ICRC, NATO AND THE U.S. – DIRECT PARTICIPATION IN *HACKTIVITIES* – TARGETING PRIVATE CONTRACTORS AND CIVILIANS IN CYBERSPACE UNDER INTERNATIONAL HUMANITARIAN LAW

IDO KILOVATY[†]

ABSTRACT

Cyber-attacks have become increasingly common and are an integral part of contemporary armed conflicts. With that premise in mind, the question arises of whether or not a civilian carrying out cyber-attacks during an armed conflict becomes a legitimate target under international humanitarian law. This paper aims to explore this question using three different analytical and conceptual frameworks while looking at a variety of cyber-attacks along with their subsequent effects. One of the core principles of the law of armed conflict is distinction, which states that civilians in an armed conflict are granted a set of protections, mainly the protection from direct attacks by the adversary, whereas combatants (or members of armed groups) and military objectives may become legitimate targets of direct attacks. Although civilians are generally protected from direct attacks, they can still become victims of an attack because they lose this protection "for such time as they take direct part in hostilities."[1] *In other words,*

[†] Cyber Fellow at the Center for Global Legal Challenges, Yale Law School; Resident Fellow Information Society Project, Yale Law School; S.J.D. Candidate, Georgetown University Law Center. I would like to gratefully acknowledge the generous support of the Minerva Center for the Rule of Law under Extreme Conditions at the Faculty of Law and Department of Geography and Environmental Studies, University of Haifa, Israel and of the Israeli Ministry of Science, Technology and Space, who made this project possible. I am thankful for the comments and support of Prof. Amnon Reichman, Prof. Rosa Brooks, Prof. Alexa Freeman, Prof. Mary DeRosa, Adv. Ido Rosenzweig, and all the attendees of the cyber sessions at the University of Haifa, who were incredibly helpful in the process of writing this paper.

[1] Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I) art. 51(3), June 8, 1977, 1125 U.N.T.S. 3 [hereinafter AP I]. In the non-international armed conflict context, see Article 13(3) of the Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of Non-

under certain circumstances, if a civilian decides to engage in hostile cyber activities (or "hacktivities"), they may well become a target of a direct lethal attack. I will argue that although the answer is highly nuanced and context dependent, the most salutary doctrinal revision that can be made in this area is that the threshold of harm must adapt to the particular intricacies of cyberspace.

INTRODUCTION TO TARGETING IN MODERN CONFLICT

In August 2015, a U.S. drone strike in Syria killed Junaid Hussain, a British hacker who was carrying out hostile cyber activities on behalf of ISIS.[2] Hussain was believed to have a leading position in ISIS' Cyber Caliphate, a hacking group that allegedly took control of US Central Command's social media accounts. This group published U.S. soldiers' and officers' identifying information, such as their full names, addresses, and photos.[3] Unlike other members of ISIS, Hussain only engaged in cyber activities which, although they were hostile, did not pose the same threat of imminent danger as the actual use of conventional military force. Hussain's death represents the first time a hacker was lethally targeted. This article predicts that the practice of targeting hackers in the context of an armed conflict will only increase because cyber-attacks will become a more frequent and substantial phenomenon in armed conflicts. The crucial question in this regard is whether individuals like Hussain are legitimate targets under international humanitarian law (IHL) in armed conflict.

The emergence of cyberspace as an instrument of warfare exacerbates the gaps and ambiguities that already exist within IHL. A particularly aggravating factor is the increased involvement of civilians in cyber-attacks. Civilians in modern armed conflicts can play an active role when it comes to cyber-attacks, as they serve as private contractors carrying out both offensive and defensive cyber operations.[4] In addition to

---

International Armed Conflicts (Protocol II), June 8 1977, 1125 U.N.T.S. 609 [hereinafter AP II].

