Emerging technology continues to revolutionize all traditional forums for human activity. Known as the Fourth Industrial Revolution,[1] the innovation of this century has been sparked by the ongoing development of artificial intelligence (AI), machine learning, and expansion of cyberspace. Within cyberspace lies unlimited potential to benefit humanity, but falls victim to a security dilemma in which continued economic and military competition is lead by technologically developed stakeholders (Buchanan 2019: 3). Cyberspace now embodies a critical area of debate centered primarily on the modern conflict landscape (Schmitt 2017: 4). On 24 June 2013, the United Nations’ (UN) Group of Governmental Experts (GGE) on Developments in the Field of Information and Telecommunications (ICT) in the Context of International Security stated that the UN Charter is applicable within cyberspace (UNGA 2013: 2). This proclamation solidified the notion that cyberspace is the fifth domain for warfare (The Economist Briefing 2010). The Tallinn Manual 2.0 on the International Law Applicable to Cyber Warfare, finalized in 2017, became the primary source on how sovereignty, state responsibility, human rights, and the law of air, space, and the sea apply to cyberspace (Schmitt 2017: 10). This piece thoroughly documents the relevance of jus ad bellum[2] to cyber conflict, but the applicability of jus in bello[3] to cyberattacks yields scant scholarship due to its lack of precedent and few available frameworks of direct relation. This paper will, therefore, utilize the publications of the International Committee of the Red Cross (ICRC) in addition to the Tallinn Manual 2.0 to establish the extent to which existing norms and principles of international humanitarian law (IHL) apply to cyberattacks in an active conflict setting.

The Cyber Threat Landscape

In order to demonstrate the human cost and consequent applicability of IHL to cyberattacks, a technical summary must precede a discussion around a weaponized cyberspace. Under a general scope, the European Union Agency for Network and Information Security (ENISA) defines the ‘Internet of Things’ (IoT) as, “a cyber-physical ecosystem of interconnected systems and actuators, which enable intelligent decision making” (ENISA 2017: 1). Cloud computing serves as the backbone of this system, as cloud computing enables citizens and states alike to store and analyze vast quantities of data in real-time while undersea fiber optic cables serve as the high speed data transmission medium (Routely 2017). Given the enormity of modern data use, companies are exploring the possibility of manufacturing more fiber optic cables in the International Space Station (The Economist Briefing 2018). This symbiotic relationship has transformed civil society, but evokes grave concerns in return. The relationship depends on the preservation of ‘grids,’ or the generation, transmission, distribution and consequent end use of electricity (US Department of Energy 2014). As personal data, energy utilities, infrastructure, healthcare facilities, transportation, and peoples’ homes are interconnected and remotely addressable (World Economic Forum 2019), the security implications include the ability to compromise all of the data underpinning a civil society’s day-to-day operations (Deloitte 2019). Therefore, distinguishing the various branches of threats within the cyber threat landscape proves useful to building upon a baseline technical understanding.

Given the ever-evolving nature of cyberspace, new threats materialize almost daily (Knake & Clarke 2012: 5). The most prevalent threats to the modern cyber ecosystem include cyber espionage, cybercrime, cyber-enabled disinformation campaigns and cyberattacks. Within the context of economic and political relations between nation-states, these practices hold a unique sphere within the emergence of modern conflict. Each topic varies in definition...
Mitigating the Human Cost of Modern Conflict: Jus in Bello and Cyberattacks

Written by Tory Igoe

depending on the consulted legal code (Rubenstein 2014), so a generalized definition will be proposed for the purpose of differentiating between branches. Cyber espionage, in short, involves the cyber-enabled theft of government or corporate intellectual property (IP) (Kessler 2017). APT40, a Chinese cyber espionage group, is a primary example, as they target strategically placed states relative to the ‘Belt and Road Initiative’ (FireEye 2019). Cybercrime is far broader, as the Budapest Convention put forth by the Council of Europe (COE) defines cybercrime as, ‘criminal acts that are committed [through] online electronic communication networks and information systems’ (Council of Europe 2001: 7). The Silk Road, a dark web marketplace dedicated to trafficking illicit goods (Nimfuehr 2018), serves as the most famous example of an act of cybercrime. Distinctive from the preceding branches, cyber-enabled disinformation campaigns are the modern incarnation of an ancient war tactic. Within the sphere of information geopolitics and propaganda, a cyber-enabled disinformation campaign involves the conscious spreading of false information to influence a specific outcome (Rosenbach & Mansted 2019). The influence campaign spearheaded by the Russian Federation against the United States’ 2016 election epitomizes the harm cyber-enabled ‘fake news’ may yield (Nye 2019). These practices embody some of the greatest hazards to cyber stability, but do not possess the equivocal physical impact day-to-day operations may feel following a cyberattack.

