Corinne Johnston
Cybersecurity Ethics Capstone
Fall 2019

Applying Just War Theory and International Humanitarian Law to Facilitate Ethical Cyberwarfare

Introduction

The world has entered the digital age and ventured into cyberspace. This development has changed almost every aspect of the world, from education to healthcare to entertainment. Warfare is not an exception to this transformation. Modern technology has transformed the way cyberspace and national defense are conceptualized, leading former CIA Chief Michael Hayden to declare cyberspace the fifth domain of warfare, along with the traditional domains of land, sea, air, and space (Finkelstein ix). Nations have begun to focus on improving their technical skills, causing some to claim that we are currently in "a cyberarms race reminiscent of the nuclear arms race of the Cold War" (Finkelstein xii). Despite this dramatic increase of international interest in cyberwar, there is not a clear consensus on how it should be governed (Allhoff).

Traditional warfare is regulated by international treaties and agreements in order to protect the innocent and prevent unnecessary suffering. These rules are collectively referred to as International Humanitarian Law (IHL) and have been deeply influenced by the Just War Theory (JWT), an ethical framework developed over centuries to help define what makes a war morally permissible or, as the name suggests, just (Emba). Over recent years, there has been much debate over whether these existing systems can be effectively applied to cyberwarfare or if it differs too much from kinetic warfare. However, both IHL and JWT focus more on the intentions and effects of attacks and less so on the means through which the attack is performed (Orend 13).

This allows these traditions to be successfully applied to new means of warfare, such as cyberwar. Though it is possible to utilize the traditional guidelines, additional thought still needs to be given to how exactly the rules can be applied to cyberwar. This application needs to be defined and accepted by the international community in order to successfully regulate cyberwar. Various experts in both the technology and military field have proposed possible ways to achieve this, some of which are more realistic than others.

What is Cyberwarfare?

In order to understand the ethical questions related to cyberwar, one must understand what cyberwar is. Cyberwar is when a nation employs "the aggressive use of advanced computer technologies in a way deliberately designed to substantially harm" another nation for a political objective. (Orend 4). Harm can be financial damage, physical destruction, or denial of resources and is substantial when it reaches the level of destruction that could be caused by a physical attack. This definition is appropriate because it considers both the intent and consequences of the act, meaning that only cyber incidents that are consciously created to threaten peace will be deemed acts of war. This is correct because declaring an attack an act of war should not be done lightly as it gives the targeted nation the right to respond, thus bringing about war between nations. Because of this gravity, many cyberattacks do not fall under the definition of cyberwarfare; these are attacks do not reach the threshold of cyberwar because they are not deliberate, not sponsored by a state, or do not cause substantial harm. This also means that attacks such as cyberespionage, cybercrime, or information warfare (disinformation and propaganda campaigns) do not constitute cyberwar. Though no nation has ever officially declared cyberwar on another state, there have been numerous attacks in the past few decades

that can be classified as acts of cyberwar. These attacks have demonstrated the significant damage that cyberwar makes possible.

One of the most prominent acts of cyberwarfare occurred in 2010. It began when more than a thousand centrifuges used in Iranian nuclear facilities self-destructed with no obvious cause (McAfee). This damaged the machines beyond recovery and baffled Iranian scientists. Eventually, the cause of the destruction was traced back to a computer worm that spread through Windows computers (IEEE Spectrum). The virus, named Stuxnet, was immense and sophisticated, suggesting its creation must have been supported by a nation-state. Two years later, experts finally traced its development back to a joint effort between the United States and Israel, though neither nation has confirmed this (IEEE Spectrum). No matter who created it, Stuxnet was revolutionary; it was one of the earliest attacks to result in physical damage. Stuxnet can be classified as an act of cyberwar because it was almost certainly backed by a nation, caused serious physical damage akin to what would be accomplished by a physical attack, and served the political purpose of considerably reducing Iran's nuclear ability.

