armed attack that has a simple and clean answer that has been overtaken by time and changing circumstances. The clearest answer to the question is the armed forces of the state. For a long time, this would have been an acceptable endpoint of the inquiry. However, in some cases, like *Nicaragua,* have widened this understanding to state-backed armed groups other than the official military.[72] The state-backed armed groups have to have *de facto* dependence on the state for their conduct to be attributable to the state.[73][74]

September 11 terrorist attacks, by the latest, would extend armed attack *ratione personae* to include armed groups not backed by a state. However, the state practice on the matter had already started to change with the United States adopting the Shultz doctrine, which foresaw a possible future in which an armed response would be aimed against a non-state armed group in the territory of another state.[75] In the aftermath of the September 11 terrorist attacks, the Security Council adopted resolutions 1368 and 1373, which both named international terrorism as a threat to international peace and security[76] and imposed obligations on all states to combat, or at least not aid and abet, terrorist organizations.[77] The unprecedented situation in which the Security Council had to react to the actions of a non-state actor in a similar fashion as that of a state solidifies the fact that armed attacks can be launched by other actors than states alone. Almost unanimous approval of Operation Enduring Freedom, the United States and United Kingdom's invasion of Afghanistan to root out those responsible for the terrorist attacks of September 11, gives further evidence that attacks by non-state actors can constitute armed attacks,

70 George Brandis, 'The Right of Self-Defense against Imminent Armed Attack in International Law' (2017) 35 Aust YBIL 60
71 *Ibid*. 62
72 *Nicaragua* (n 53) 103
73 *Ibid*. 62
74 *Case Concerning Application of the Convention on the Prevention and Punishment of the Crime of the Crime of Genocide (Bosnia and Herzegovina v. Serbia and Montenegro)*, (Merits, Judgment, 26 February 2007), ICJ Reports 2007 205
75 Tom Ruys (n 50) 422-423
76 UNSC Res 1368 (2001) UN Doc S/RES/1368
77 UNSC Res 1373 (2001) UN Doc S/RES/1373

triggering the state's right to self-defense.[78] The reasoning behind this conclusion is that the premise of the operation was that the United States had become the target of an armed attack and that they and the United Kingdom were practicing collective self-defense sanctioned by Article 51 of the UN Charter.

3.3 Protection Granted to Objects

Following the delineation of objects and attacks, the inquiry shifts towards elucidating the actual protection provided by international humanitarian law concerning these objects in relation to attacks. Initially, one might be tempted to assert that non-military objectives should not be subjected to attacks, and indeed, this is a fundamental premise. However, the protection afforded by international humanitarian law entails nuanced considerations that extend beyond merely prohibiting direct attacks on non-military objectives. Direct attacks are not the only way to endanger non-combatants; therefore, international humanitarian law needs to account for that. For example, the armed forces on one side can directly kill civilians, or they could cause deaths within the civilian population via indirect means such as the destruction of water treatment infrastructure. The spirit of the law cannot stand for allowing these indirect offenses.

If the protection from attacks were interpreted as an absolute proscription of military operations in the vicinity of non-combatants, this would be viewed as overly restrictive by armed forces worldwide. Such an approach could risk rendering the law, and consequently the protection it offers, ineffective due to non-compliance. This underscores the delicate balance between military necessity and humanitarian concerns, a balance that is intrinsic to international humanitarian law.

This equilibrium is achieved through several core principles of international humanitarian law, which include the principle of distinction between combatants and non-combatants, the principle of necessity in conducting military actions, and the principle of proportionality in the selection of weapons for an attack. While these principles are not explicitly named, they find their basis in Article 57 of Additional Protocol I.

3.3.1 Principle of Distinction

Article 57 of Additional Protocol I encompasses a set of obligations for individuals involved in planning or deciding upon attacks, and one of its pivotal provisions stipulates that they must take all

78 Tom Ruys (n 50) 436

feasible measures to distinguish between civilians or other protected classes as defined by the Protocol and legitimate military targets. While this places a substantial burden on those orchestrating attacks, it is essential to note that these obligations are not contingent on the rank or position of the individuals within the attacking party. The phrase "plan or decide upon an attack" extends this obligation to all personnel involved, irrespective of their role or status. Even individual soldiers, when faced with the micro-level decision to engage in combat, bear the responsibility to differentiate between combatants and non-combatants. Although this interpretation is not explicitly endorsed by the Commentary of 1987 to the Protocol, it recognizes the need to involve lower-ranking commanders in the field.[79] Consequently, every combatant participating in combat actions can be considered bound by the principles of international humanitarian law.

The Commentary to the Protocol provides further insights into the responsibilities of military commanders and individual soldiers concerning target identification. It emphasizes that the Article necessitates not merely acknowledging the need for target differentiation but also verifying the nature of the intended target and eliminating any doubts before initiating an attack.[80] The term "everything feasible" is interpreted to encompass all practical means of ascertaining information about the target, contingent upon the specific circumstances at the time and place of the attack. This includes considering the impact of information gathering on the success of the military operation.[81] While incorporating the impact on military success as a factor in determining the feasibility of information gathering might raise concerns, it is likely intended to strike a balance between military necessity and humanitarian considerations.

An integral aspect of the principle of distinction revolves around the exceptional cases in which civilians or civilian objects may lose their protection against attacks.[82] Although such circumstances are rare and regarded as extraordinary deviations from the norms of international humanitarian law, they can render attacks on non-combatants and civilian objects legal. The only circumstance recognized in Additional Protocol I is when civilians take a direct part in hostilities.[83] If captured while participating

79 International Committee of the Red Cross, 'Precautions in Attack' (Commentary of 1987, 1 June 2023) <https://ihl-databases.icrc.org/en/ihl-treaties/api-1977/article-57/commentary/1987?activeTab=undefined> accessed 15 April 2024

80 *Ibid*.

81 *Ibid.*

82 Salem Aessa Farhat and others, 'Attacks Against Civilian Objects: An Analysis Under International Humanitarian Law' [2022] 8(1) Hasanuddin Law Review 65

83 Additional Protocol I art 51

in hostilities, they do not revert to civilian status but instead become prisoners of war.[84] It's crucial to note that the loss of protection under these circumstances is temporary and applies only as long as direct participation persists.

3.3.2 Principle of Proportionality

The principle of proportionality plays a vital role in determining the methods and means employed by belligerent parties at various stages of a conflict. Although not explicitly named in Additional Protocol I, the Protocol provides content for this principle. Article 57 of the Protocol emphasizes that belligerents must exercise caution when selecting methods or means of attack to minimize the risk to non-combatants and civilian objects. The language of the Article carefully navigates the delicate balance between military necessity and humanitarian concerns. It stipulates that belligerents should aim to "avoiding, and in any event minimizing [civilian casualties]," recognizing that complete elimination of incidental civilian damage is often unattainable.

The Commentary to the Protocol underscores that this principle primarily concerns the methods and means of warfare.[85] For instance, it illustrates the principle using the choice of bombs as an example. It argues that there is no need to use a 10-ton bomb when a smaller one can effectively achieve the destruction of the target. However, the assessment of what method or means is proportionate to a specific objective is a complex endeavor. The battlefield is often characterized by confusion and the "fog of war," where accurate reconnaissance may be compromised due to hostile airspace or other factors. The Commentary acknowledges these challenges, which commanders must grapple with in real-time decision-making.[86] Another example it provides pertains to minefields, booby traps, and other devices that remain active long after the conflict's conclusion. While no specific bans or obligations are imposed, the Commentary suggests measures such as mapping minefields and employing timed mechanisms for self-defusing these devices.[87] Analogously, in the realm of cyber warfare, logic bombs (pieces of code designed to activate upon specific commands) can be likened to booby traps. If logic bombs lack self-deletion or remote deletion features, they may persist in infected systems, potentially causing harm and re-activating hostilities should the damage suffice, even after the conflict has ended.

84 *Ibid.*
85 Precautions in Attack (n 78) 2200
86 *Ibid.*
87 *Ibid.*

However, this area is still relatively uncharted due to its novelty, and further investigation exceeds the scope of this thesis.

Regarding "feasible precautions," the Commentary defers to the 1980 Convention of Prohibitions or Restrictions on the Use of Certain Conventional Weapons (CCW).[88] While the CCW does not provide an exhaustive list of precautions, it offers general guidelines. According to these guidelines, "feasible precautions" encompass all measures that are practical or practically possible, taking into account the circumstances at the time.[89] Although in this case the CCW specifically pertains to minefields, this framework can be applied more broadly when evaluating the proportionality of methods and means chosen by battlefield commanders. It ties the practicability of precautions to the moment when the decision to employ a specific method or means of warfare is made, allowing commanders flexibility in their choices while providing a clear framework for *post-facto* analysis.

While it may seem that the principle of proportionality permits belligerents to disregard some civilian casualties during combat operations, this is not the case. The principle acknowledges that, given the multitude of factors influencing the danger posed by a particular combat operation to civilians, it is virtually impossible to entirely eliminate incidental danger.[90] The calculation that military commanders must undertake does not involve calculating the actual damage caused to civilians or civilian objects but rather the risk of such damage occurring. Commanders are expected to engage in this calculation in good faith, employing common sense.[91] An attack that fails to adhere to this standard is deemed indiscriminate and is prohibited, constituting a grave breach of the Protocol and a war crime.[92]

When assessing attacks, the Commentary recognizes that there is a significant difference between individual attacks aimed at specific targets and theater-wide attacks targeting multiple objectives over a broad front. Both types of operations should be considered as individual attacks. However, the assessment of incidental danger to civilians is more object-based than attack-based. In other words, in broad attacks involving multiple objectives, the determination of incidental danger to civilians must be conducted separately for each military objective.[93]

---

88 *Ibid.*
89 Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons which may be deemed to be Excessively Injurious or to have Indiscriminate Effects (adopted 10 October 1980 entered into force 2 December 1983) 1342 UNTS 22495 (CCW) art 3(4)
90 Precautions in Attack (n 78)
91 *Ibid.*
92 *Ibid.*
93 *Ibid.*

Sumanadasa highlights the nebulous nature of the principle of proportionality when applied on the battlefield.[94] There are no precise guidelines on how a military commander should determine whether a specific response is proportionate in a given situation. The only examples given are of two very extreme cases of the spectrum, such as shooting at a sniper's nest versus aerial bombardment of a building.[95] The actual decision is left to the individual judgment of the commander. Moreover, Sumanadasa identifies two distinct aspects of the principle: a protective aspect, aimed at safeguarding combatants and non-combatants from excessive use of force, and a defensive aspect, used to justify military operations when they offer concrete military advantages.[96] Consequently, the principle can be invoked both in support of and against the legality of military actions taken by belligerent parties.

It is crucial to note that the principles of international humanitarian law are interconnected rather than isolated. The principle of proportionality is linked to the principle of distinction by enhancing protection for civilians through restrictions on the use of force against legitimate targets. Additionally, proportionality connects with the principle of military necessity by requiring a judgment based on the balance between incidental damage and necessary military advantage.[97]

3.3.3 Principle of Military Necessity

The principle of necessity in international humanitarian law stipulates that, when selecting military objectives for combat operations, the one posing the least danger to civilians and civilian objects should be chosen.[98] This principle operates on the assumption that, in situations where prospective targets offer an equal military advantage, the selection should prioritize minimizing risks to non-combatants and their property. The Commentary of Additional Protocol I provides limited commentary on this principle, offering examples such as roads and railroads that can be targeted in a manner that hampers enemy logistics without endangering civilians.

The principle of necessity is considered a customary rule of international humanitarian law, as it has been incorporated into the military manuals of multiple states and is supported by official statements

94 W.A.D.J. Sumanadasa, 'Principle of Proportionality: The Criticized Comprising Formula of International Humanitarian Law' (2010) 10 ISIL YB Int'l Human & Refugee L 26
95 *Ibid*. 27
96 *Ibid*. 33
97 Salem Aessa Farhat and others (n 81) 66-67
98 Additional Protocol I art 57(3)

affirming its obligation.[99] This recognition extends to states that were not parties to Additional Protocol I when it was adopted, reinforcing the notion that the principle is part of customary international humanitarian law. However, it's worth noting that the interpretation of the principle may vary among states. For instance, the United States asserts that the principle is not an absolute obligation but allows for the analysis of commanders concerning its feasibility in relation to mission completion and acceptable levels of risk.[100]

Hayashi identifies three distinct contexts in which military necessity is relevant: material, normative, and juridical.[101] Material military necessity pertains to the amoral necessity dictated by battlefield conditions, where goals, means, and circumstances intersect.[102] Normative military necessity assesses whether the evil demanded by material military necessity is legitimate and whether it should be permitted or prohibited.[103] This evaluation considers the balance between the means (e.g., killing enemy soldiers) employed and the purpose (depleting the military power of the enemy) for their use, determining the legitimacy of material military necessity. Additional factors, such as whether or not enemy combatants are surrendering, can influence this evaluation.[104] Juridical military necessity, military necessity in context of valid norms,[105] involves judging whether a conduct aligns with the norms of international humanitarian law and is therefore allowed or prohibited. This assessment is based on the existence of conduct necessary to achieve a purpose that conforms to international humanitarian law norms.[106]

In essence, the principle of necessity underscores the importance of selecting military objectives with the least harm to civilians and their property, particularly when comparable military advantages can be achieved through alternative targets. Additionally, it limits the gamut of operations available to the actor to only those necessitated by operational realities. This principle serves as a crucial component of international humanitarian law, striking a balance between military requirements and humanitarian concerns on the battlefield.

---

99 International Committee of the Red Cross, 'Target Selection' (Customary International Humanitarian Law, 28 January 2023) <https://ihl-databases.icrc.org/en/customary-ihl/v1/rule21> accessed 15 April 2024
100 *Ibid*.
101 Nobuo Hayashi, 'Contextualizing Military Necessity' (2013) 27 Emory Int'l L Rev 193
102 *Ibid*. 195
103 *Ibid.* 223
104 *Ibid*. 234-237
105 *Ibid.* 254
106 *Ibid*. 254-256

4. Computer Data

The investigation into the constraints imposed by international humanitarian law on the conduct of cyber warfare naturally begins with an examination of the primary subject of cyber warfare: computer data. In this chapter, the focus will be on defining computer data and distinguishing between different forms of data, such as content data and software. The analysis delves into the micro-level characteristics of computer data and juxtaposes them with the relevant norms in international humanitarian law, particularly those pertaining to objects, with a special emphasis on civilian objects. This inquiry aims to determine whether computer data, in and of itself, can be considered an object under the purview of international humanitarian law.