Issued by Rule 30 of the Tallinn Manual 2.0, ‘a cyberattack is a cyber operation, whether offensive or defensive, that is reasonably expected to cause injury or death to persons or damage or destruction to objects’ (Schmitt 2017: 94). This definition fails to highlight the variances in cyberattacks, but emphasizes the crippling outcome if done effectively. Stemming from its umbrella definition, the most common forms of cyberattacks include denial-of-service (DDoS), man-in-the-middle (MitM), phishing, drive-by attack, password cracking, SQL injection, cross-site scripting (XSS), eavesdropping attack, birthday attack, and malware attacks. This strata of methodologies demonstrate how cyberattacks vary in approach and severity, but may damage civil society across a range of operations.

Independent from severity, all of the previously mentioned attacks can be checked at any point of the ‘cyber kill chain.’ Published by Lockheed Martin Corporation in 2011 as an intelligence driven defense model, the seven steps of the cyber kill chain – reconnaissance, weaponization, delivery, exploitation, installation, command & control (c2), and actions on objectives – describe the path a cyberattack must take to achieve its ultimate objective (Lockheed Martin Corporation 2019: 3). Governments and private sector entities now understand that stopping adversaries at any point of the seven stages ends the attack in its entirety (Panda Security 2019: 5). Unfortunately, copious forms of cyberattacks – both well-known and those of enhanced complexity – reach their objective successfully. Highlighted by Microsoft Corporation’s recent announcement that state actors targeted 10,000 customer accounts in the past year (Burt 2019), the mechanisms in place to protect civilians, states and critical infrastructure fall short. Thus, further analysis into weaponizing cyberspace within a pre-existing conflict sparks alarm.

To preface the discussion of cyberattacks capable of astounding harm, six cyber nation-state powers exist within the modern cyber threat landscape. These states include the United States (US), the United Kingdom (UK), Russia, China, North Korea, and Israel (Vavra 2017). Despite varying cyber capabilities, each embrace cyberattacks as a method of statecraft due to the dependence of civil society on an electrical grid, the symbiotic relationship between IoT and cloud computing, in addition to the existence of billions of mobile devices. Thus, cyber offensive capabilities make a lucrative military strategy. An example of a targeted cyberattack on critical infrastructure undertaken by a state against another state occurred on 23 December 2015. Presumed Russian hackers employed the use of BlackEnergy 3 malware in a phishing attack on three major energy providers within the Ivano-Frankivsk region of Ukraine. When employees downloaded the malicious emails, hackers were able to plant KillDisk malware into the electrical grid causing a six-hour blackout for 225,000 people (SANS 2016: 5-8). ‘CRASH OVERRIDE,’ an identical cyberattack focused on Kiev, occurred one year later (Greenberg 2017). These incidents marked more than the first use of a cyberattack against a power grid (International Risk Management 2016), as they highlight profound concerns for weaponizing cyberspace.

These events demonstrate the human cost of cyberattacks. Within the Ivano-Frankivsk region and Kiev exist numerous hospitals, power plants, water treatment centers, financial institutions, and government agencies all interconnected and dependent on the same grid. Albeit incurring no deaths, an attack of this nature holds the potential to unravel the fundamental day-to-day operations of a society. Without a functioning hospital, people die.
Mitigating the Human Cost of Modern Conflict: Jus in Bello and Cyberattacks
Written by Tory Igoe

Without water treatment plants, disease spreads. Financial institutions and industries unable to operate for an extended period of time paralyzes economies. As the consequences mount, a state comes apart at the seams. In addition to these dire straits, the attribution dilemma looms as an ancillary concern. When states cannot perfectly attribute attacks to other states in a pre-existing conflict setting, the creation of a ‘grey zone’ undermines domestic and international legal systems (Wolitzky, de Mesquita, & Baliga 2019: 12). Thus, a world of growing interconnectivity and rapid technological development without reasonable clarity on global governance mechanisms yields a bleak prediction for the development of modern conflict.