Another significant attack (sometimes referred to as "Web War I") was in 2007 when Estonia decided to move a statue celebrating the Russian military from the center of Tallinn, Estonia's capital city (Davis). This enraged many Russians, including those high up in the Russian government. Within days, Estonia was hit with a devastating attack that effectively shut down the majority of the country's internet (Davis). The attack focused primarily on essential online services such as banks, the government, and media sites (McGuinness). This was particularly disastrous for Estonia because of its emphasis on being an "e-government" (Davis). Eventually, the cyberattack ended after several weeks of the country being under digital siege. The Russian attack on Estonia served the sole purpose of making a political statement and

resulted in serious harm through the denial of access to the country's critical infrastructure, making it an act of cyberwar.

Then, in 2012, another act of cyberwar was carried out. Iran targeted Saudi Aramco, an oil company owned by the Saudi Arabian government (Perlroth). The Shamoon virus erased the data of two-thirds of the company's computers. The company was forced to shut down its internal network and it took a week to recover basic services and several more to recover fully from the attack (Perlroth). Saudi Aramco is one of the world's most valuable companies and the attack had a significant financial impact (Perlroth). This attack did considerable financial damage and was motivated by the Iranian's objection to the Saudi Arabian government, allowing the attack to be classified as cyberwar.

These are just a few examples of the many attacks that have occurred in the past few decades. As nations become more technically advanced, the level of sophistication and destruction rises. The frequency of acts of cyberwar is also rising. Such attacks are most commonly used by the US, China, Russia, Great Britain, France, India, Israel, and Pakistan (Finkelstein xii). However, they are not limited to these countries. In 2007, security firm McAfee estimated that 120 countries had developed ways to use the internet to target financial markets, government computer systems, and utilities (Finkelstein xii).

One of the reasons that cyberwar is appealing to nations is its differences from traditional kinetic warfare. Cyberwar allows for attacks from a physical distance: "cyber represents the complete loss of the physical battlefield" (Finkelstein xiv). This means significantly lower risk for the attacker because attacks do not have to put any of their own combatants in danger. However, the nature of cyberwar also increases the impact on civilians; a large portion of attacks purposely target essential infrastructure in order to cause widespread damage (Lucas 74).

Traditional Law of War

This disproportionate effect on civilians is one of the many moral dilemmas surrounding cyberwarfare that make it clear that the ethical frameworks Just War Theory (JWT) and International Humanitarian Law (IHL) need to be adhered to. There has been much more focus on the application of JWT to cyberwarfare, even though both frameworks are important and only IHL is legally recognized. This could be due to the substantial overlap between the two; IHL is heavily influenced by JWT. JWT is a tradition of ethics that can be traced back as far as ancient Egypt (Emba). It requires that a war meet certain criteria in order to be classified as just. It is split into two main categories: Jus ad Bellum (the conditions under which a nation is justified to go to war) and Jus in Bello (the limitations under which a nation must operate during a war) (Emba).

Jus ad Bellum requires that war must be publicly declared by the proper authority and that the nation must have the right intention, as well as a probability of success. The only motive should be to "resist, repulse, and punish aggression" (Orent 13). A nation must also have a just cause. There are only three scenarios in which this requirement is met. The most straightforward is self-defense from aggression, where aggression is defined as "any unjustified use of force against another country" (Orent 11). The classic example is an unprovoked armed attack that crosses a national border. Aggression authorizes war because it violates people's most basic right of being able to live their lives peacefully (Orent 11). Nations also have just cause when they are protecting collective security, such as when a nation goes to war to assist another country that suffered an act of aggression. The only other time a nation meets the just cause criteria is when it has the prior approval of the United Nations Security Council to use force against another nation.

War must be proportional. This requirement is to encourage world leaders to think about any other existing ways to achieve the same goal without resorting to violence. In order for a war to be proportional, the situation must be severe enough that the only effective and equivalent response is war. This means that the reason for the war must be proportional to the expected harm that the war will inevitably cause. War is horrific; therefore, there are very few cases where the appropriate response to a situation is war. Once a nation is engaged in war, the harm of all military maneuvers must be less than or, at most, equal to the benefit. All attacks must directly serve the purpose of winning the war. It is not just to cause suffering and risk the lives of both nations' combatants in exchange for a minor and meaningless victory.