The second part of this chapter explores the concept of data as an integral component of larger systems. This encompasses various contexts, such as data within banking systems, medical facilities, and industrial processes. The examination considers the analogous nature of data and assesses the consequences of targeting data within these systems compared to physically destroying the systems themselves. This analysis is guided by the overarching spirit of international humanitarian law, which considers the effects and repercussions of damaging or threatening data within these systems. Specific focus is placed on several critical systems, including medical computers and data, data essential to agricultural and water supply systems, and data as part of objects indispensable for the survival of civilian populations.

Furthermore, this chapter investigates the positions of three states—Finland, the United Kingdom, and the United States—with regard to data and its classification as an object under international humanitarian law. This exploration seeks to gain insights into the extent to which these states are willing to adhere to international legal standards. It is crucial to acknowledge that the foundation of all international law rests on the consent and willingness of states to abide by its provisions.

4.1 ‘Objecthood’ of Computer Data

When initiating an inquiry into the concept of "objecthood" as it relates to an object and the potential implications thereof, it is logical to commence by examining the substance of the object in question. In this context, the substance of an object refers to what constitutes the object in the physical world,

distinguishing it from other coexisting entities. Understanding this substance enables the comparison of the object against different sets of criteria for various purposes. In the context of this thesis, the focus is on scrutinizing the substance of computer data in relation to the understanding of an object within the framework of international humanitarian law.

Computers and computer systems primarily utilize binary code, a system of representing information using two digits: 0 and 1. These indivisible elements, known as bits, form the fundamental building blocks of the binary computing system. Nearly everything stored on a computer or computer system is encoded in bits or strings of bits, such as bytes (comprising eight bits). Although theoretical models of quantum computing offer exceptions to this rule,[107] these exceptions remain largely theoretical, with quantum computers not yet in widespread use. Therefore, this thesis concentrates on digital data stored in binary form.

At its most basic level, computer data consists of individual values represented as bits. These bits only transform into usable data entities when processed by a computer.[108] These data entities can take a multitude of forms, ranging from text to video and audio. Data can also manifest as sets of machine instructions, known as software. While software is itself a form of digital data, it differs significantly from non-software data, especially in its purpose and function. This distinction is important for several reasons. One such reason is that software is meant for the machine to interpret; content data is meant to be viewed by the user of the machine. The differences between content data and software are numerous. First, software is designed primarily to serve as a tool for specific tasks, whereas non-software data typically represents empirical information.[109] Second, software has the capacity to be executed, meaning that the instructions contained within it can be run by a computer.[110] Data cannot be run on a computer without the requisite software. Third, software often relies on non-software data to function effectively.[111] Fourth, software is subject to copyright protection, unlike most non-software data. The difference in this is that software is the result of a creative process, whereas data is most often the result of an empirical one.[112] Fifth, while both software and data can experience a phenomenon

107 IBM, 'Introduction' (Quantum Computing, 5 December 2023) <https://learning.quantum.ibm.com/tutorial/explore-gates-and-circuits-with-the-quantum-composer> accessed 15 April 2024

108 Neil Selwyn, 'Data entry: towards the critical study of digital data and education' [2015] 40(1) Learning, Media and Technology 65

109 Daniel S. Katz, 'Software vs Data' (Software vs Data Open Article, 9 December 2016) <https://github.com/danielskatz/software-vs-data> accessed 22 February 2023

110 *Ibid.*

111 *Ibid.*

112 *Ibid.*

known as "bit rot,"[113] which is the degradation of data or software due to changes in storage conditions or the environment, the effects are the result of different processes such as degradation of storage medium or support for certain data formats.[114] Data is more susceptible to changes in the former,[115] whereas the latter more often affects software.[116] Lastly, the lifespan of software is typically determined by its utility, whereas data can persist indefinitely as long as its storage medium remains intact.

To further differentiate software from data, it is helpful to categorize them as operational data (software) and content data (non-software).[117] Operational data, such as software, serves as a tool for specific functions and tasks, whereas content data represents information and evidence of various phenomena. The divide is very similar to that between software and non-software.

The physical manifestation of bits varies depending on the storage medium, such as holes in punch cards, changes in magnetic charge on cassette tapes, or differing electric charges in transistors in modern electronic systems. In information theory, information, of which bits are the smallest measurable units, can be linked to thermodynamics through the Landauer principle. This principle suggests that the destruction of information necessitates the dissipation of heat, implying a physical connection between information and energy.[118]

Therefore, information, including computer data as a subset, possesses physical attributes and is intertwined with the material world. The Landauer principle even provides an estimate of the mass of hypothetical information particles, which, while likely lighter than observable particles, indicates that information and data have physical relevance.[119] This understanding has significant implications when viewed through the lens of international humanitarian law.

The primary consequence of this perspective is that computer data should be considered an object in its own right. This conclusion aligns with one of the key requirements for an object as stated in the Commentary to Additional Protocol I, which specifies that an object is "something that can be

113 *Ibid.*

114 Vinton G. Cerf, 'Avoiding "Bit Rot": Long-Term Preservation of Digital Information' [2011] 99(6) Proceedings of the IEEE 915-916

115 Karol Król and Dariusz Zdonek, 'Peculiarity of the bit rot and link rot phenomena' [2020] 69(1) Global Knowledge, Memory and Communication 21

116 *Ibid.* 22

117 Heather A. Harrison Dinniss, 'The Nature of Objects: Targeting Networks and the Challenge of Defining Cyber Military Objectives' (2015) 48 Isr L Rev 39

118 Rolf Landauer, 'Irreversibility and Heat Generation in the Computing Process' [1961] 5(3) IBM Journal of Research and Development 186

119 Edward Bormashenko, 'The Landauer Principle: Re-Formulation of the Second Thermodynamics Law or a Step to Great Unification?' [2019] 21(10) Entropy 918

perceived; a material thing."[120] This implies that, to qualify as an object, a thing must possess substance and be perceptible in some form. As demonstrated earlier, information and computer data meet this requirement as they possess physical substance and exist within the realm of physical reality.

Hence, it is reasonable to assert that computer data, by virtue of its physical properties and substance, can be classified as an object within the norms of international humanitarian law. Consequently, computer data should be entitled to the protection afforded to objects, shielding it from the adverse effects of armed conflicts. Classifying computer data as an object would not hamper the combatants' ability to attack military targets, including data, as international humanitarian law already divides objects in realspace into military and non-military objects. Nevertheless, it is essential to acknowledge that, within man-made systems and legal frameworks, definitions do not always strictly adhere to physical reality, especially when dealing with concepts rooted in micro-particles and theoretical physics.

4.1.1 Tallinn Manual and Arguments against Interpreting Data as an Object

The Tallinn Manual, a comprehensive document that provides an extensive interpretation of international law in the context of cyber activities, adopts a predominantly effects-based perspective. This perspective is notably evident in Rule 92, which defines a cyber attack as a combat operation aimed at causing injury or death to individuals or damage or destruction of material objects. Although the Manual explicitly acknowledges that cyber operations targeting computer data, categorized as non-physical entities, are not excluded from its scope, it insists that a real-world effect of injury, damage, death, or destruction must occur for an operation to qualify as an attack and fall under the jurisdiction of international humanitarian law.[121] Cyber operations lacking this requisite effect are not automatically categorized as attacks under the framework provided by the Manual.

It is essential to recognize that the Tallinn Manual's definitions are not rigidly fixed. For instance, the Manual's definition of injury to personnel encompasses illness and severe mental distress,[122] while its interpretation of damage and destruction of objects is effects-driven. Temporary disruptions to an

120 International Committee of the Red Cross, 'General Protection of Civilian Objects' (Commentary of 1987, 16 March 2023) <https://ihl-databases.icrc.org/en/ihl-treaties/api-1977/article-52/commentary/1987?activeTab=undefined> accessed 15 April 2024

121 Michael N. Schmitt, Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations (2nd edn, Cambridge University Press 2017) 416

122 *Ibid.* 417

object's functionality through cyber means can constitute an attack if restoring the object's functionality necessitates replacing physical components. Some experts even argue that operations targeting data within a system, such as the operating system, that compromise its functionality until data is restored should be considered attacks. However, this perspective underscores the importance of the real-world effects of an attack on data, not the attack on data itself. In cases where an attack causes other disturbances but primarily targets data, the Manual refers to alternative legal frameworks, such as the prohibition of collective punishment.[123] However, in such cases, the prohibited conduct is the event caused by targeting data, not the attack on data itself.

Nevertheless, the distinction between physical and non-physical objects remains a foundational aspect of the Manual's approach, which is open to criticism.[124] This distinction relies on an outdated definition of objects found in the Commentary of 1987 to Additional Protocol I and is susceptible to arbitrariness. This issue raises questions about whether current definitions and delineations within the Manual are internally consistent, thus highlighting the need for *lege ferenda* to address these concerns.

The Tallinn Manual explicitly states that attacks on data as standalone entities are not considered attacks because the term "object" in the Manual does not encompass computer data according to its "ordinary meaning," as defined in the Commentary of 1987 to Additional Protocol I.[125] The concept of ordinary meaning originates from the Vienna Convention on the Law of Treaties. The article it originates from states that interpretation of the terms of the treaty must follow the ordinary meaning in its context and in light of the object and purpose of the treaty.[126] The parties are free to agree on the meaning of the terms other than what would be their ordinary meaning. The ordinary meaning can also be disregarded should it lead to a “manifestly absurd or unreasonable” result.[127] In the International Law Commission’s (ILC) Third Report on the Law of Treaties, special rapporteur Sir Humphrey Waldock presents draft articles on the application, effects, revision, and interpretation of treaties. The general rule of interpretation given is that terms of a treaty must follow the natural and ordinary meaning in the whole context of the treaty and in the context of international law at the time of the conclusion of the treaty.[128] The interpretation can rely on preparatory works, the circumstances of the conclusion, and the subsequent practice of the parties when confirming the natural and ordinary

123 *Ibid.* 418
124 *Ibid.* 437
125 *Ibid.*
126 Vienna Convention on the Law of Treaties (adopted 23 May 1969 entered into force 27 January 1980) 1155 UNTS 18232 (VCLT) art 31
127 *Ibid.* art 32(b)

meaning of the terms.[129] The interpretation of treaties is not frozen in time. While the context of international law at the time of the conclusion of the treaty is the general rule, changing times and new norms of international law have an influence on how terms of a treaty are interpreted.[130]

Some experts within the Manual's group argued for the inclusion of critical datasets, such as tax records, social security data, and bank accounts, as objects under international humanitarian law. Their argument is that leaving critical datasets unprotected would be against the spirit of international humanitarian law.[131] This view, however, is not reflected in *lege lata*, the existing legal framework.

An exception to this exclusion of data as an object under international humanitarian law is evident in cases involving cultural property. Rule 142 of the Manual emphasizes the importance of respecting and protecting cultural property susceptible to cyber operations or located in cyberspace. This exception implies that cultural property, even if composed entirely or partially of data, can be regarded as an object. Some Tallinn experts expressed the opinion that other forms of intangible protected objects, such as intellectual property, exist.[132] However, this exemption is limited to cultural property existing solely in digital form or limited digital copies of cultural property that is at risk of disappearing.[133]

While there is a prevailing view, at least in *lege lata*, that international humanitarian law does not encompass computer data as objects, notable scholars like Schmitt acknowledge the limitations of this perspective. They recognize the potential under-inclusiveness of this view and the evolving nature of the concept of objects. However, they also argue that a blanket inclusion of all data as objects could lead to over-inclusiveness, and until consensus is reached among states, a cautious approach is advisable.[134] Schmitt draws an analogy between signal jamming and DDoS attacks, which both affect the means of communication rather than the message.[135] Similar methods can be used in psychological

128 ILC, 'Third Report on the Law of Treaties, by Sir Humphrey Waldoc, Special Rapporteur' (1964), UN Doc A/CN.4/167 and Add.1-3 art 70(1)
129 *Ibid.* art 71(2)(a)
130 *Ibid.* art 72
131 Tallinn Manual (n 120) 417
132 *Ibid.* 535
133 *Ibid*.
134 Michael N. Schmitt, 'Notion of Objects during Cyber Operations: A Risposte in Defence of Interpretive and Applicative Precision' (2015) 48 Isr L Rev 84
135 *Ibid.* 93-94

operations, with similar results.[136] Schmitt agrees that there are cases in which the current *lege lata* is not sufficient; however, many of these cases can be solved without giving data the status of an object.[137]

Schmitt further emphasizes the importance of assessing the effects of cyber operations when determining their legality. He contends that the question of whether data is considered an object is secondary to evaluating whether a cyber operation has the consequences of causing injury, death, or significant damage, as these are the only permissible targets of an attack.[138] If a cyber operation causes these effects, it is considered an attack, regardless of whether data is an object. Attacks on data are permissible if the system containing it or its functionality are not explicit targets of the attack.[139] Schmitt's analysis aligns with the cautious approach he advocates for under existing international law.

McCormack notes that there is no consensus among states regarding whether computer data can be considered objects under international humanitarian law.[140] He suggests that the determination of whether data is an object remains within the purview of states, but he also highlights that the need for such specification is limited, given that most large-scale cyber operations exclusively targeting data fall into categories such as data exfiltration or ransomware attacks, which do not directly damage the data.[141]

Pomson, in his analysis, emphasizes the ordinary meaning of the word 'object,' contending that an object must be tangible and touchable, as per one of the definitions in the Oxford English dictionary. He explicitly disregards all other possible meanings of the word.[142] However, this strict interpretation disregards the existence of intangible entities, such as gases, which are considered objects despite being imperceptible to the naked human senses alone.

The United States adopts an ambiguous stance, considering the consequences of a cyber operation to assess whether it can be considered an attack under international humanitarian law. The consequences of an operation need to include either injury, death, significant damage, or the threat thereof to be read as attacks.[143] This stance implies that computer data is not inherently an object but can become one

136 Michael N. Schmitt, 'Wired Warfare 3.0: Protecting the Civilian Population during Cyber Operations' (2019) 101 Int'l Rev Red Cross 342
137 Michael N. Schmitt (n 133) 98, 100
138 *Ibid*. 101
139 *Ibid.*
140 Tim McCormack, 'International Humanitarian Law and the Targeting of Data' (2018) 94 Int'l L Stud Ser US Naval War Col 239-240
141 *Ibid*. 234, 236
142 Ori Pomson, 'Objects'? The Legal Status of Computer Data under International Humanitarian Law (2023) 11-12
143 UNGA Res 76/136 (13 July 2021) UN Doc A/76/136 137

when the operation produces certain effects.[144] The United Kingdom's approach is similar, focusing on the consequences of a cyber operation to determine its status as an attack, with a clear stance that only operations with analogous consequences to kinetic attacks can be considered attacks.[145] This stance also implies that computer data is not inherently an object within the framework of international humanitarian law.