International Humanitarian Law (IHL) and Cyberattacks

In accordance with the mission of the ICRC, IHL serves as a set of rules to limit the effects of armed conflict for humanitarian purposes (ICRC 2004: 2). The principles of IHL are the Martens Clause, the principle of distinction, the principle of proportionality, and the principle of military necessity (ICRC 2019). Regarding the codification of these principles, IHL derives its powers from the four Geneva Conventions of 1949, the Additional Protocols of the Geneva Conventions adopted in 1977 (ICRC 2010), and various treaties banning the development and usage of certain weaponry (Melzer 2011: 5). Regarding cyberattacks, the Fourth Geneva Convention and subsequent Additional Protocols prove most applicable within this domain. Stated in Article 22 of The Hague Convention, ‘the right of belligerents to adopt means of injuring the enemy is not unlimited’ (ICRC 1907) offers the fundamental idea behind combatting an ever-evolving threat landscape. Albeit central to the conceptualization of jus in bello, a clear hindrance lies in Article 36 of the 1977 Additional Protocol stating that, ‘in the study, development, acquisition or adoption of a new weapon, means or method of warfare, a High Contracting Party is under an obligation to determine whether its employment would, in some or all circumstances, be prohibited by international law’ (UNGA 1977). An assertion meant to address this ever-evolving threat landscape, a lack of widespread understanding around cyberspace undermines the notion from the start. States have entered a digital arena with unclear adversaries, a lack of universal definitions, and a cleavage between engineers and policy makers. The current open-ended working group (OEWG) of the General Assembly sponsored by Russia and Group of Governmental Experts (GGE) sponsored by the United States prove two of these points: the lack of a multistakeholder approach in addressing these gaps in global governance and placing state competition before widespread stability (Grigsby 2018). Given the state of developing cyber norms, present frameworks remain the best springboard for discussing applicability.

In continued attempts to transpose a body of law designed for traditional modes of conflict to a virtual domain, the Tallinn Manual 2.0 presents the sole piece of widely accepted literature addressing the topic. Within Part IV of the Tallinn Manual 2.0, it is established that IHL applies when undertaking a ‘cyber operation’ within a pre-existing conflict (Schmitt 2017: 68). In the words of the International Court Tribunal for the Former Yugoslavia (ICTY), this is true when the ‘operation’ is undertaken with a clear ‘nexus’ between the armed conflict at hand and the operation in question (ICRC 2000). Noting the role of cyberattacks in an international armed conflict and non-international armed conflict, the Tallinn Manual 2.0 affirms the perspective of the ICRC in stating that the legal attributability of a cyberattack is guided by the general international law of state responsibility (Schmitt 2017: 35). The official codification in Rule 14 of Tallinn 2.0 states that, ‘[A] State bears international responsibility for a cyber-related act that is attributable to the State and that constitutes a breach of an international legal obligation,’ (Schmitt 2017: 59). This view includes a person acting on behalf of a state in a de jure or de facto capacity (Melzer 2011: 24). Attribution is fundamental to legal responsibility, but attributing a cyberattack proves difficult for the best cyber security experts (SearchSecurity 2019). This fact undermines legal responsibility in its entirety (Rid & Buchanan 2015: 32). When attributing an attack to a state party to a conflict is not possible, IHL bears no explanation on how to proceed; consequently, the lack of clarity around state responsibility destabilizes all underlying principles of IHL in the process.

The law of state necessity, derived from the law of state responsibility, invokes the principle of military necessity. This principle ‘permits’ the necessary measures needed to achieve a military objective without infringing upon IHL (ICRC 2001). Given the ‘downright dangerous’ lack of clarity around cyber offensive capabilities (Serbu 2018) and added lack of precedent, applying the concept of ‘military necessity’ to cyberspace remains in its infancy. Referring to the discussion of attributability, military necessity requires a level of predetermined threat. If a party to a conflict finds themselves at the receiving end of a cyberattack, an inability to ‘distinguish’ the adversary may lead to an
Mitigating the Human Cost of Modern Conflict: Jus in Bello and Cyberattacks
Written by Tory Igoe