A nation must have exhausted all other potential solutions before it declares war, making warfare a last resort. This means attempting to utilize the other three "basic tools in their foreign policy tool-box: diplomacy, economic incentives, and sanctions" in order to resolve the tension before declaring war. (Orent 12). War is just only in the case where all three other tools fail.

The nation cannot utilize any means or weapons that are malum in se. This translates to "wrong or evil in itself" (Orent 15). There are certain actions and weapons that the international community have deemed inherently evil such as nuclear weapons, rape, and biological weapons.

In order to be a just war, nations must distinguish between combatants and civilians. This means that all attacks must focus on a legitimate target that is part of "the military-industrial-political complex which guides the war and fights it" (Orent 16). Attacks on such targets would damage the opposing side's ability to fight the war. Examples of legitimate targets are soldiers, military equipment, means of transportation, supply and communication lines, etc. Attacks cannot be directed against illegitimate targets which are things that are not "engaged in military supply, or military activity" (Orent 16). Illegitimate targets are things such as civilians, hospitals,

farms, non-military industrial sites, residential areas, etc. This also applies to prisoners of wars; they must be treated fairly and cannot be made to fight against their own side. The principle of distinction does not require that no illegitimate target be impacted by an attack, but instead requires that all possible efforts were made to protect and avoid illegitimate targets. Targets that are utilized by both the opposing nation's military and civilians, such as bridges, radio/TV networks, etc. are called dual-use targets. They are forbidden from being targeted, but historically this rule has not been obeyed. However, it is much more controversial and criticized to target basic and critical infrastructure such as food supply, power grids, etc.

Under Just War Theory, a nation must meet all of the above requirements for the entirety of the war in order to wage a just war. There is no consensus on if this has ever been truly achieved or if every war ever fought has been unjust under JWT (Orent 13). The requirements exist to serve both as guidelines to the proper and moral behavior in regard to war and as a reminder of the horrors of war that all nations entering into conflict inflict and suffer from.

Like JWT, International Humanitarian Law "is designed to minimize human suffering in war" (United Nations). IHL's roots date back to the ancient civilizations, but the written, internationally recognized IHL originated in the 19th century (United Nations). The majority of the IHL is contained in the four Geneva Conventions of 1949 which were supplemented in 1977. Almost every single country in the world has agreed to be bound by these laws (United Nations). IHL pertains only to armed conflict and applies equally to all sides of the conflict, regardless of who the aggressor was. IHL incorporates all of the criteria of JWT, except for right intention and probability of success. This is because both are very difficult to prove and the principle of probability of success is biased towards powerful nations. IHL serves to protect non-combatants (civilians, medical and religious military personnel) as well as those who can no longer

participate in the fighting (prisoners of war, wounded soldiers, etc.). It requires that all these groups be treated humanely and bans the wounding or killing of anyone who falls under these categories. IHL also places restrictions on the acceptable weapons and tactics of war: "The right of parties in an armed conflict to choose methods and means of warfare is not unlimited" (United Nations). Weapons that do not discriminate between legitimate and illegitimate targets, cause excessive suffering, or cause long-lasting damage to the environment are forbidden under IHL. Violations of IHL are considered war crimes and can result in individuals being tried and punished for their involvement. Though IHL's relation to cyberwarfare is less commonly examined than JWT's, IHL is still an essential part of the discussion because it is already recognized on an international level. Therefore, IHL is much easier to enforce than JWT.

How Traditional Law of War Applies to Cyberwarfare

Regulations and norms surrounding cyberwarfare need to adhere to the JWT and IHL in order to promote the most moral and least harmful warfare possible. Understanding JWT and IHL makes it possible to attempt to determine how they should be applied to cyberwarfare. The biggest challenge that arises from this is defining what constitutes aggression in cyberwar. Aggression in traditional war occurs when a nation violates another's sovereignty through force. This can be difficult to conceptualize in cyberspace. However, as cyberwarfare has advanced, the damage cyberwar can cause equals or even surpasses that of kinetic warfare. In 2011, the Pentagon declared cyberattacks from a foreign nation an act of war because they "increasingly serve the function that kinetic attacks have historically served" (Finkelstein xv). The Pentagon did not give a clear definition of what they consider an act of war in cyberspace and instead reserved the right to decide on "a case-by-case and fact-specific basis" (Aftergood). However, they did clarify that they would particularly examine attacks that "result in a significant loss of