4.1.2 Arguments for Interpreting Data as an Object

Schmitt's influence on the discourse surrounding the "objecthood" of data is unquestionable, but it is crucial to acknowledge that he is not the sole voice in this discussion. Numerous academics have contributed their perspectives to this complex matter. One of the prominent critics of the consensus put forth in the Tallinn Manual is Mačák, who offers an alternative viewpoint. Mačák's analysis focuses on the context and purpose of the treaty rather than solely relying on the ordinary meaning of terms, which is the approach adopted by the Tallinn Manual and Schmitt. The analysis is seated in the rationale within Article 31(1) of the Vienna Convention on the Law of Treaties (VCLT).[146]

Mačák contends that the requirement of objects being "visible and tangible," as suggested in the Commentary of 1987 to Additional Protocol I, was originally intended to distinguish objects from abstract goals of military operations rather than limiting the status of objects.[147] He argues against the strict interpretation of treaties based solely on the ordinary meanings of words used at the time of drafting. Mačák asserts that this approach is overly broad and points out that both the International Law Commission (ILC) and the International Court of Justice (ICJ) have evolved their interpretations over time.[148] He cites the *Navigation Rights* case as an example of evolving treaty interpretation, emphasizing that certain treaties, like the Geneva Conventions and their Additional Protocols, are meant to endure indefinitely. In *Navigation Rights,* one question the ICJ had to answer revolved around the evolving meaning of the word ‘commerce’ which was deemed to be a generic term.[149] The answer given by the ICJ was that the generic terms of a treaty are meant to evolve with the realities of the

144 Brian J. Egan, 'International Law and Stability in Cyberspace' [2016] 35(1) Berkeley Journal of International Law 178
145 A/76/136 (n 142) 116, 119
146 Kubo Macak, 'Military Objectives 2.0: The Case for Interpreting Computer Data as Objects under International Humanitarian Law' (2015) 48 Isr L Rev 66-67
147 *Ibid.* 67-68
148 *Ibid*. 69-70
149 *Dispute regarding Navigational and Related Rights (Costa Rica v. Nicaragua)*, Judgment, ICJ Reports 2009 (*Navigational Rights*) 243

world.[150] The ICJ also presumes that when generic terms are used, the parties to a treaty enter into it with tacit knowledge that the terms are meant to evolve.[151] A similar presumption can be made when the parties enter into a treaty that is meant to be in effect for a long period of time or for an indefinite period.

Mačák's position is that treaty interpretation should evolve with changing circumstances and align with the object and purpose of the treaty. In the case of the Geneva Conventions and their Additional Protocols, which aim to improve the protection of victims of armed conflicts, excluding attacks against data would be contrary to this purpose. He argues that data should be considered an object under international humanitarian law in light of this evolving interpretation.[152] This is also reflected in jurisprudence.[153]

To support his argument, Mačák examines the different official languages of the treaty and notes that there are two distinct camps of thought. While some versions, including English, Arabic, Russian, and Chinese, align with the "visible and tangible" requirement, others, such as the French and Spanish versions, use terms that, within their respective domestic contexts, explicitly include intangible elements.[154]

Mačák further contends that only objects can be damaged or destroyed, and he refutes the notion that cyber operations targeting data are akin to psychological operations. Although the damage to data may not always be immediately observable, it is analogous to damage that goes unnoticed by a belligerent until later, such as damage to a bridge that remains undetected for a time.[155]

Moreover, Mačák argues for the inclusion of data as objects based on the object and purpose of the Geneva Conventions and their Additional Protocols. To improve the protection of victims in armed conflicts, it is essential to consider attacks against data as a potential violation of international humanitarian law.[156] He also emphasizes that the right of belligerents to cause harm is not unlimited, supporting a more restrictive interpretation.[157]

---

150 *Ibid*.
151 *Ibid*.
152 Kubo Mačák (n 145) 69-70
153 *Ibid.* 71
154 *Ibid.* 71-72
155 *Ibid*. 72-74
156 *Ibid. 77-78*
157 *Ibid*. 79

In addition to Mačák, other scholars contribute diverse viewpoints to the discussion. Mavropoulou challenges the Tallinn Manual's reading of the "visible" and "tangible" requirements, arguing that these terms were meant to distinguish objects from abstract concepts rather than exclude data.[158] Furthermore, arguing that data cannot be an object because it can be restored would strip objecthood from other objects easily restored.[159] Todd examines national cybercrime legislation to support the idea that data can be considered an object, pointing out that cybercrime and cyber warfare affect computer data similarly.[160] For instance, section 1030 of title 18 of the United States' code defines damage in cyberspace as "any impairment to the integrity or availability of data, a program, a system, or information."[161] Meanwhile, Dinniss criticizes the emphasis on tangibility, highlighting that data can fulfill the requirements of a legal military objective.[162] Furthermore, when attacking dual-use objects, the military objective must be defined at the minimum level.[163] In computer systems and networks, the minimum level is data.

The Finnish position differs somewhat from that of the United States and the United Kingdom, as it leaves open the possibility of other types of consequences, such as long-lasting, significant economic damage beyond death, injury, and substantial material damage, when assessing cyber operations as armed attacks.[164] There are also additional requirements for sufficiently severe consequences, such as territorial requirements for the state claiming an armed attack.[165] In the realm of *jus in bello*, Finland explicitly includes civilian data as part of civilian objects to be safeguarded against damage.[166]

In summary, the debate surrounding the status of data as an object under international humanitarian law is multifaceted, with various scholars offering diverse interpretations and arguments. The discussion is influenced by evolving treaty interpretation, the object and purpose of the relevant treaties, and linguistic differences in treaty texts. Different states also hold varying positions on this complex issue.

158 Elizabeth Mavropoulou, 'Targeting in the Cyber Domain: Legal Challenges Arising from the Application of the Principle of Distinction to Cyber Attacks' (2015) 4 JL & Cyber Warfare 49-50
159 *Ibid.* 50-51
160 Graham H. Todd, 'Armed Attack in Cyberspace: Deterring Asymmetric Warfare with an Asymmetric Definition' (2009) 64 AF L Rev 82-83
161 *Ibid*. 83
162 Heather A. Harrison Dinniss (n 116) 47-49
163 *Ibid*. 51
164 Ministry of Foreign Affairs of Finland, International Law and Cyberspace Finland's National Positions (2020) 6
165 *Ibid*.
166 *Ibid*. 7

4.2 Data as a Part of a Wider System

As elucidated in the preceding discourse, the question of whether data, in its intrinsic form, can be considered an object and consequently be subject to the protective tenets of international humanitarian law remains a subject of unresolved contention within the realm of scholarly discourse. However, when delving into the issue of targeting data within larger systemic contexts and the attendant rules that govern such actions, a more clearly delineated framework emerges. This is chiefly due to the fact that targeting data as an element of a broader system is not contingent upon whether data itself attains the classification of an object under international humanitarian law. In this context, the focal point of an attack remains the overarching system, and the criteria for ascertaining its legitimacy as a military objective are correspondingly centered on the system rather than the specific data components that serve as vectors for the attack.

Data, as a fundamental component, can be integrated into a wide spectrum of systems reliant on computer technology or networked infrastructure. This expansive purview encompasses diverse domains, ranging from the computerized systems inherent in healthcare facilities, such as patient record systems, to automated systems within sectors like agriculture and power generation. The susceptibility of these systems to the ramifications of cyber operations underscores their indispensability in contemporary society.

As previously touched upon in the examination of the Tallinn Manual's perspective on data objecthood, operations explicitly targeting data within a system can be broadly classified as an attack if the overarching aim of the operation is to inflict harm or fatality upon individuals or to cause damage or destruction to objects. This categorization persists irrespective of whether actual harm or injury materializes and applies equally in cases where the operation was not originally intended to cause such effects but inadvertently does so.[167] This serves as a foundational guideline for identifying whether an operation may be characterized as an attack. However, it represents the introductory phase of analysis when assessing broader systems and the potential protection they may warrant from cyber operations.

A salient exemplar of systems in which data assumes an integral role is the realm of medical systems. Hospitals, for instance, extensively utilize computer technology for myriad functions, spanning from the routine task of patient record-keeping to more intricate applications like computer-assisted surgical procedures. These applications depend on computers to employ, store, and transmit data in various

167 Tallinn Manual (n 120) 418-419

formats, the compromise of which can result in catastrophic consequences. In line with international humanitarian law, these systems, along with the data encapsulated therein, are afforded general protection commensurate with that granted to medical personnel, equipment, and units. It is noteworthy that this safeguarding does not imply universal protection but rather extends to medical computers, data repositories, and the networks that constitute integral facets of medical unit operations and administration.[168] This protection encompasses various types of data, including information essential for the operation of medical equipment and individual patient medical records.[169] However, this protection may be diminished in cases where content-level medical data, such as patient records, are hosted on servers that also house military-related data.[170] Naturally, the protection is nullified if the medical personnel or units actively participate in actions detrimental to the enemy. It is imperative to note that the assumption of the capability to partake in such actions without actual engagement does not inherently lead to the nullification of protection. As an illustrative example, consider a medical computer with the latent capability to engage in a distributed denial of service (DDoS) attack.[171] The aforementioned principles are also extended to encompass religious personnel and the data associated with their functions, as per the stipulations articulated in Tallinn Manual Rule 131.

Regarding installations containing dangerous forces, the protection afforded by the Tallinn Manual interpretation is less strict than that given in Additional Protocol I. Where the Protocol prohibits attacks on such installations even when qualifying as a military objective, the Manual interprets there being a duty of care to guarantee that the forces contained do not escape during an attack.[172] The reasoning behind this is based on Rule 42 of the ICRC Customary IHL Study and is viewed as the baseline states must adhere to. The stricter wording of the Protocol binds only those states party to it specifically.[173] What limits this rule in the protection it gives is that it is specifically limited to apply only in cases of dykes, dams and nuclear power plants, and military objectives within their immediate vicinity.[174] Again, the limitation of computers, networks, and data is integral to the functioning and containment of the forces within the installation and the supporting functioning thereof. The reality of this rule is that it only calls for special care on the part of the attacker when attacking these installations. It is clear why

168 *Ibid.* 515
169 *Ibid*.
170 *Ibid*.
171 *Ibid*.
172 *Ibid*. 529
173 *Ibid*.
174 *Ibid.* 530-531

the international group of experts would defer to the less strict customary rule on the matter, however. The aim of the Manual to begin with is to stay within the *lex lata* of international humanitarian law. It is far easier to hold existing customary law as the *de facto* floor of the regulation on the matter since it is binding to all states equally, unlike the Additional Protocol, which lacks some key players in the field, like the United States, from its state parties.

The issue of cyber operations with the potential to harm data systems used in critical sectors such as agriculture or water supply is addressed within the framework of Rule 107 of the Tallinn Manual, which pertains to the prohibition of using starvation as a method of warfare. This rule explicitly forbids the use of cyber operations in a manner that could endanger the essential resources needed for the sustenance of a civilian population, including water, with the intent of weakening or causing harm to them.[175] It's important to note that this prohibition applies specifically to the use of starvation as a method of warfare, which somewhat narrows the scope of protection provided by the rule. However, it's crucial to avoid drawing overly asymmetrical conclusions by considering that incidental starvation or suffering due to war-related factors is not strictly prohibited under international humanitarian law in general. Instead, such incidental consequences are subject to the principles of proportionality and precaution, as will be discussed later in this thesis.

Regarding the safeguarding of objects indispensable to the survival of civilian populations, the interpretation provided by the Tallinn Manual aligns with the broader principles of international humanitarian law. Rule 141 of the Manual stipulates that objects meeting this criterion cannot be intentionally targeted, destroyed, removed, or rendered useless. Nevertheless, this rule's application is limited to cases where the explicit objective of a cyber operation is to deprive civilian populations of essential resources necessary for their survival. Consequently, the scope of protection offered by this rule is constrained, as most damage to objects indispensable for civilian survival typically arises as incidental harm during the course of armed conflict. Furthermore, the rule exclusively pertains to objects that are strictly necessary for civilian survival, excluding those that enhance the conditions of survival. For example, access to the internet is explicitly not considered an indispensable object, although networks and data integral to electricity, water, and food supply systems could potentially qualify as such depending on the circumstances.[176] It's important to note that the protection afforded by this rule does not encompass data unless the damage inflicted extends beyond the data itself.

175 *Ibid*. 459-460
176 *Ibid.* 533

The natural environment and ecological systems are not primary targets of cyber operations, but there are scenarios in which cyber operations on broader systems can lead to environmental harm. International humanitarian law, as elucidated in Tallinn Manual Rule 99, classifies nature and the natural environment as civilian objects, affording them protection from direct attack. Additionally, Rule 143 of the Manual imposes further obligations on states parties to Additional Protocol I by prohibiting means and methods of warfare that result in, or are expected to result in, widespread, long-term, and severe damage to the natural environment. To illustrate this concept, consider a cyber operation targeting a military fuel and petroleum system that subsequently results in the release of large quantities of fuel, causing environmental contamination. Similarly, a cyber operation leading to an oil leak into natural waterways would fall under this framework.[177]

In certain cases, data can play a role in impartial humanitarian assistance efforts. In these contexts, parties to a conflict are obligated not to unduly interfere with such efforts through cyber means.[178] This obligation is based on the overarching principle of international humanitarian law, which mandates that belligerents allow and refrain from interfering with humanitarian assistance efforts without regard to geographic limitations.[179] "Unduly interfering" via cyber means encompasses any cyber activities that obstruct or impede legitimate humanitarian assistance efforts aimed at providing relief to civilians affected by the hardships of war.[180] This prohibition extends beyond the targeting of data used in relief operations and encompasses other forms of cyber interference that hinder humanitarian efforts.[181]

177 *Ibid*. 538
178 *Ibid.* 540
179 *Ibid.* 541
180 *Ibid.* 542
181 *Ibid*.

5. Cyber Operations, Attacks, and Weapons

The concept of "attack" in the context of international humanitarian law is a crucial element in understanding how this body of law regulates and restricts the use of cyber operations in warfare. This chapter delves into the nuanced understanding of what constitutes a cyber attack in the legal sense, distinguishing it from the broader term "cyber attack," often used colloquially. It highlights that while both everyday language and international humanitarian law may refer to military operations as cyber attacks, only those cyber operations that meet the legal criteria for an attack under international humanitarian law can be deemed as such. To be considered a legal cyber attack, an operation must cross the threshold of causing substantial damage, whether the damage is purported or actual. This chapter explores the differentiation between the two senses of the term "cyber attack" and provides support for the argument that legal cyber attacks involve a level of damage that warrants their classification as attacks under international humanitarian law.

Additionally, this chapter addresses the concept of an "armed attack" under the *jus ad bellum* framework in the context of cyber operations. It examines how cyber operations fit within the definition of an armed attack and considers the perspectives of relevant states on this matter. Understanding when a cyber operation qualifies as an armed attack is crucial in determining the threshold for the lawful use of force in cyberspace.