unpermitted escalation in conflict under IHL, or an act transgressing the principle of proportionality. Thus, the interwoven nature of the principles of IHL allow for an ineffective framework. This dilemma segues into deciding whether or not a cyberattack constitutes an act of ‘armed force.’ The question itself is twofold, as it raises whether or not a cyberattack can be justified or exacerbate pre-existing conflicts. According to Article 51 of the UN Charter, states hold the right of self-defense in the face of an armed attack (UNGA 1945). Conversely, a cyberattack crosses the threshold of an act of ‘armed force’ only when the effects are deemed equivocal to an act of kinetic armed force (Serralvo & Dormann 2014: 712). As attribution and lack of precedent remains a pitfall of this argument, interpreting when it is appropriate for a state to repudiate a cyberattack is left to the state’s discretion under present norms of conflict. In this case, cyberattacks hold the potential to escalate towards or engage in more severe forms of kinetic conflict evoking grim prospects for technologically advanced conflict (Lin 2012: 51).

The purpose of IHL is to safeguard civilians from the barbarity of war. Under the principle of distinction, parties are required to distinguish between combatants and non-combatants, military objects and civilian objects (ICRC 1977). The concept of combatancy derives from the The Hague Regulations and their consequent adoption of Article 4A within Geneva Convention III (Schmitt 2017: 84). According to Rule 26 of the Tallinn Manual, ‘members of the armed forces of a Party to the conflict who, in the course of cyber operations, fail to comply with the requirements of combatant status lose their entitlement to combatant immunity’ (Schmitt 2017: 84). As the International Group of Experts for the Tallinn Manual concluded that individual members adopting cyberattacks on civilian infrastructure would fall under the category of unprivileged belligerents, said individuals would be prosecuted under domestic statute (Schmitt 2017: 87). Distinction in cyberspace, as touched on in 2014 by Microsoft’s proposed international cybersecurity norms, requires a new range of categories not yet covered by IHL (Microsoft 2014: 4). Aside from understanding one’s adversary, undertaking a cyberattack on networks and infrastructure implicates the surrounding populace (Tsagourias & Buchan 2014: 357), as the Internet itself is 90% civilian infrastructure (ICRC 2019: 29). This notion does not imply that all cyberattacks are indiscriminate (ICRC 2019: 36), but the interconnected character of cyberspace challenges the principle of distinction at its core.

The heart of IHL itself exists in the Martens Clause, or the principle of humanity. The Martens Clause states that as treaties and norms develop, citizens and combatants remain under the protection of international law as international law itself is an extension of the ‘public conscience’ of ‘civilized nations’ (Peace Palace Library 2019). Given the ongoing developments in analyzing how IHL applies to cyberattacks, scholars and policymakers must remember the human element behind any technology. The ICRC serves as the global vanguard for this principle, and spearheaded a clear analysis centered on the industries most vulnerable to cyberattacks in conflict. The ‘human cost’ of cyberattacks and key pieces of infrastructure threatened by cyberoffensive measures are systems impacting the delivery of essential healthcare, supervisory control and data acquisition (SCADA) industrial control systems, internet services and cloud service providers (ICRC 2019: 10). Regarding medical care, a hospital at full-operability possesses two network systems – those dedicated to administrative data and those embedded in medical devices (ICRC 2019: 18-19). Within a conflict setting, patient information may be exploited to find a specific adversary while an attack centered on medical devices hold the potential to sabotage a needed procedure (ICRC 2019: 19). Noting the threat of cyberattacks against SCADA industrial control systems, this technology regulates valves, motors, and various industrial processes (Goodwin 2018). If targeted within a conflict setting, the system collapse of energy and heavy industry complexes can wreak havoc on the surrounding area (ICRC 2019: 23-24). Moreover, an attack of this nature yields potential for impacting the grid of a city, spurring widespread blackouts mimicking the case in Ukraine. The consequent events drive up the human cost of cyberattacks. Lastly, cloud infrastructure and internet services offers those party to a conflict a high reward target based on interconnectivity (ICRC 2019: 29-30). As previously noted, the internet itself is 90% civilian infrastructure (ICRC 2019: 29), and the legal ramifications prove enormous while invoking the principle of distinction and humanity. Moreover, the symbiotic relationship between IoT and cloud computing allow for real-time analysis and vast data storage for billions of devices. The consequences, on a large scale, cut communication and wireless protocols civilians and military personnel alike depend on (Shea 2019). From cellphones to industry, an attack of this caliber holds the potential to disrupt and dismantle pieces of infrastructure needed to safeguard civilians within an active conflict setting (ICRC 2019: 31). Therefore, in a digital domain wrought with vulnerabilities and mechanisms for attack, assessing the framework of IHL against the use of cyberattacks begs more questions than answers.
Mitigating the Human Cost of Modern Conflict: Jus in Bello and Cyberattacks