life, injury, destruction of critical infrastructure, or serious economic impact" (Aftergood). Accepting cyberwar as a form of aggression leads to the question of proportionality. Specifically, what does a proportionate response to cyberwar look like? Does it allow the targeted nation to respond with either their own cyberattack or a kinetic attack? There is still much debate about this, but a growing portion of experts agree that an attack in cyberspace does allow for a kinetic attack because they can "produce the death, damage, destruction or high-level disruption that a traditional military attack would cause" (Gorman). Or, as one US military official put it: "If you shut down our power grid, maybe we will put a missile down one of your smokestacks" (Gorman). One exception to this should be noted: a cyberattack that slows down a nonessential service but does not deny it entirely does not constitute aggression and therefore does not justify a kinetic response. This is because it does not block access to a resource and therefore its damage cannot be equivalent to that of a kinetic attack. However, if the attack denies a critical service, such as power, for a substantial amount of time, the attack would result in a loss of life equivalent to a physical attack and would be classified as an act of war.

Another significant question about ethical cyberwar is how to distinguish between civilians and combatants, as is required by both JWT and IHL. The majority of cyberwar incidents have disproportionally affected civilians, which makes the attack unjust. For example, the Russian attack on Estonia affected every civilian in the country and was not aimed towards a legitimate target (Davis. This focus on damaging infrastructure is a trend in cyberwar, particularly for acts of cyberwar between Russia and America (Cheravitch). Intentionally causing disproportionate damage to civilians and critical infrastructure violates JWT and IHL because it does not follow the principle of distinction and causes unnecessary suffering to innocents. Therefore, no attacks of this nature should be permitted, even though this includes the majority

of current cyberwar. However, it is possible to have a successful attack that does not affect civilians; for example, Stuxnet purposefully avoided affecting any computer that was not being run in a nuclear facility (McAfee). This extra effort put in by Stuxnet's developers guaranteed that no civilians were unfairly affected.

The norms of cyberwar as they stand today also violate the principle of public declaration as well as last resort. In the history of cyberwarfare, there has never been an act of cyberwar that was publicly announced beforehand. This is because the vast majority of attacks rely on exploiting a bug in software and prior warning would give the target the chance to fix the bug, rendering the attack useless. Also, most cyberwar attacks have not been the last resort. Instead, the international norm is to use them as warnings or displays of power and anger. For example, the Russians did not attempt any form of diplomacy, incentives, or sanctions to prevent the Estonians from moving the statue; instead, they launched an attack to show their displeasure and to highlight Estonia's vulnerability to their skills (McGuinness).

Beyond Law of War: Other Proposals for Cyberwar

The past few decades have produced important critical thought about potential ways to regulate and utilize cyberwar. Some of these proposals support JWT and IHL, but some directly violate both frameworks. For example, John Arquilla, an expert in cyberwarfare, recommends adopting a "declaratory doctrine of 'no first use' of information warfare against largely civilian targets" (Orent 18). This would mean that nations cannot target civilians in an act of cyberwar unless their own civilians were attacked first. This does address JWT and IHL's concern with distinction, but it violates the principle because this rule would allow for targeted attacks on civilians once one side has done so. This promotes unnecessary suffering among innocents and therefore should not be adopted.

A more acceptable solution was proposed by the RAND Corporation. It suggests creating an international organization for cyber attribution in order to get an unbiased determination of attribution of cyberwar. They also proposed an agreement between countries to include a digital signature on attacks to make identification much easier. This would follow the requirement that cyberwar be publicly declared without rendering all cyberwar useless by necessitating prior notice to an attack. Another potential solution developed by Brian Orent, an expert in JWT, suggests that cyberweapons themselves be limited and regulated through the creation of treaties and international agreements, similar to the regulations surrounding nuclear weapons (Orent 32).