Moving forward, the chapter shifts its focus to various vectors of cyber operations, including distributed denial of service (DDoS) attacks, hacking, and technical analysis of malware used in cyber operations. It analyzes the extent to which these vectors can adhere to the limitations imposed by international humanitarian law. By examining the capabilities and characteristics of different cyber vectors, the thesis aims to identify the boundaries within which these vectors can operate in compliance with international humanitarian law and where the line may be drawn between legal and illegal means of cyber warfare. This analysis provides insights into how states and non-state actors can engage in cyber operations while adhering to the legal constraints set by international humanitarian law.

5.1 Cyber Attack vs. Cyber ‘Attack’

5.1.1 *Jus in Bello* and Ordinary Understanding of the Word “Attack”

In common language and media discourse, the term "cyber attack" is frequently used without a full appreciation of its legal significance. The legal context of the term "attack" in international humanitarian law differs significantly from its colloquial usage. In legal terms, "attack" carries a specific meaning and weight, slightly differing between the realms of *jus ad bellum* and *jus in bello*. Therefore, using the term casually can lead to misunderstandings, as what may be considered an attack in everyday language may not meet the legal criteria for an attack under international humanitarian law, and vice versa.

As established in Chapter 2 of this thesis, an "attack" in the Geneva Conventions and their Additional Protocols is defined as violence, either offensive or defensive, against an adversary. This definition is not limited by geographical boundaries but requires the violence to be under the control of an adverse party. The Commentary of 1987 further clarified the term by introducing the concept of "combat action," emphasizing that attacks must be considered at multiple scales. While attackers at the operational level often engage in what is traditionally seen as an attack, defenders may also conduct combat operations as part of their defense. This distinction is essential in understanding that some actions, like signal interference, may not cause direct physical harm and therefore may not qualify as an attack under international humanitarian law.

However, there are notable weaknesses in the Additional Protocol's understanding of attacks. For instance, violence inflicted by an adverse party upon itself, known as the "scorched earth" defense, is not categorized as an attack,[182] provided the violence is confined to the pre-conflict area of the adverse party in question. In the borderless realm of cyberspace, this presents challenges, particularly regarding the location of servers and networks. If the target nation has a theoretical cyber defense system that includes a "wiper" module to destroy data, it might affect servers located outside the target nation's territory. This raises the possibility of a "digital scorched earth" scenario, which might be considered an attack only if the wiped system could be seen as national territory under sovereign immunity principles.[183] It's important to note that such systems would require robust backup and data restoration

182 Definition of Attacks and Scope of Application (n 38)
183 Sean Watts & Theodore Richard, 'Baseline Territorial Sovereignty and Cyberspace' (2018) 22 Lewis & Clark L Rev 824

mechanisms, and their effects would likely be localized, but the theoretical possibility of wider damage exists.

The Tallinn Manual, as discussed in Chapter 3, defines an "attack" in cyberspace based on the physical damage or threats of damage caused by a cyber operation. The manual's Rule 92 emphasizes that cyber operations, whether offensive or defensive, are considered attacks if they are "reasonably expected" to result in injury, damage, death, or destruction. This rule excludes purely psychological operations and operations equivalent to espionage from the definition of an attack.[184] The focus is on the consequences of the cyber operation rather than the operation itself. By focusing on the consequences, a concession is made for non-kinetic operations such as chemical, biological, and radiological to be considered attacks.[185] This decouples the link between attacks and the release of kinetic forces. The manual also lists specific cases in which attacks on computer data may cross the threshold of harm to qualify as attacks under international humanitarian law. These cases include cyber operations causing disease outbreaks, severe mental distress similar to terrorizing the population,[186] targeting digital cultural property,[187] or interfering with the functionality of objects to the extent where hardware repairs are needed to restore their functionality.[188]

It's important to recognize that not all military operations in cyberspace can be classified as attacks in the legal sense. Cyber operations encompass various forms, many of which do not meet the current definitions of an attack. Some operations, such as Titan Rain,[189] are espionage-like intrusions, and DDoS attacks have equivalents outside cyberspace, espionage and signal interference, respectively, that are not considered attacks. Titan Rain is a name given to a series of network intrusions and data exfiltrations on American companies such as Lockheed Martin and state departments such as the Department of Energy.[190] Titan Rain targeted systems containing industrial designs and retrieved copies of them without causing damage. This makes them cases of cyber espionage rather than cyber warfare. The origin of the attack was in China, and the malware used by the hackers was similar to that found in other network intrusion cases also traced back to China.[191] The APT responsible for the attacks was

184 Tallinn Manual (n 120) 415
185 *Ibid*. 415-416
186 *Ibid.* 417
187 *Ibid*.
188 *Ibid.*
189 William T. Hagestad, 21st Century Chinese Cyberwarfare (IT Governance Publishing 2012) 27-29
190 Marieke Lomans, "Investigating Titan Rain" Cyber Security & Cyber Operations (2017) 1
191 *Ibid.* 6

connected to the People's Liberation Army (PLA) of China by an investigation done by the American government.[192] The aforementioned operations, being the most common types of cyber operations, serve as an important safety mechanism in international humanitarian law. Not being recognized as attacks means that the targeted state does not have the right to resort to force in responding to them.

For the ruling understanding of *lege lata,* a cyber operation is to be considered a cyber attack if it involves a violent act with the potential for harm, occurring through a computer or computer network. This definition strikes a balance between encompassing cyber operations that cause harm and excluding those that do not result in significant damage or threat. Including non-destructive cyber operations as attacks would constrain states' existing coercive measures in cyberspace, lead to non-adoption of legal rules via states rejecting them as binding, and could lead to unwarranted escalation of conflicts by giving the target state of a DDoS attack the right to use military means to end such an attack. The meaningful distinction between cyber and kinetic attacks would effectively disappear.

The disruption or denial of access to data is another aspect of the cyber attack definition that some legal scholars include.[193] While this may not align with the traditional understanding of attack, it can be considered a form of cyber attack under the Tallinn Manual, especially if it leads to tangible effects in the physical world. Regardless, this means that operations that are so-called pure cyber attacks—attacks without effects in the realspace such as the destruction of medical records or banking information—cannot be seen as attacks in a legal sense. It can be argued that pure cyber attacks are only theoretical constructs since everything online has a connection to something offline; lost medical records can lead to injuries or deaths due to the wrong kind of care; and lost banking information has effects on how well people can navigate their lives.

The purpose of the Geneva Conventions' Additional Protocols' to alleviate the negative consequences of war must be held in view. Within this purpose, the balancing of humanity and military necessity reigns. Disruption or denial of access to military data cannot be included in the category of actions prohibited by the Protocols since the definite military necessity to reduce the capabilities of the enemy forces exists. However, this can be done with means that are akin to mild and temporary inconveniences to the civilian population, should their effect be wider than anticipated. One such avenue of action is the use of encrypting malware like WannaCry in 2017. Such malware does not

192 *Ibid.* 8-9
193 Nicholas Tsagourias, 'Cyber Attacks, Self-Defence and the Problem of Attribution' (2012) 17 J Conflict & Sec L 229

destroy the targeted data; it only encrypts it in a way that can be reverted.[194] The effects are tantamount to a temporary inconvenience that civilians must, in most cases, endure. Anything else would shift the balance too much to the side of humanity at the cost of military necessity and would risk erosion of the protection afforded to civilians as states withdraw their consent to be bound by legal norms.[195] While the aforementioned could be seen as a *carte blanche* for the militaries to do as they like, viz. disruption and denial of access to data, this analysis would be wrong. Some datasets, such as medical data, must at all times be not disrupted and readily available for use and cannot, under any circumstances, be subjected to attacks of any kind. The protection such datasets enjoy is grounded in their being constituent parts of specially protected entities.[196]

Under the paradigm explored above, system intrusions and data exfiltration would never cross the threshold of an attack. However, there is the issue of cyber pillaging, under which conditions of unlawful conduct would be met. Cyber pillaging is defined as 'the expropriation of property by cyber means by a member of the armed forces for private or personal use in the context of an armed conflict'.[197] Cyber pillage can be anything from stealing the trade secrets of a company working for the enemy to requisitioning the personal data of a civilian for personal use later. The conditions of cyber pillage disappear if the data is taken for the use of the armed forces of the belligerent or if the original owner of the data is remunerated.[198]

5.1.2 *Jus ad Bellum*

Jus ad bellum concepts of use of force and armed attack are instructive when defining attacks *jus in bello*. Armed attacks are cross-border attacks that give rise to the target state's right to use military force.[199] While cyber operations may constitute the use of force, armed attacks typically involve significant physical damage or the threat of such damage. In the Charter of the United Nations, Article 2(4) prohibits the "threat or use of force" against territorial integrity or political independence.[200] Later, the Charter explicitly addresses armed attacks under the states' right to self-defense in Article 51.[201] The

194 Kaspersky Lab, 'What is WannaCry Ransomware?' (Resource Center, 20 April 2020) <https://www.kaspersky.com/resource-center/threats/ransomware-wannacry> accessed 11 May 2020
195 Michael N. Schmitt (n 135) 349
196 *Ibid*. 346
197 Christopher Greulich and Eric Talbot Jensen, 'Cyber Pillage' (2020) 26 Sw J Int'l L 281
198 *Ibid.* 285-286
199 Charter of the United Nations (adopted 26 June 1945 entered into force 24 October 1945) art 51
200 *Ibid.* art 2(4)
201 *Ibid*. art 51

division of the two into specific categories, only one of which explicitly grants the target state the right to answer in self-defense, means that situations in which the coercive actions of a state may be seen as the use of force but not an armed attack. Cyber operations can satisfy either category. The political independence of a state can be attacked via DDoS attacks, system intrusions, and influencing public opinion with information technology. All of the aforementioned have their respective analogs outside of cyberspace. Some legal scholars posit that national cyber infrastructure physically located within the borders of a state is an extension of national sovereignty, and unauthorized access to such infrastructure would be tantamount to breaching that sovereignty.[202] The Tallinn Manual does affirm that, though states cannot claim sovereignty over cyberspace per se, they do exercise sovereignty over cyber infrastructure within their borders and cyber activities conducted by persons and entities within their borders.[203] Internally, the sovereignty of states is all-encompassing. States not only have control and sovereignty over the cyber infrastructure but can also restrict and mandate the use of specific technical protocols to access cyberspace therein.[204] Additionally, states can limit the scope of activities within the cyberspace of the infrastructure they exercise sovereignty over.[205] Externally, all states have the right to engage in cyber activities, subject to limitations of international law such as prohibitions on violating other states' sovereignty, intervention, and use of force.[206]

The Tallinn Manual's Rule 71 reaffirms the states' inherent right to self-defense in case of an armed attack in cases where a cyber operation crosses the threshold of an armed attack. How the threshold is determined lies in the evaluation of the cyber operations' scale and effects. The Rule prescribes that a cyber-armed attack must have a trans-border element.[207] Having a trans-border element means that the origin of the attack must be computer infrastructure from outside the target state's borders. The requirement is also met when a state launches or, not dissimilar to the case in *Nicaragua*, instructs non-state actors outside its borders to launch a cyber operation. Proving such connections is a difficult matter due to the myriad ways actors on the internet can spoof their digital identities, such as IP addresses. The problem one often runs into is the problem of attribution. Put simply, it states that there are little fool-proof ways of attributing a cyber operation to a single actor or a group of actors outside of an actor taking credit for the operation. This is doubly true when trying to attribute a cyber operation

202 Sean Watts & Theodore Richard (n 183) 823
203 Tallinn Manual (n 120) 12-13
204 *Ibid.* 14
205 *Ibid.*
206 *Ibid.* 17
207 *Ibid.* 340

to a state-controlled actor, as the connection between the actor and the controlling state has similar difficulties as connecting the actor and the operation. Burkadze points out similar issues with the definition of armed attack and the use of force in cyberspace. There are no clear tests for states to conduct, but she outlines a similar heuristic to the Tallinn Manual's scale and effect stipulation. The factors that should be assessed are the context of the adverse event, the actor perpetuating the action, the target and location, the effects of the action, and the intent of the perpetrator.[208]

Todd points to the tendency of cyber attacks to have long intervals between the attack itself and the full effects of the attack, and how this reveals weaknesses in the effects based analysis model of cyber-armed attacks.[209] Instead, he advocates for a weapons-based analysis model. This model of analysis focuses on the means used in the operation rather than its effects. Todd justifies using this model with reference to the UNGA's definition of aggression, which includes the phrase "any weapon."[210] Furthermore, such an approach is analogous to the way criminal law approaches criminal acts; the action determines whether or not a crime has happened, and the effects only determine the severity of the crime.[211] The definition of what constitutes a weapon is the crux around which the weapons-based model revolves. Traditional definitions of weapons often limit themselves to kinetic weapons, with some exceptions given to some classes of non-kinetic weapons such as chemical and biological weapons. However, the definition of a weapon is a living concept that changes with time and the development of novel technologies used in warfare.