Written by Tory Igoe

Conclusion

The principles of IHL offer the strongest mechanism for protecting civilians amidst kinetic conflict (ICRC 207: 35). Its powers under the Geneva Conventions and Additional Protocols protect all parties to conflict, as conflict itself continues to evolve into a multi-domain endeavor. However, the countless ambiguities in cyberspace, specifically cyberattacks, grant unique challenges to the applicability of IHL (ICRC 2019: 38). It can be deduced that the principles of military necessity, distinction, proportionality, and humanity fall short when addressing the difficulties in attack attribution, lack of precedent and the asymmetric nature of ‘cyber conflict’ en masse. Given that the majority of cyber operations land below the threshold of kinetic aggression, continued discussion exists as to how international institutions will mitigate combat these ‘known threats’ (ICRC 2019: 39). Given that these ‘known threats’ exist in the ‘grey zone’ between peacetime and conflict, global governance mechanisms in place prove inadequate when addressing the varying scale and present uses of cyberattacks. In spite of this fact, IHL applies to cyberattacks within the context of armed conflict, but to a nominal extent.

The prospect of a conflict occurring solely within cyberspace is far from conceivable at this point in human history (Rees 2018: 85), so international institutions must continue to focus on what is a known security dilemma. Present shortcomings in mitigating the use of cyberattacks demonstrate how technologically advanced states exploit the novelty and lack of regulation around cyber statecraft for political and economic gain. Further, the lack of cooperation between government entities, international institutions and private sector innovation demonstrates a clear vulnerability in developing new norms of state behavior in cyberspace. The imperfections of technology and legal frameworks are due their being an extension of human capability. Humanity can adapt to technological advancement, so global governance frameworks must do the same. If international institutions fail to forge a new digital convention applicable to all contexts and capacities in which cyberattacks occur under the auspices of policy makers and engineers alike, the fabric of humanity’s new, interconnected sphere will come apart at the seams.

Bibliography


Mitigating the Human Cost of Modern Conflict: Jus in Bello and Cyberattacks
Written by Tory Igoe


Mitigating the Human Cost of Modern Conflict: Jus in Bello and Cyberattacks
Written by Tory Igoe


Mitigating the Human Cost of Modern Conflict: Jus in Bello and Cyberattacks

Written by Tory Igoe


Notes

[1] The fundamental change in the way humanity lives, works and relates following extraordinary technological
development (World Economic Forum 2019).

[2] The conditions under which states may resort to war or the use of armed force (ICRC 2019).


[4] China’s proposed infrastructure development and investment initiative that seeks to connect East Asia to Europe vis-à-vis the ancient Silk Road (McBride & Chatzky 2019).

[5] An attack meant to shut down a machine or network, making it inaccessible to its intended users by flooding it with traffic (FireEye 2019).

[6] An attack that relays and possibly alters the communications between two parties who believe they are directly communicating with each other (FireEye 2019).

[7] A social engineering attack used to steal user data (FireEye 2019).

[8] Common method used to spread malware by means of ‘drive-by’ downloads on insecure websites (FireEye 2019).

[9] Recovering passwords from data stored or transmitted by a computer system (FireEye 2019).

[10] Injecting code that may destroy your database (FireEye 2019).


[12] An attack that takes advantage of unsecured network communications to steal data from computers, smartphones and other mobile devices (FireEye 2019).

[13] Cryptographic attack that exploits the math behind the birthday problem in probability theory (FireEye 2019).

[14] Malicious software that takes over a person’s computer, make it a botnet, and send malware to other computers (FireEye 2019).


[16] Provisional recognition given by a government if seen to fulfill the requirements of permanence, popular support, and prior international obligations (Qadir 2012).

Written by: Tory Igoe
Written at: Northeastern University
Written for: Professor Denise Garcia
Date written: July 2019