There are a wide variety of proposed solutions existing in the world, some better than others. It is important that the ethical theories are featured prominently when solutions are discussed and decided upon in order to use the guidelines to minimize the suffering and damage that are inherent to war.

Future Recommendations

Existing JWT and IHL are sufficient to apply to and cover cyberwarfare. Therefore, a new treaty is not required. However, a formal and internationally recognized definition of cyberwarfare is necessary in order to facilitate ethical cyberwarfare. The main ethical issue currently surrounding cyberwarfare is not the lack of ethical standards and guidelines as these are provided by both JWT and IHL, but the lack of commitment to and following of these rules. The international community needs to create a standardized set of norms of what is acceptable and what is not acceptable in cyberwar, using JWT and IHL. These norms need to be followed by large, powerful nations that will lead by example. Once the big players are adhering to these rules, it will be much more difficult for other nations to reject them. These norms can also be

enforced through economic incentives and sanctions to promote further adherence to them. This could be done by utilizing existing international organizations such as NATO.

These norms would not be easy to establish, nor would they solve all the potential ethical issues in cyberwarfare, but they are a step in the right direction. They apply the ethical guidelines of JWT and IHL to the complicated nature of cyberwar in order to reduce suffering and increase the morality of cyberwar. However, the current international view of cyberwar is that there are no real norms or official regulations. This allows individual nations to create their own rules for behavior in cyberspace, resulting in wildly differing understanding of the ethics of cyberwar. This means that civilians will suffer the most disruption and destruction in their lives until the international community can jointly recognize that the already existing IHL does in fact apply to cyberwarfare. Once this recognition is made and behavior in cyberspace is regulated, nations can work on creating explicit laws and agreements that are specifically aimed at cyberwarfare.

Greater clarity and broader consensus would have many benefits. For example, one of the most important of these norms that come from JWT is the active avoidance of targeting civilians. Instead of the shift towards cyberwar that targets critical infrastructure, cyberwar should move towards more precise targeting of only legitimate government and military targets. This will greatly lower the unnecessary suffering that would be subjected on uninvolved citizens. Stuxnet is a prime example of this; it focused exclusively on specific, legitimate government targets.

Another norm that should be implemented is that all cyberwar should be designed to be reversible whenever possible. This will allow the attack to be effective in the moment because the damage has been done, but after the attack has terminated and tensions have decreased, the damage can be undone. For example, the attack could “be in the form of an encryption of critical data or programs using a secret key known only to the attacker, so performing a decryption could

repair the damage" (Schmitt). This would make cyberwar more ethical than traditional attacks because the damage is not permanent.

One final possible norm could focus on regulating the post attack process. Once the attack has concluded, the attacker should be required to aid the targeted nation. This could be reversing the attack's damage or paying fines to resolve any negative financial impact. The attacking nation should also publicly acknowledge their role in the development of the attack, something that almost never occurs currently. This would give the targeted nation the autonomy to make a clear, informed decision on the next steps, whether that is negotiations or a return attack. If an attack is severe and malicious enough, the regulations could even go so far as to limit the attacking nation's access to and development of cyberweapons for a set period of time, similar to denuclearization. This would ensure that the aggressive nation wouldn't be able to attack again and would punish them for violating IHL. However, doing so would be difficult and would require oversight from an international organization such as the United Nations.

Conclusion

The adoption of such norms would necessitate the international community agreeing on a comprehensive definition of what cyberwarfare is and how an act of cyberwar is distinct from a simple cyberattack. In order to do this, the traditional idea of aggression needs to be translated to cyberspace. This requires the recognition that cyberwar can generate at least as much damage as traditional war. Therefore, cyberattacks that deal such a level of destruction have to be considered an act of war.