The state under a cyber-armed attack can choose to respond in any way it would to an armed attack of non-cyber character. What this entails is that states can use kinetic weapons in cyber operations should the threshold of armed attack be crossed.[212] The reasoning for such a liberty of means is based on the ICJ's *Nuclear Weapons* advisory opinion. The ICJ concludes that UN Charter Article 51 does not limit itself, viz. any specific weapon type but is weapons-agnostic and clearly meant to apply in all cases where a threshold of armed attack is crossed.[213] Furthermore, tying the concept of armed attack with the release of kinetic forces would lead to an absurd situation in which military operations that would

208 Khatuna Burkadze, 'A Shift in the Historical Understanding of Armed Attack and Its Applicability to Cyberspace' (2020) 44 Fletcher F World Aff 40
209 Graham H. Todd (n 159) 77
210 *Ibid*. 78
211 *Ibid*.
212 Tallinn Manual (n 120) 340
213 *Legality of the Threat or Use of Nuclear Weapons*, (Advisory Opinion, 8 July 1996), ICJ Reports 1996 (*Nuclear Weapons*) 244

clearly cross the threshold of armed attack otherwise, such as chemical attacks, would not do so, leaving the target state without a legal right to retaliate. The right to answer in a style of their choosing is not free of the limitations that principles of necessity and proportionality put on all attacks, as was seen in *Nicaragua*.[214]

Regarding the meaning of 'scale and effects' on which the definition of a cyber-armed attack relies, the Manual explicitly refers back to *Nicaragua* and how the term is used therein.[215] The Manual adopts the stance taken by the ICJ that the scale and effects must be sufficiently grave so that the mere use of force can be considered an armed attack. What sufficiently grave scale and effect mean, remains largely unsettled. However, the Manual does point to cases that are unambiguous, such as intelligence gathering and theft, as well as cyber operations that result in "brief or periodic interruptions of non-essential cyber services." What is considered a 'non-essential' cyber service the manual does not delve into; however, referring to the Manual's conception of a cyber attack, it can be said that defacing or DDoS attacks of websites are included in that category. In all of the above cases, the threshold for an armed attack is never crossed.[216] Reversely, cyber operations that lead to injury or death of people or significant damage or destruction of property are cases in which the 'sufficiently grave' requirement is always satisfied. Radziwill disputes this, pointing out that whether or not strikes resulting in singular or few fatalities are considered armed attacks by states has not been uniform.[217] On the question of when a cyber operation becomes a cyber-armed attack, the international group of experts penning the Manual was split. The ICJ ruled in *Nicaragua* that "mere border incidents" do not qualify as an armed attack,[218] and later in *Oil Platforms* ruled that armed attacks can be as narrowly targeted as a singular installation.[219] The Manual specifically examines the case of Stuxnet, and in that case, the International Group of Experts is divided in their opinion as to whether or not it could be considered an armed attack. On the matter of whether many smaller-scale cyber operations can, in aggregate, constitute an armed attack, the experts agree. Should multiple smaller-scale cyber operations originate from the same source, and in aggregate, cross the threshold of sufficiently grave scale and effects, they can be seen as an armed attack, and the target state can act in self-defense.[220]

214 *Nicaragua* (n 53) 94
215 Tallinn Manual (n 120) 341
216 *Ibid.*
217 Yaroslav Radziwill, Cyber-attacks and the Exploitable Imperfections of International Law (Brill Nijhoff 2015) 143
218 *Nicaragua* (n 53) 103
219 *Oil Platforms* (n 55) 189-190
220 Tallinn Manual (n 120) 342

Some effects of a cyber attack can be as catastrophic as an armed attack without causing injury or death to people or damage to or destruction of property. Imagine a widespread cyber attack on the state's whole banking system or central stock exchange that causes catastrophic financial losses and destroys all confidence in the national economy. The scale would be massive, and the effects, though temporary, would not be dissimilar to an actual outbreak of hostilities. Yet the attack has no human casualties, and there is no tangible destruction of property. The International Group of Experts was divided on whether to extend the definition to a scenario such as described.[221] The arguments ranged from not wanting to make the nature of the consequence of an operation the locus of defining armed attacks but focusing on the extent of the consequences to disagreement on whether or not financial damage is under the ambit of property damage requisite of an armed attack. Similar disagreements characterize the Manual's treatment of the effects of an operation one should focus on when determining the existence of an armed attack. The Manual makes it clear that the foreseeable consequences of an operation are unequivocally included in the determination.[222] A cyber attack against water treatment infrastructure leads to a lack of clean drinking water, which leads to the emergence of diseases like cholera, leading to injury and death of people. In such a case, the causal chain consists of foreseeable events and therefore must be taken into account when appraising the legal status of the cyber operation; in this case, the operation would be an armed attack. Burkadze lists three different scenarios in which a cyber-armed attack has indubitably happened: a cyber operation leading to nuclear meltdown, a cyber operation that opens a dam in a populated area, and a cyber operation disabling air traffic control and causing an airplane to crash.[223] Chang approaches the issue of non-death/destruction cyber operations through hypothetical scenarios in lieu of sufficient state practice.[224] First is the hypothetical of including cyber operations resulting in mere disruption as armed attacks. This would be preferable to states since it would protect their vital cyber infrastructure as disrupting them would risk an armed response and include them under the protective umbrella of international humanitarian law.[225] Such inclusion could be justified by the object and purpose of international humanitarian law, that is, to protect civilians from the effects of war. Without such wide protection, there would be a *lacuna* in legal protection – a space in which states would be free to cause mayhem to each other and their populace. The second

221 *Ibid.* 342-343
222 *Ibid.* 343
223 Khatuna Burkadze (n 207) 40
224 Zen Chang, 'Cyberwarfare and International Humanitarian Law' (2017) 9 Creighton Int'l & Comp LJ 32
225 *Ibid*. 32-33

hypothetical approach is to focus on the intentions and perceived threat of the operation.[226] This approach reflects the current state practice of choosing whether or not to react to the use of armed violence by another state, whether by mistake or otherwise.

The intentionality of the consequences of a cyber operation does not matter when assessing whether or not an armed attack took place.[227] Accordingly, the possibility of an operation that was not meant to produce consequences concomitant with armed attacks does so and results in the target state gaining the right to act in self-defense and use armed force. The possibly resulting armed response must be proportionate to the damage caused, however.[228] Similar questions of intentionality of consequences are present in cases of bleed-over effects of a cyber operation. Bleed-over effects are consequences of an operation that happen as a consequence of an operation elsewhere. An example of bleed-over effects would be state A releasing malware into the network of state B and the malware infecting and compromising systems in state C that were not the target of the operation. In the aforementioned situation, should the consequences for state C be severe enough in scale and damage, an armed attack has occurred, and the state suffering the consequences can answer with proportionate military force.[229]

Another pertinent question is that of the originator of an armed attack, or who can launch operations that could be considered armed attacks. The obvious originators are the state and its organs and, as seen in *Nicaragua,* armed groups acting on behalf of or sent by a state.[230] By this logic, hacker groups and individual hackers can be seen as the originators of armed attacks when operating by orders of or on behalf of a state, should the scale and effects of the consequences cross the threshold of an armed attack. In such cases, the targeted state has the right to a proportionate military answer.[231] On the matter of non-state actors not acting on behalf of a state, there is no consensus within positive international law or international case law.[232] However, what is clear is that state practice points to the existence of a right to self-defense against armed attacks on the territory of a state by unaffiliated non-state actors.[233] The International Group of Experts compiling the Tallinn Manual agreed that terrorist attacks by groups without state backing via cyber means could constitute an armed attack if their effects rise to the

226 *Ibid.*
227 Tallinn Manual (n 120) 343-344
228 *Ibid.*
229 *Ibid.* 344
230 *Nicaragua* (n 53) 103-104
231 Tallinn Manual (n 120) 344
232 *Ibid.* 345
233 *Ibid.*

threshold of an armed attack.[234] However, the Group could not settle further questions on the matter of unaffiliated armed groups, such as whether or not such groups need to be sufficiently organized or whether a single individual could launch an operation entitling a state to respond with violence.

The Manual next turns to the question of territory, viz. cyber operations crossing the threshold of armed attack. The Group of Experts agrees that the most obvious case of a fulfilled territoriality requirement is when a cyber operation crossing the threshold of an armed attack is trans-border in character, meaning that it originates from a different state than the targeted state.[235] Such being the case, it does not matter whether or not the damaged property or injured persons are public or private; it all counts as an armed attack against the state on whose territory the damage happens. Regarding the question of whether a state can claim self-defense in cases where damage is done to persons or property under its nationality but outside its territory, the Group of Experts remains undecided. The answer is dependent on multiple factors that need to be weighed, such as the extent of damage, the status of harmed property and persons, and whether or not the attack was politically motivated or whether targets were chosen based on their nationality.[236] One case where every factor listed above is present is an attack against property or representatives of a government other than the territorial sovereign. Such operations are considered armed attacks against the represented state if other factors, such as sufficient damage caused, are satisfied. Bobrowski points out the difficulty of attributing the origin of a cyber attack to a singular state since the attacker can route their attack to go through the cyber infrastructure of different states than the state they originally inhabit.[237] The target state's right to retaliate against the original attacker is not diminished by such routeing; however, whether or not it can affect the cyber infrastructure of a third state is not clear. In cases in which the third state has unequivocally failed its duty of due diligence to prevent its territory from being used for deleterious activities against another state, the answer is yes.[238]

The state under armed attack is not unlimited in its options to retaliate against cyber-armed attacks. The response employed must be necessary in order to stop the attack, and proportionate to the attack as a whole. With the general limitations these requirements impose on the retaliating state come specific requirements that need to be met before resorting to force. The retaliating state must be sure that an

234 *Ibid.*
235 *Ibid.* 346
236 *Ibid*.
237 Krzysztof Bobrowski, 'Conventional Attack vs Digital Attack in the Light of International Law' (2021) 10 Polish Rev Int'l & Eur L 92
238 *Ibid*. 91-92

armed attack has occurred or is imminent, and it must be sure of the identity of the attacker. Both of these assessments are subject to the information at hand at the time, independent of any conflicting information that might later be revealed. In principle, the retaliating state can aim its countermeasures at armed attacks within its borders, within the borders of the state the attack originated from, in international waters, airspace, or outer space.[239] The right to aim countermeasures at an attacker in another state is affirmed by the discipline of state responsibility as well. While in general the conduct of armed groups outside the control or instruction of a state is not considered an act attributable to the state itself,[240] the condition of self-defense is recognized in the ILC draft articles on state responsibility. According to the draft articles, the wrongfulness of an act by a state, in this case the military countermeasures, is precluded when an act is in conformity with and pursuance to the right of self-defense enshrined in the United Nations' Charter.[241] The preclusion of wrongfulness is only limited to the existence of the countermeasures, leaving intact the norms and obligations the responding state has, viz. international humanitarian law and human rights law.

The rules on defensive cyber operations are less strict. The Manual states that defensive cyber operations make an exception to the principle of sovereignty in cyberspace and can initiate, employ assets in, and be launched from a state that is neither the originator nor the victim state.[242] Naturally, this is only pertinent to non-consensual defensive actions since states can give permission to use their sovereign territory for military actions to others. The Group of Experts could not find consensus on how far this right for non-consensual action goes. Some tied it to the principle of necessity combined with the state in question being unwilling or unable to effectively stop its territory or cyber infrastructure from being used in an armed attack, while others were more strict and tied the right to the authorization of the Security Council.[243]

5.1.3 States' Positions

The material from which the positions of Finland, the United Kingdom, and the United States are derived relies on the official release of the ministry of foreign affairs from 2020 in the case of Finland and on a United Nations compendium of national positions on international law's applicability, viz.

239 Tallinn Manual (n 120) 347
240 ILC, 'Draft Articles on Responsibility of States for Internationally Wrongful Acts' (2001), UN Doc A/56/10 50
241 *Ibid.* 74
242 Tallinn Manual (n 120) 347
243 *Ibid.*

information and communications technology, released in 2021 for the United Kingdom and the United States. These are the most up-to-date releases at the time of writing and can be read as binding to the states as they are given in an official capacity by the organs of the respective states.

All states examined in this thesis are in agreement that, in cases of armed attacks triggering the states' right to respond in self-defense, mere damage within cyberspace does not cross the threshold of damage. The states are rather unanimous in their view that for a cyber operation to be considered an armed attack in the *jus ad bellum* sense, some form of physical, realspace damage must occur, or at least there has to be a significant threat of such damage.[244] All states also agree that the damage, whether real or threatened, must be significant. An often-used threshold is that the consequences of a cyber operation must be similar to those of kinetic operations. Having such stringent requirements for an armed attack is reasonable since an armed attack and the subsequent triggering of the right of self-defense lead to the possibility of resorting to an armed response.

As alluded to above, the national positions examined differ in their views on the *jus in bello* definition of cyber attack. To reiterate, the United Kingdom and the United States limit their conception of attack, ergo cyber attack, to the consequences of the attack in realspace. Ergo, a cyber attack can be said to be, in both states' opinion, a cyber operation whose consequences are similar to those of a kinetic operation. Finland differs in this, with its conception that civilian data should be offered additional protection. Therefore it is safe to say that the Finnish conception of cyber attack is wider than that of the United Kingdom and the United States.

In summary, the definition of a cyber attack in international humanitarian law hinges on the concept of damage or the potential for damage. It is a term with specific legal implications, and its understanding may vary among states. While the threshold for an armed attack in the context of *jus ad bellum* is high and typically requires significant physical damage or threat, the definition of a cyber attack in *jus in bello*, as outlined in the Tallinn Manual, focuses on the consequences of cyber operations, whether they result in physical damage or other harmful effects. Different states may have varying interpretations of what constitutes a cyber attack, with some allowing for a wider definition than others.

244 Ministry of Foreign Affairs of Finland (n 164) 6
A/76/136 (n 142) 116, 137

5.2 Restrictions Applicable to Cyber Attacks

The principles of distinction, proportionality, and precautions in attack become especially salient when one operates within the framework provided by the Tallinn Manual's definition of an attack, which notably excludes the characterization of data as an object with inherent protection. Within the confines of this definition, the principles of distinction and proportionality naturally come into play. This logical progression stems from the Tallinn Manual's explicit stance that data, in and of itself, does not qualify as a protected object, thereby necessitating the application of these principles to discern the legality of cyber operations. It is important to note that this stance does not grant a carte blanche for arbitrary data destruction without real-world consequences; rather, it acknowledges that there exist instances where the destruction of digital property may not directly translate into tangible, real-world harm.

However, if the scope of protection from attack were to be extended to encompass data as an object deserving safeguarding, this expansion would inevitably encompass pure cyber attacks within its purview. This broader interpretation would render numerous known cyber weapons, and potentially unknown ones, illegal under the provisions of international humanitarian law. The application of these three cardinal principles—distinction, proportionality, and precautions in attack—is unquestionable. However, they merit a more in-depth examination, particularly within the unique context of cyber warfare.

5.2.1 Distinction

The Tallinn Manual underscores the pivotal role of the principle of distinction within the context of cyber conflict. When designating a cyber operation as an "attack," the principle of distinction assumes immediate relevance, imposing stringent constraints on the conduct of such operations. This principle primarily serves the critical purpose of delineating hostilities to ensure that only combatants and legitimate military objectives become the targets of hostile actions. The Manual aptly highlights that even beyond the realm of cyber warfare, there exist operations that primarily target the civilian population of an adversary, which may nonetheless find legal sanction. As an illustrative example, the Manual cites various forms of adversarial propaganda campaigns, such as the broadcasting of radio signals, which, while targeting civilians, do not qualify as attacks in the absence of impending harm or injury.[245]

245 Tallinn Manual (n 120) 423

Central to the principle of distinction is the dichotomy it establishes between combatants and civilians, with the aim of affording protection to both individual civilians and the civilian population as a whole, as delineated in Tallinn Manual Rule 94. It is essential to note, however, that this safeguard is circumscribed by the Manual's interpretation of the term 'attack,' which confines its purview to instances of direct harm inflicted upon civilians. In cases where civilians experience injury or fatality as collateral consequences of a cyber attack directed at a legitimate military target, the principle of distinction is deemed to have been upheld.[246] Nonetheless, there exist other complementary principles, such as those of proportionality and precautions in attack, that address incidental harm or loss of life among civilians.

A critical aspect emphasized by the Tallinn Manual, specifically articulated in Rule 95, pertains to the determination of the status of the purported target. In the context of cyber warfare, the Manual underscores the significance of elucidating whether the target qualifies as a combatant or civilian entity. The responsibility for clarifying the status of the target remains a contentious issue, with divisions among experts regarding whether this onus should rest solely upon the attacker or whether the defender should also assume a role in facilitating this distinction.[247] The complexity of ascertaining the target's status in the cyber domain is further compounded by the substantial overlap in the utilization of cyber infrastructure by both civilian and military entities. This ubiquity extends across various strata of cyber infrastructure, encompassing scenarios where civilians may utilize the same computing equipment as legitimate military targets or share connections within local networks, sometimes even taking part in hostilities under the cover of an anonymity that is hard to penetrate. Furthermore, the interconnected nature of the internet enables any device connected to the network to potentially interact with others, further complicating the task of differentiation. The Manual acknowledges these challenges and emphasizes the need for achieving a reasonable degree of certainty, considering the available resources and conditions at the time of the cyber attack.[248] It is notable that the principles governing distinction apply similarly to both civilian individuals and objects.

In line with the principle of distinction's application in the cyber context, the prohibition of indiscriminate attacks assumes significance.[249] Such attacks, by their very nature, constitute a stark violation of the principle of distinction, as they lack any form of discrimination in target selection. The

246 *Ibid*. 423
247 *Ibid*. 424
248 *Ibid.* 424-425
249 *Ibid*. 455, 467

prohibition against indiscriminate attacks extends to the indiscriminate use of cyber warfare means that have the potential for precise targeting.[250] The core tenet underlying this prohibition is that the attack employed must not be devoid of specific targeting and, by virtue of its indiscriminate nature, must not endanger civilians. Additionally, there exist separate, standalone rules delineating scenarios where attacks are specifically aimed at civilian targets, and these too are unequivocally prohibited.[251]

Regarding the positions of the states compared on the principle of distinction in cyber warfare, unanimity prevails in acknowledging that once a cyber operation surpasses the threshold of an attack, the principle of distinction becomes germane and must be meticulously observed by the attacking party. Both the British and American positions emphasize the congruity of limitations applied to cyber and kinetic attacks, emphasizing that comparable constraints should be applicable irrespective of the method of attack.[252] Conversely, the Finnish position concurs on the fundamental need to adhere to the principle of distinction and its derivative rules during cyber attacks without emphasis on consequences.[253]

5.2.2 Proportionality

Rule 113 of the Tallinn Manual, which addresses proportionality in the context of cyber warfare, establishes a crucial principle: the prohibition of incidental loss and damage to civilians and civilian objects that exceeds the concrete and direct military advantage anticipated from a given operation. This rule emphasizes that while some level of "collateral damage" may be inevitable and the parties to the conflict should tolerate it, the right to use means of cyber warfare that cause it has limits. Essentially, Rule 113 aims to prevent excessive use of force during an attack.[254]

The rule's scope extends not only to damage caused during the actual cyber attack but also to potential harm incurred while the attack vector traverses civilian cyber infrastructure.[255] This precaution is particularly relevant due to the inherent dual-use nature of much of this infrastructure. Typically, when an object has dual-use status, it could be considered a legitimate military target in its entirety. However, given the vast and interconnected nature of the internet and other cyber infrastructure, interpretations

250 *Ibid*. 468
251 *Ibid.* 422
252 A/76/136 (n 142) 119, 138
253 Ministry of Foreign Affairs of Finland (n 164) 7
254 Tallinn Manual (n 120) 471
255 *Ibid*.

must be tempered to avoid rendering the protections of international humanitarian law ineffective in cyberspace. For instance, the Tallinn Manual offers the example of GPS systems, the targeting of which could result in damage to civilian shipping and air travel, raising complex proportionality concerns.[256]

The determination of what constitutes intolerable collateral damage depends on a comparative assessment rather than fixed, quantifiable criteria.[257] This assessment hinges on the military advantage anticipated from the operation. In essence, the greater the expected military advantage, the more incidental civilian damage may be deemed acceptable, whereas operations with minimal anticipated advantage cannot justify significant collateral harm. Crucially, this assessment must be made in advance and should not rely on speculative evaluations by commanders. The advantage sought must be "concrete and direct" and must encompass the entire operation, including its cyber component.

National positions align with the principle of proportionality, asserting its applicability whenever a cyber operation reaches the threshold of an attack under *jus in bello*. While the Finnish position maintains conciseness, the British perspective emphasizes the similarity in consequences between cyber and kinetic operations. In contrast, the American position offers a more detailed perspective, highlighting the interconnectedness of machines within a network as a specific concern for cyber attacks. This interconnectedness necessitates additional caution to prevent incidental harm to civilian devices that share the same network as the targeted military objective.[258]

5.2.3 Precautions in Attack

Numerous additional precautions must be adhered to by belligerents engaging in cyber operations under the umbrella of international humanitarian law. These precautions serve to safeguard civilians and civilian objects, minimize harm, and uphold ethical standards in cyber warfare. Some of these crucial precautions include:

Constant Care for the Civilian Population[259]: Commanders and planners of operations are obligated to consider the potential impact of their cyber operations on civilians and civilian objects. This duty underscores the need for continuous vigilance in assessing the consequences of military actions on non-combatants.

256 *Ibid*. 471-472
257 *Ibid.* 473
258 A/76/136 (n 142) 138
259 Tallinn Manual (n 120) 476

Belligerents must take all feasible measures to ascertain the status of their targets.[260] This ensures that attacks are directed only against legitimate military objectives, preventing harm to civilians and civilian objects resulting from misidentification. Additionally, belligerents engaging in cyber warfare must choose means and methods that minimize incidental damage and injury to civilians and civilian objects.[261] This precaution emphasizes the importance of selecting cyber tools and tactics with precision and restraint. While these are notable examples, it's important to note that there are several other precautions in an attack stipulated by international humanitarian law. Many of these precautions, such as those considering the choice of targets based on projected military advantage, closely resemble principles applied to operations in physical space. One precaution deserving special attention is the obligation placed on the defending party to take measures against cyber attacks targeting them. This obligation distinguishes itself by being directed at the defending party rather than the aggressor, which is a departure from the majority of rules in the Tallinn Manual that primarily address the active party in an operation.

This rule asserts that parties involved in an armed conflict must adopt all necessary and feasible precautions to shield civilians from the dangers posed by cyber attacks.[262] Notably, this rule primarily pertains to passive defensive measures rather than active wartime actions. For instance, these measures could include the effective separation of military and civilian cyber infrastructure, the establishment of robust civilian data backup systems, and the implementation of antivirus measures.[263]

While this rule draws inspiration from Article 58(c) of Additional Protocol I, which mandates states to take necessary precautions to protect against the dangers arising from military operations, it differs in one crucial aspect. While Article 58 pertains broadly to protection from dangers resulting from military operations, the Tallinn Manual's rule narrows the focus specifically to attacks. At first glance, this might seem to limit the rule's protective scope. However, when viewed through the lens of the Manual's understanding of cyber attacks, it becomes evident that the balance between military necessity and civilian protection remains adequately preserved.

The rule acknowledges the limitation of obligations to the maximum feasible extent, which is not dissimilar to similar constraints in Article 58 of Additional Protocol I. Practical limitations also affect

260 *Ibid*. 478
261 *Ibid.* 479-480
262 *Ibid.* 487
263 *Ibid.* 488

the feasibility of protective measures. For example, some systems cannot be realistically segregated into distinct civilian and military systems and must remain dual-use systems, making them legal military targets. One such system is air traffic control, where operational imperatives necessitate a dual-use approach.[264]

Crucially, if the defending party fails to take the necessary precautions, this does not impede the attacker's right to engage in a legal cyber attack.[265] The principles of proportionality, distinction, and precautions in attack remain in effect, but the responsibility for failure lies squarely with the defending party.

5.3 Analyzing Some Means of Cyber Warfare from the Purview of International Humanitarian Law

Cyber warfare encompasses a broad spectrum of activities, ranging from relatively simple file access through psychological manipulation to more complex and destructive operations that can disrupt enemy combatants and civilians. This section provides a concise technical overview of some prominent methods employed in cyber warfare, including distributed denial of service (DDoS) attacks, hacking, and various cyber warfare suites and malware. Additionally, it examines whether these methods can conform to the limitations imposed by international humanitarian law or whether such laws even apply to them. In cases where the latter is true, the technical analysis offers insights into why this exemption exists.

DDoS Attacks (Distributed Denial of Service): DDoS attacks involve overwhelming a target's online services or website by flooding them with an excessive volume of traffic. This flood of traffic, generated by a network of compromised devices called botnets, can render the target's services inaccessible to users. DDoS attacks primarily disrupt online services and are often not considered attacks under international humanitarian law because they typically don't involve lasting realspace effects on physical objects or individuals. In essence, DDoS attacks are analogous to radio signal interference with a stronger signal. However, if a DDoS attack results in collateral damage or indirect harm to civilians or civilian infrastructure, it could raise legal questions on whether the operation crosses the threshold for attack.

264 *Ibid.* 489
265 *Ibid*. 491

Hacking: Hacking is the act of gaining unauthorized access to a computer or computer network. It encompasses various techniques used to gain unauthorized access to computer systems, networks, and data. These activities can range from stealing sensitive information to disrupting or manipulating systems. Hacking operations can potentially qualify as cyber attacks under international humanitarian law if they cause physical harm or direct damage to civilian infrastructure. The distinction depends on the nature and objectives of the hacking operation.

Cyber Warfare Suites and Malware: Cyber warfare suites and malware refer to software tools designed for offensive cyber operations. These tools can include viruses, worms, Trojans, and sophisticated malware packages. Their functionality can vary widely, from data theft to the destruction of critical systems. Whether these tools are subject to international humanitarian law depends on their specific use. If they are employed to cause harm to civilians or civilian objects, they could be categorized as cyber attacks and subject to legal scrutiny.

The applicability of international humanitarian law to these cyber warfare methods hinges on several factors, including the intent of the operation, the scale of damage or harm caused, and whether the operation meets the legal threshold for a cyber attack. Additionally, the evolving nature of cyber warfare poses challenges in applying traditional laws of armed conflict to this domain. Clearer legal frameworks and definitions are needed to address the unique characteristics of cyber warfare and provide guidance on how international humanitarian law applies to cyber operations.

5.3.1 DDoS Attack

The distinction between distributed denial of service (DDoS) attacks and cyber attacks that qualify as such under international humanitarian law is crucial and can be elucidated further.

A DDoS attack's primary objective is to overwhelm a target's online services with excessive internet traffic, effectively causing a traffic jam called denial of service.[266] This attack operates by establishing a botnet, a network of compromised internet-connected machines under the attacker's control. While DDoS attacks can disrupt online services and websites, they do not inherently pose a direct threat of physical damage to equipment or harm to individuals. This is a fundamental criterion for an action to be considered a cyber attack under international humanitarian law.

266 Cloudflare, 'What is DDoS Attack?' (Learning, 31 July 2017) <https://www.cloudflare.com/learning/ddos/what-is-a-ddos-attack/> accessed 16 May 2023

DDoS attacks often resemble a form of interference similar to jamming communication signals, such as radio interference. In both cases, the aim is to disrupt the normal functioning of a system without causing physical damage. Under international humanitarian law, acts that merely interfere with the functioning of systems or communications without posing a threat of harm or injury to civilians or civilian objects do not meet the legal criteria for an attack.

In essence, the key distinction lies in the potential for harm. DDoS attacks primarily disrupt online services but do not inherently threaten physical damage or harm to individuals. In contrast, cyber attacks that qualify as such under international humanitarian law involve actions that go beyond disruption and have the potential to cause direct harm or damage, whether physical or virtual, to civilians or civilian objects.

Therefore, the rules of international humanitarian law governing cyber attacks, such as the principles of distinction and proportionality, do not typically apply to DDoS attacks because DDoS attacks do not meet the threshold of posing a direct threat of harm or damage as required by these legal principles.

5.3.2 Hacking

Hacking, unlike distributed denial of service (DDoS) attacks, encompasses a wide range of activities aimed at unauthorized access to computer systems.[267] This section provides a closer look at hacking methods and their interaction with international humanitarian law.

Hacking involves exploiting vulnerabilities within a system to gain unauthorized access. The methods employed for such access vary considerably, with some showing no immediate signs of cyber warfare activity. The weakest link in the chain, in many cases, is the end user. Social engineering techniques are often employed to deceive users, with phishing links being a common method. Phishing links appear legitimate but lead to fraudulent websites designed to capture a user's credentials, essentially tricking the user into willingly providing access.

On a more technical level, hacking may involve cracking system passwords. Brute-force attacks, a technique where attackers systematically try various combinations to guess passwords, represent one approach.[268] The time required for such attacks varies based on factors like password complexity and the computing power used. More powerful hardware can significantly reduce cracking times. For

267 Kaspersky Lab, 'What is Hacking? And How to Prevent It' (Resource Center, 12 August 2022) <https://www.kaspersky.com/resource-center/definitions/what-is-hacking> accessed 17 May 2023

instance, a password with two billion possible combinations can take a powerful CPU more than two years to crack; add to this a powerful GPU, and the cracking time drops to 3.5 days.[269]

From an international humanitarian law perspective, hacking has several dimensions. The initial act of gaining unauthorized access to a system, akin to espionage or reconnaissance, often falls outside the limitations related to attacks. However, the legality of hacking is contingent on the actions taken after gaining access.

If the hacker's objectives are limited to data exfiltration or psychological warfare without causing physical harm, the act may not be considered an attack, and the law's limitations may not apply. For instance, hacking to display war footage on television channels as an act of psychological warfare does not inherently qualify as an attack. An example of this would be the hacking of Russian streaming services and TV channels to show them war footage from Ukraine by hackers sympathetic to Ukraine.[270]

However, if the hacker's intent is more malicious and results in damage akin to a kinetic attack, such as hacking into the enemy's air traffic control system and causing an accident or causing an accident at a critical factory like an ammunition plant, international humanitarian law and its constraints come into play. The reason why the principles of international humanitarian law become effective is because the effects of the attack cross the threshold of an attack by most of the rubrics examined above. The principles of distinction and proportionality, as well as precautions in attack, must be observed in such cases. Applied to the example above, this would mean that the hacker must verify that the air traffic control does not cause an accident to civilian aviation. The means chosen must be specific enough to target only legal targets. In the latter example, the hacker ought to be cognizant of the effects an explosion at an ammo plant can have on its surroundings.

The control of the hacking operation rests with the hacker, allowing for compliance with the legal principles of distinction as the hacker knows what system they are gaining access to; proportionality as the hacker must be cognizant of the reasonable effects of their action; and precautions in attack since the knowledge of the former two informs the hacker of possible spill-over effects. Consequently,

268 Kaspersky Lab, 'Brute Force Attack: Definition and Examples' (Resource Center, 10 April 2019) <https://www.kaspersky.com/resource-center/definitions/brute-force-attack> accessed 17 May 2023

269 *Ibid.*

270 Radio Free Europe, 'Russian TV Channels Hacked to Show Independent Coverage of War in Ukraine' (News, 7 March 2022) <https://www.rferl.org/a/russian-tv-hacked-ukraine-anonymous/31740663.html> accessed 17 May 2023

hacking can be conducted in a manner that aligns with the limitations set by international humanitarian law, provided the hacker adheres to these principles, meaning that hacking can be done in a manner that does not outright disqualify it as an option the belligerents have. In cases where the threshold of attack is not crossed, the limitations of international humanitarian law do not apply since no attack is taking place.