Once a consensus has been reached, the international community as a whole needs to acknowledge that all cyberwar must adhere to JWT and IHL in order to be ethical and to avoid unnecessary suffering. This should, in theory, be straightforward because the vast majority of

nations already agree to IHL, which can be directly applied to cyberwar. The next step is to enforce these frameworks and hold noncompliant nations accountable through organizations such as the United Nations as well as other means such as economic sanctions. With the implementation of guidelines that follow JWT and IHL, cyberwarfare will be ethically enforced. Both JWT and IHL limit cyberwarfare and force those waging cyberwars to consider their actions and their consequences. JWT and IHL allow other nations to hold each other accountable and to utilize just cyberwar. The procedures surrounding cyberwarfare need to adhere to the JWT and IHL; only then will the harmful effects of cyberwar be minimized.

Works Cited

Aftergood, Steven. "What Is an Act of War in Cyberspace?" *Federation Of American Scientists*, 16 Oct. 2017, https://fas.org/blogs/secrecy/2017/10/war-cyberspace/.

"Aid Operations under Increasing Threat as State, Non-State Combatants Ignore International Law, Humanitarian Affairs Chief Warns Security Council | Meetings Coverage and Press Releases." *United Nations*, United Nations, https://www.un.org/press/en/2019/sc13760.doc.htm.

Allhoff, Fritz, and Ryan Jenkins. "When Is a Real-World Response to a Cyberattack Justifiable?" *Slate Magazine*, Slate, 11 June 2014, https://slate.com/technology/2014/06/cyberwar-ethics-when-is-a-real-world-response-to-a-cyberattack-justifiable.html.

Cheravitch, Joe. "Cyber Threats from the U.S. and Russia Are Now Focusing on Civilian Infrastructure." *RAND Corporation*, 23 July 2019, https://www.rand.org/blog/2019/07/cyber-threats-from-the-us-and-russia-are-now-focusing.html.

Davis, Joshua. "Hackers Take Down the Most Wired Country in Europe." *Wired*, Conde Nast, 5 June 2017, https://www.wired.com/2007/08/ff-estonia/.

Emba, Christine. "Just War Theory: A Primer." *The Washington Post*, WP Company, 31 Mar. 2019, https://www.washingtonpost.com/news/in-theory/wp/2015/11/30/just-war-theory-a-primer/.

Finkelstein, Claire Oakes, et al. *Cyberwar: Law and Ethics for Virtual Conflicts*. Oxford University Press, 2015.

Gorman, Siobhan, and Julian E. Barnes. "Cyber Combat: Act of War." *The Wall Street Journal*, Dow Jones & Company, 31 May 2011, https://www.wsj.com/articles/SB10001424052702304563104576355623135782718?ns=prod/accounts-wsj.

"International Humanitarian Law." *International Justice Resource Center*, https://ijrcenter.org/international-humanitarian-law/.

Lucas, George R. "Permissible Preventive Cyberwar: Restricting Cyber Conflict to Justified Military Targets." *The Ethics of Information Warfare*, Springer, 2016, pp. 73–84.

McGuinness, Damien. "How a Cyber Attack Transformed Estonia." *BBC News*, BBC, 27 Apr. 2017, https://www.bbc.com/news/39655415.

Orend, Brian. "Fog in the Fifth Dimension: The Ethics of Cyberwar." *The Ethics of Information Warfare*, Springer, 2016, pp. 3–24.

Perlroth, Nicole. "In Cyberattack on Saudi Firm, U.S. Sees Iran Firing Back." *The New York Times*, The New York Times, 24 Oct. 2012, https://www.nytimes.com/2012/10/24/business/global/cyberattack-on-saudi-oil-firm-disquiets-us.html.

Schmitt, Michael N. *Tallinn Manual 2.0: On the International Law Applicable to Cyber Operations*. Cambridge University Press, 2017.

"The Real Story of Stuxnet." *IEEE Spectrum: Technology, Engineering, and Science News*, https://spectrum.ieee.org/telecom/security/the-real-story-of-stuxnet.

"UN Charter." *United Nations*, United Nations, https://www.un.org/en/sections/un-charter/un-charter-full-text/.

"The United Nations and International Law." *The United Nations and International Law*, UN, Department of Public Information, 1991.

What Is International Humanitarian Law?, International Committee of the Red Cross, 2004.

"What Is Stuxnet?" *McAfee*, https://www.mcafee.com/enterprise/en-us/security-awareness/ransomware/what-is-stuxnet.html.