5.3.3 Cyber Weapons and Malware

The definition of a cyber weapon revolves around three key concepts: context, purpose, and means or tool.[271] When these elements are combined, a cyber weapon can be defined as "attacks via information technology, used to cause damage in a conflict situation." This comprehensive definition encompasses various forms and methods through which these attacks can be executed, including "parts of equipment, devices, or sets of computer instructions."[272]

Parts of equipment refer to components within devices like communications equipment, where manufacturers may have intentionally incorporated backdoors to facilitate easier access for national or military intelligence entities.[273] These backdoors can serve both non-destructive and destructive cyber operations.

Among the most notable cyber weapons are malware, which are malicious software programs designed to infiltrate systems and inflict damage.[274] Commonly referred to as worms or viruses, these are specific types of malware categorized by their methods of propagation. Worms can spread by themselves whereas viruses use a third file for propagation. There exists a wide array of malware forms, each with varying levels of command and control that operators can exert over them once they infiltrate a target. Some malware includes dedicated command and control modules that lie dormant until further instructions are received, while others propagate automatically but require additional commands for specific functions. Some malware operates independently from its operators.

---

271 Stefano Mele, 'Legal Considerations on Cyber-Weapons and Their Definition' (2014) 3 JL & Cyber Warfare 58

272 *Ibid.*

273 New York Times, 'NSA: Breached Chinese Servers Seen as Spy Peril' (Asian News, 23 March 2014) <https://www.nytimes.com/2014/03/23/world/asia/nsa-breached-chinese-servers-seen-as-spy-peril.html> accessed 2 November 2020

274 McAfee, 'What is the Difference Between Malware and a Virus?' (Enterprise Security, 1 July 2020) <https://www.mcafee.com/enterprise/en-us/security-awareness/ransomware/malware-vs-viruses.html> accessed 2 November 2020

Due to the extensive variety of cyber warfare suites and tools, it is impossible to provide a comprehensive analysis encompassing all of them. However, the more automatic in its propagation and execution the malware is, the closer it gets to being an indiscriminate weapon prohibited by international humanitarian law should the effects of its execution cross the threshold of an attack. This thesis focuses on a select few cyber warfare suites, providing a cursory examination of their potential compliance with the rules of international humanitarian law. The technical aspect of this analysis is based on original assessments conducted by cybersecurity laboratories and antivirus service providers, including Symantec, LogRhythm, and McAfee.

5.3.3.1 Stuxnet

Stuxnet stands out as one of the most prominent cyber weapons in history, having been used to target the Iranian nuclear program's centrifuges. Although its impact was felt in the physical world, Stuxnet provides an illustrative example of the challenges posed by worm-style malware and potential strategies to mitigate those challenges.

Stuxnet's ability to escape its original target network demonstrates the virulence of purpose-made cyber weapons.[275] This uncontrolled spread could potentially raise concerns about its indiscriminate nature, as indiscriminate weapons violate the principle of distinction. However, Stuxnet managed to avoid being categorized as indiscriminate by having a very specific set of targets it was designed to affect.[276] This specificity made it a highly targeted weapon with limited potential for causing damage beyond its intended scope.

Furthermore, Stuxnet included a built-in kill switch that triggered its self-termination on a specific date. This additional feature further restricted its destructive capabilities, as infections were limited to the timeframe before the kill switch's activation.

275 Ars Technica, 'Stuxnet Was Never Meant to Propagate in the Wild' (Tech Policy, 1 June 2012) <https://arstechnica.com/tech-policy/2012/06/confirmed-us-israel-created-stuxnet-lost-control-of-it/> accessed 2 November 2020

276 A specific set of Programmable Logic Controllers used in nuclear facilities, see Broadcom, 'Endpoint Protection' (Library Documents, 22 September 2010) <https://community.broadcom.com/symantecenterprise/communities/community-home/librarydocuments/viewdocument?DocumentKey=ad4b3d10-b808-414c-b4c3-ae4a2ed85560&CommunityKey=1ecf5f55-9545-44d6-b0f4-4e4a7f5f5e68&tab=librarydocuments> accessed 2 November 2020

In summary, Stuxnet serves as a prime example of both an effective cyber weapon and a cyber weapon that can align with the rules of international humanitarian law. The possibility of it spreading to other centrifuges containing the specified PLC and causing damage to them existed; however, it can be argued that such consequences are part of the defender's duty to protect against. All in all, the operation of infecting a closed network is distinct enough by itself, and a reasonable mind would expect the operators of the network to exercise good device hygiene. The operation is proportionate since it only caused the centrifuges to suffer mechanical failure without the risk of the uranium spreading outside the facilities. Furthermore, Stuxnet having a timed termination self-contains itself from spreading further, which is a necessary precaution while also serving the purpose of denying its code to the enemy.

5.3.3.2 Petya and NotPetya

Petya and NotPetya, both first observed in Ukrainian and Russian systems in the late 2010s, represent a unique category of cyber threats known as "wipers." These types of malware are designed with the primary purpose of destroying (or wiping) files on infected systems. While Petya was initially categorized as ransomware, malware that encrypts the files on the target system and demands payment, often in bitcoin, for an encryption key to undo the encryption, its true nature raised doubts about its motivations. Unlike traditional ransomware attacks, Petya lacked certain characteristics, such as the use of anonymous email services like Tor and a means to deliver encryption keys, which are typically associated with financially motivated attacks.[277] Instead, it appeared that Petya's encryption was designed to be permanent, suggesting alternative objectives beyond financial gain. While Petya does not fit the traditional definition of a wiper, it possesses the potential to cause damage that goes beyond the threshold of an attack should the encryption of data cause effects similar to kinetic attacks.

NotPetya, on the other hand, initially appeared to be a distinct malware entity but was later revealed to be a modified variant of Petya.[278] NotPetya masqueraded as ransomware, but its true intent was to irreversibly destroy the data it claimed to encrypt,[279] classifying it as a destructive cyber weapon. NotPetya exhibited a high level of virulence, capable of infecting an entire network of machines

277 Filip Truta, 'Everything you need to know about the Goldeneye/Petya attack' (BitDefender Security Blog, 28 June) <https://www.bitdefender.com/blog/hotforsecurity/everything-you-need-to-know-about-the-goldeneye-petya-attack/> accessed 18 May 2023

278 LogRhythm, NotPetya Technical Analysis (LogRhythm Labs 2017) 3

279 McAfee, 'What is Petya and NotPetya Ransomware?' (Enterprise Security, 22 May 2020) <https://www.mcafee.com/enterprise/en-us/security-awareness/ransomware/petya.html> accessed 2 November 2020

through a single entry point and adapting its infection path based on the specific anti-virus products present on target systems.[280]

Both Petya and NotPetya lacked the specificity of highly targeted cyber weapons like Stuxnet, and they exhibited a propensity to spread to any and all vulnerable systems.[281] This indiscriminate nature raises significant concerns about their legality under international humanitarian law. While they primarily targeted data, the potential for real-world damage and harm resulting from their uncontrolled spread is substantial. Consequently, it is reasonable to assert that these types of cyber weapons are incompatible with the rules of international humanitarian law, particularly due to their indiscriminate and highly destructive characteristics.

5.3.3.3 Drovorub

Modular malwares, such as the Drovorub software suite, represent a distinct category of cyber weapons that provide a high degree of control to their operators. Drovorub, specifically designed for infecting and controlling Linux-based machines, comprises four key modules: establishing a connection between the infected machine and the operator's command and control infrastructure; concealing the malware to evade detection; enabling port-forwarding capabilities for propagation; and facilitating file transfer and root access to the compromised system.[282] With root access and remote shell capabilities, the operator gains complete control over the infected system.

The crucial distinction with modular malware like Drovorub is that the operator has full control over its use. Consequently, it cannot be classified as an indiscriminate weapon, even if it propagates to multiple machines without inherent limitations. The mere act of gaining access to a system, as discussed in the definition of a cyber attack, does not inherently constitute an attack under international humanitarian law. Whether the operation exceeds the threshold of an attack largely depends on the actions taken by the operator and the realspace effects of them.

In general, modular malware like Drovorub has the capacity to operate within the limitations established by international humanitarian law, provided that the operator adheres to those constraints.

280 LogRhythm (n 277) 5

281 McAfee Enterprise, 'New Variant of Petya Ransomware Spreading Like Wildfire' (McAfee Labs, 27th June) <https://www.mcafee.com/blogs/mcafee-labs/new-variant-petya-ransomware-spreading-like-wildfire/> accessed 11 February 2020

282 National Security Agency and Federal Bureau of Investigations, Russian GRU 85th GTsSS Deploys Previously Undisclosed Drovorub Malware (United States' Department of Justice 2020) 1-2

The key factor in compliance with these rules is the operator's behavior and intent when controlling the infected systems, as the malware itself merely serves as a tool that can be employed for various purposes, both legal and illegal. In essence, suites like Drovorub are in the same situation as hacking; their legality is solely dependent on the actions of the operator and their consequent effects. Alone, the constraints of international humanitarian law do not apply to these suites, as they only provide access to the system, which by itself has no effect outside of it.

6.0 Conclusions

The conclusion of the thesis at hand is that, both *jus in bello* and *jus ad bellum*, current international humanitarian law applies to cyber warfare and puts on similar restrictions as to kinetic warfare. What could limit the protection given by the law is wholly contingent on the understanding of the word "attack", which the Tallinn Manual and all three states observed define via effects-based analysis; the cyber operation becomes a cyber attack via having similar consequences as traditional, kinetic attacks. Such a definition creates an incongruence between the lay understanding of the word and the legal understanding. The reasoning behind this is sound, as it is based on the nearest analogs outside cyberspace. Cyber operations consisting of only DDoS attacks on an internet service are tantamount to radio interference, and operations involving only data exfiltration are cyber espionage rather than cyber warfare.

The principles of distinction, proportionality, and precautions in attack apply to all cyber attacks. The belligerents must use all reasonable measures to clarify the status of the target. When weighing the direct military advantage of an operation, overt wanton destruction must be avoided. Indiscriminate cyber attacks are as prohibited as their kinetic counterparts. When deciding to respond to a cyberarmed attack, the necessity of the operation must be calculated. The limitations on cyber operations are not weakened by the dual-use nature of the internet. The calculations of distinction and proportionality are made harder by this, but never frees the cyber operatives and their commanders from their international humanitarian law obligations.

On the matter of how international humanitarian law limits some of the common forms of cyber warfare, the law does not outright prohibit any form it might take. DDoS attacks do not qualify as attacks as they do not, traditionally, have physical consequences. Therefore, they are not limited by international humanitarian law but are inconveniences the people, both civilian and military, must endure. The limitations of hacking are fully dependent on the consequences the tampering with the breached system has. While hacking can take the form of cyber espionage, as noted above, it can have deadly consequences that the hacker in control of the system must be cognizant of and is liable for. As per the issue of cyber weapons, international humanitarian law has the most to limit. Cyber weapons can be both indiscriminate and unproportional, and as seen, they are hard to control even in a closed network. It is clear that the law prohibits the use of such weapons as they are unable to follow the

principles of distinction and proportionality. However, cyber weapons can be made specific to the target and self-destruct in particular circumstances, making them an allowed form of cyber warfare again. An example from the realm of kinetic weapons would be anti-personnel mines. By their nature, they are incapable of distinguishing between civilians and combatants. It is common to mine an area using artillery-fired vehicles that randomly spread an area full of mines, making it hostile to life and unfit for human use until the time of mine clearance. However, more modern anti-personnel mines can be set to explode after a certain time, making their effect on an area temporary.

The subquestion of the objecthood of data also has a clear answer. At the time of writing, data itself, by default, is not considered an object under international humanitarian law. The understanding of its not-objecthood is based on its intangibility and invisibility. Some datasets, such as medical data, create an exception; however, the protection is not specific to the data itself but relies on the more general protection given to specific classes. The closest current law comes to protecting pure data is in the case of digital cultural property. The current understanding is not without its critics. However, their arguments are often based on the evolving meaning of terms and, as such, can be categorized more as *lege ferenda*. As international law currently stands, data does not have objecthood.

The current state of things in which international humanitarian law directly applies to cyber warfare is the balance of military necessity and civilian need for protection. The possible weaknesses of limiting understanding of attacks only through physical consequences are not well pertinent in the current state. There are only few cases, such as that of digital cultural property, where protection of data as such would be immediately warranted. However, the *lege ferenda* may see to it that this understanding is expanded as the gamut of human activities further digitizes. It remains to be seen how willing the states are to accept the changes if and when they come.

The thesis has been partially successful in answering the questions laid out in its introduction. *Lege lata* has been fairly thoroughly explored, which was the main goal of the thesis. However, the states' practice on the matter was explored only in a cursory manner. State practice is also the focus of the revised Tallinn Manual being collated at the time of writing. The matter of artificial intelligence was wholly outside the purview of the thesis. Further research, and revised research will be needed.

Bibliography

Books and Journal Articles

Blank, Laurie R., 'Irreconcilable Differences: The Thresholds for Armed Attack and International Armed Conflict' (2020) 96 Notre Dame L Rev 249

Bobrowski, Krzysztof, 'Conventional Attack vs Digital Attack in the Light of International Law' (2021) 10 Polish Rev Int'l & Eur L 77

Bormashenko, Edward, 'The Landauer Principle: Re-Formulation of the Second Thermodynamics Law or a Step to Great Unification?' [2019] 21(10) Entropy 918

Brandis, George, 'The Right of Self-Defense against Imminent Armed Attack in International Law' (2017) 35 Aust YBIL 55

Burkadze, Khatuna, 'A Shift in the Historical Understanding of Armed Attack and Its Applicability to Cyberspace' (2020) 44 Fletcher F World Aff 33

Cerf, Vinton G., 'Avoiding "Bit Rot": Long-Term Preservation of Digital Information' [2011] 99(6) Proceedings of the IEEE 915-916

Chang, Zen, 'Cyberwarfare and International Humanitarian Law' (2017) 9 Creighton Int'l & Comp LJ 29

Clausewitz, Carl von, On War (Princeton University Press c1976)

Dinniss, Heather A. Harrison 'The Nature of Objects: Targeting Networks and the Challenge of Defining Cyber Military Objectives' (2015) 48 Isr L Rev 39

Egan, Brian J., 'International Law and Stability in Cyberspace' [2016] 35(1) Berkeley Journal of International Law 169-180

Farhat, Salem Aessa and others, 'Attacks Against Civilian Objects: An Analysis Under International Humanitarian Law' [2022] 8(1) Hasanuddin Law Review 60-78

Greulich, Christopher and Jensen, Eric Talbot, 'Cyber Pillage' (2020) 26 Sw J Int'l L 264

Guerrero-Saade, Juan Andres, and others, Penquin’s Moonlit Maze: The Dawn of Nation-State Digital Espionage (Kaspersky Lab 2015)

Hagestad, William T., 21st Century Chinese Cyberwarfare (IT Governance Publishing 2012)

Hayashi, Nobuo, 'Contextualizing Military Necessity' (2013) 27 Emory Int'l L Rev 189

Król, Karol and Zdonek, Dariusz, 'Peculiarity of the bit rot and link rot phenomena' [2020] 69(1) Global Knowledge, Memory and Communication 20-37

Landauer, Rolf, 'Irreversibility and Heat Generation in the Computing Process' [1961] 5(3) IBM Journal of Research and Development 183-191

LogRhythm, NotPetya Technical Analysis (LogRhythm Labs 2017)

Lomans, Marieke, "Investigating Titan Rain" Cyber Security & Cyber Operations (2017)

Macak, Kubo, 'Military Objectives 2.0: The Case for Interpreting Computer Data as Objects under International Humanitarian Law' (2015) 48 Isr L Rev 55

Mavropoulou, Elizabeth, 'Targeting in the Cyber Domain: Legal Challenges Arising from the Application of the Principle of Distinction to Cyber Attacks' (2015) 4 JL & Cyber Warfare 23

McCormack, Tim, 'International Humanitarian Law and the Targeting of Data' (2018) 94 Int'l L Stud Ser US Naval War Col 221

Mele, Stefano, 'Legal Considerations on Cyber-Weapons and Their Definition' (2014) 3 JL & Cyber Warfare 52

Ministry of foreign affairs of finland, International Law and Cyberspace Finland's National Positions (2020)

National Security Agency and Federal Bureau of Investigations, Russian GRU 85th GTsSS Deploys Previously Undisclosed Drovorub Malware (United States' Department of Justice 2020)

Pomson, Ori ‘Objects’? The Legal Status of Computer Data under International Humanitarian Law (2023)

Ottis, Rain, Analysis of the 2007 Cyber Attacks Against Estonia from the Information Warfare Perspective (NATO Cooperative Cyber Defence Command of Excellence 2018)

Yaroslav Radziwill, Cyber-attacks and the Exploitable Imperfections of International Law (Brill Nijhoff 2015)

Ruys, Tom, Armed Attack and Article 51 of the UN Charter: Evolutions in Customary Law and Practice (Cambridge University Press 2013)

Sassòli, Marco, Legitimate Targets of Attack under International Humanitarian Law (International Humanitarian Law Research Initiative 2003)

Schmitt, Michael N., “Attack” as a Term of Art in International Law: The Cyber Operations Context. in Czosseck and others (eds), 4th International Conference on Cyber Conflict (NATO Cooperative Cyber Defence Command of Excellence 2012)

Schmitt, Michael N., 'Notion of Objects during Cyber Operations: A Riposte in Defence of Interpretive and Applicative Precision' (2015) 48 Isr L Rev 81

Schmitt, Michael N., Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations (2nd edn, Cambridge University Press 2017)

Schmitt, Michael N., 'Wired Warfare 3.0: Protecting the Civilian Population during Cyber Operations' (2019) 101 Int'l Rev Red Cross 333

Selwyn, Neil, 'Data entry: towards the critical study of digital data and education' [2015] 40(1) Learning, Media and Technology 64-82

Stiennon, Richard, A Short History of Cyber Warfare. in James A. Green (ed), Cyber Warfare: A Multidisciplinary Analysis (Routledge 2015)

Sumanadasa, W.A.D.J., 'Principle of Proportionality: The Criticized Comprising Formula of International Humanitarian Law' (2010) 10 ISIL YB Int'l Human & Refugee L 21

Todd, Graham H., 'Armed Attack in Cyberspace: Deterring Asymmetric Warfare with an Asymmetric Definition' (2009) 64 AF L Rev 65

Tsagourias, Nicholas, 'Cyber Attacks, Self-Defence and the Problem of Attribution' (2012) 17 J Conflict & Sec L 229

Tzu, Sun, Art of War (Allandale Online Publishing c2000)

Watts, Sean & Richard, Theodore, 'Baseline Territorial Sovereignty and Cyberspace' (2018) 22 Lewis & Clark L Rev 771

Zamir, Noam 'Distinction Matters: Rethinking the Protection of Civilian Objects in Non-International Armed Conflicts' (2015) 48 Isr L Rev 111

International Treaties

1980 Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons which may be deemed to be Excessively Injurious or to have Indiscriminate Effects (adopted 10 October 1980 entered into force 2 December 1983) 1342 UNTS 22495

1977 Protocol Additional to the Geneva Conventions of 12 August 1949 and relating to the protection of victims of non-international armed conflicts (adopted 8 June 1977 entered into force 7 December 1978) 1125 UNTS 17513

1977 Protocol additional to the Geneva Conventions of 12 August 1949, and relating to the protection of victims of international armed conflicts (adopted 8 June 1977 entered into force 7 December 1978) 1125 UNTS 17512

1969 Vienna Convention on the Law of Treaties (adopted 23 May 1969 entered into force 27 January 1980) 1155 UNTS 18232

1949 Geneva Convention relative to the treatment of prisoners of war (adopted 12 August 1949 entered into force 21 October 1950) 75 UNTS 972

1949 Geneva Convention for the amelioration of the condition of the wounded and sick in armed forces in the field (adopted 12 August 1949 entered into force 21 October 1950) 75 UNTS 970

1945 Charter of the United Nations (adopted 26 June 1945 entered into force 24 October 1945)

International Case Law

2009 *Dispute regarding Navigational and Related Rights (Costa Rica v. Nicaragua)*, Judgment, ICJ Reports 2009

2007 *Case Concerning Application of the Convention on the Prevention and Punishment of the Crime of the Crime of Genocide (Bosnia and Herzegovina v. Serbia and Montenegro)*, (Merits, Judgment, 26 February 2007), ICJ Reports 2007

2003 *Case Concerning Oil Platforms (Islamic Republic of Iran v. United States of America)*, (Merits, Judgment, 6 November 2003), ICJ Reports 2003

1996 *Legality of the Threat or Use of Nuclear Weapons*, (Advisory Opinion, 8 July 1996), ICJ Reports 1996

1986 *Case Concerning Military and Paramilitary Activities in and Against Nicaragua (Nicaragua v. United States of America)*, (Merits, Judgment, 27 June 1986), ICJ Reports 1986

Documents from International Organizations

2021 UNGA Res 76/136 (13 July 2021) UN Doc A/76/136

2001 ILC, 'Draft Articles on Responsibility of States for Internationally Wrongful Acts' (2001), UN Doc A/56/10

2001 UNSC Res 1373 (2001) UN Doc S/RES/1373

2001 UNSC Res 1368 (2001) UN Doc S/RES/1368

1974 UNGA Res 3314 (XXIX) (14 December 1974)

1964 ILC, 'Third Report on the Law of Treaties, by Sir Humphrey Waldoc, Special Rapporteur' (1964), UN Doc A/CN.4/167 and Add.1-3

Other Material

Ars Technica, 'Stuxnet Was Never Meant to Propagate in the Wild' (Tech Policy, 1 June 2012) <https://arstechnica.com/tech-policy/2012/06/confirmed-us-israel-created-stuxnet-lost-control-of-it/> accessed 2 November 2020

Broadcom, 'Endpoint Protection' (Library Documents, 22 September 2010) <https://community.broadcom.com/symantecenterprise/communities/community-home/librarydocuments/viewdocument?DocumentKey=ad4b3d10-b808-414c-b4c3-ae4a2ed85560&CommunityKey=1ecf5f55-9545-44d6-b0f4-4e4a7f5f5e68&tab=librarydocuments> accessed 2 November 2020

Cloudflare, 'What is DDoS Attack?' (Learning, 31 July 2017) <https://www.cloudflare.com/learning/ddos/what-is-a-ddos-attack/> accessed 16 May 2023

Daniel S. Katz, 'Software vs Data' (Software vs Data Open Article, 9 December 2016) <https://github.com/danielskatz/software-vs-data> accessed 22 February 2023

Filip Truta, 'Everything you need to know about the Goldeneye/Petya attack' (BitDefender Security Blog, 28 June) <https://www.bitdefender.com/blog/hotforsecurity/everything-you-need-to-know-about-the-goldeneye-petya-attack/> accessed 18 May 2023

IBM, 'Introduction' (Quantum Computing, 5 December 2023) <https://learning.quantum.ibm.com/tutorial/explore-gates-and-circuits-with-the-quantum-composer> accessed 15 April 2024

International Committee of the Red Cross, 'Definition of attacks and scope of application' (Commentary of 1987, 14 March 2023) <https://ihl-databases.icrc.org/en/ihl-treaties/api-1977/article-49/commentary/1987?activeTab=undefined> accessed 15 April 2024

International Committee of the Red Cross, 'General Protection of Civilian Objects' (Commentary of 1987, 16 March 2023) <https://ihl-databases.icrc.org/en/ihl-treaties/api-1977/article-52/commentary/1987?activeTab=undefined> accessed 15 April 2024

International Committee of the Red Cross, 'Precautions in Attack' (Commentary of 1987, 1 June 2023) <https://ihl-databases.icrc.org/en/ihl-treaties/api-1977/article-57/commentary/1987activeTab=undefined> accessed 15 April 2024

International Committee of the Red Cross, 'Property of Prisoners' (Commentary of 2020, 14 December 2022) <https://ihl-databases.icrc.org/en/ihl-treaties/gciii-1949/article-18/commentary/2020?activeTab=undefined> accessed 15 April 2024

International Committee of the Red Cross, 'Target Selection' (Customary International Humanitarian Law, 28 January 2023) <https://ihl-databases.icrc.org/en/customary-ihl/v1/rule21> accessed 15 April 2024

Kaspersky Lab, 'Brute Force Attack: Definition and Examples' (Resource Center, 10 April 2019) <https://www.kaspersky.com/resource-center/definitions/brute-force-attack> accessed 17 May 2023

Kaspersky Lab, 'What is Hacking? And How to Prevent It' (Resource Center, 12 August 2022) <https://www.kaspersky.com/resource-center/definitions/what-is-hacking> accessed 17 May 2023

Kaspersky Lab, 'What is WannaCry Ransomware?' (Resource Center, 20 April 2020) <https://www.kaspersky.com/resource-center/threats/ransomware-wannacry> accessed 11 May 2020

McAfee Enterprise, 'New Variant of Petya Ransomware Spreading Like Wildfire' (McAfee Labs, 27th June) <https://www.mcafee.com/blogs/mcafee-labs/new-variant-petya-ransomware-spreading-like-wildfire/> accessed 11 February 2020

McAfee, 'What is Petya and NotPetya Ransomware?' (Enterprise Security, 22 May 2020) <https://www.mcafee.com/enterprise/en-us/security-awareness/ransomware/petya.html> accessed 2 November 2020

McAfee, 'What is the Difference Between Malware and a Virus?' (Enterprise Security, 1 July 2020) <https://www.mcafee.com/enterprise/en-us/security-awareness/ransomware/malware-vs-viruses.html> accessed 2 November 2020

Nato Cooperative Cyber Defence Centre of Excellence, 'Cyber Attacks Against Estonia (2007)' (Cyber Law, 15 October 2018) <https://cyberlaw.ccdcoe.org/wiki/Cyber_attacks_against_Estonia_(2007)> accessed 13 February 2024

Nato Cooperative Cyber Defence Centre of Excellence, 'Cyber operations against government systems in Ukraine (January 2022)' (Cyber Law, 28 February 2022) <https://cyberlaw.ccdcoe.org/wiki/Cyber_operations_against_government_systems_in_Ukraine_(January_2022)> accessed 14 February 2024

Nato Cooperative Cyber Defence Centre of Excellence, 'HermeticWiper malware attack (2022)' (Cyber Law, 27 May 2022) <https://cyberlaw.ccdcoe.org/wiki/HermeticWiper_malware_attack_(2022)> accessed 14 February 2024

Nato Cooperative Cyber Defence Centre of Excellence, 'NotPetya (2017)' (Cyber Law, 15 October 2018) <https://cyberlaw.ccdcoe.org/wiki/NotPetya_(2017)> accessed 14 February 2024

Nato Cooperative Cyber Defence Centre of Excellence, 'Power grid cyberattack in Ukraine (2015)' (Cyber Law, 20 December 2018)

<https://cyberlaw.ccdcoe.org/wiki/Power_grid_cyberattack_in_Ukraine_(2015)> accessed 14 February 2024

Nato Cooperative Cyber Defence Centre of Excellence, 'Stuxnet (2010)' (Cyber Law, 18 December 2018) <https://cyberlaw.ccdcoe.org/wiki/Stuxnet_(2010)> accessed 14 February 2024

Nato Cooperative Cyber Defence Centre of Excellence, 'The Shadow Brokers publishing the NSA vulnerabilities (2016)' (Cyber Law, 20 December 2018) <https://cyberlaw.ccdcoe.org/wiki/The_Shadow_Brokers_publishing_the_NSA_vulnerabilities_(2016)> accessed 14 February 2024

Nato Cooperative Cyber Defence Centre of Excellence, 'Use of malware to track and target Ukrainian artillery units (2014-2016)' (Cyber Law, 29 August 2023) <https://cyberlaw.ccdcoe.org/wiki/Use_of_malware_to_track_and_target_Ukrainian_artillery_units_(2014-2016)> accessed 14 February 2024

New York Times, 'NSA: Breached Chinese Servers Seen as Spy Peril' (Asian News, 23 March 2014) <https://www.nytimes.com/2014/03/23/world/asia/nsa-breached-chinese-servers-seen-as-spy-peril.html> accessed 2 November 2020

Radio Free Europe, 'Russian TV Channels Hacked to Show Independent Coverage of War in Ukraine' (News, 7 March 2022) <https://www.rferl.org/a/russian-tv-hacked-ukraine-anonymous/31740663.html> accessed 17 May 2023

Royal Navy, 'Cyber Operative' (Job Role, 22 September 2020) <https://www.royalnavy.mod.uk/careers/roles/cyber-operative> accessed 30 January 2024

West point lieber institute, 'The Definition of an “Attack” under the Law of Armed Conflict' (Articles of War, 3 November 2020) <https://lieber.westpoint.edu/definition-attack-law-of-armed-conflict-protection/> accessed 8 January 2023

Yale Law School, 'The Caroline' (Yale Avalon Project, 19 April 2009) <https://avalon.law.yale.edu/19th_century/br-1842d.asp#ash1> accessed 20 November 2023