[2] Spencer Ackerman, Ewan MacAskill & Alice Ross, *Junaid Hussain: British Hacker for ISIS Believed Killed in US Air Strike*, THE GUARDIAN (Aug. 27, 2015, 12:28 EDT), http://www.theguardian.com/world/2015/aug/27/junaid-hussain-british-hacker-for-isis-believed-killed-in-us-airstrike.

[3] Julia Glum, *ISIS Hacker Junaid Hussain Confirmed Dead After US Airstrike on Islamic State in Syria: Pentagon*, INTERNATIONAL BUSINESS TIMES (Aug. 28, 2015, 2:52 PM), http://www.ibtimes.com/isis-hacker-junaid-hussain-confirmed-dead-after-us-airstrike-islamic-state-syria-2073451.

[4] *See* Heather Harrison Dinniss, *Cyber Warriors, Patriotic Hackers and the Laws of War*, *in* INTERNATIONAL HUMANITARIAN LAW & THE CHANGING TECHNOLOGY

these private contractors, there are patriotic hackers who make similar contributions.[5] While the traditional notion of an armed conflict includes members of regular armed forces ("combatants"), this paradigm is no longer the case on the modern battlefield due to the growing involvement of civilians in cyber hostilities. As Susan Brenner, a leading expert on cyber conflict, stated, the "integration of civilians into military efforts can create uncertainty as to whether someone is acting as a 'civilian' (noncombatant) or as a military actor (combatant)."[6] After all, it is far easier to possess a computer with an active internet connection than a conventional weapon.[7] This is because computer technology is readily accessible and carries a low cost of entry and use.[8]

In fact, in the aftermath of the cyber-attacks on Estonia in 2007, the Estonian government decided to recruit civilian volunteers to serve as "cyber warriors" should another major cyber-attack on Estonia occur.[9] Sean Watts, Professor of Law at Creighton University, explained this phenomenon of the privatization of cyber operations by describing how today, many companies provide the expertise required for the development and employment of computer network attacks.[10]

The versatility of computers and cyberspace has caused difficulties for countries that wish to utilize computers militarily, while retaining their civilian properties. This is sometimes referred to as the

---

OF WAR 269 (Dan Saxon ed., 2013) (noting that civilian contractors are increasingly involved in cyber hostilities).

[5] *Id*.

[6] SUSAN BRENNER, CYBERTHREATS, THE EMERGING FAULT LINES OF THE NATION STATE 197 (2009).

[7] *See* Roger Barnett, *A Different Kettle of Fish: Computer Network Attack*, 76 J. INT'L. L. STUD. 1, 22 (2002) (stating that the "entry costs to conduct a strategic information attack are insignificant" and requite only "an inexpensive computer, some easily obtainable software, and a simple connection to the Internet.").

[8] *See* MARCO ROSCINI, CYBER OPERATIONS AND THE USE OF FORCE 202 (2014) (noting that the "low cost and ease of access to technology" allows civilians to easily conduct cyber operations).

[9] David Blair, *Estonia Recruits Volunteer Army of 'Cyber Warriors'*, THE TELEGRAPH (Apr. 26, 2015, 6:58 PM BST), http://www.telegraph.co.uk/news/worldnews/europe/estonia/11564163/Estonia-recruits-volunteer-army-of-cyber-warriors.html.

[10] *See* Sean Watts, *Combatant Status and Computer Network Attack*, 50 VA. J. INT'L L. 391, 402–03 (2010) ("Reports indicate that few information operations experts currently serve as active duty soldiers. Many private companies have employed the skills of those with expertise in the various weapons commonly used in [Computer Network Attacks].").

"dual-use" characteristic of cyberspace.[11] Dual-use means that computer systems and networks may be used for legitimate, benign, civilian purposes, but at the same time, these systems and networks may be used by the military, or by civilians, for hostile purposes.[12] While this characteristic is clear to countries, their civilians are often unaware of the potential consequences of their involvement in cyber operations in armed conflict, the importance of which may be life or death. Such was the case with Evgeny Morozov, who participated in the cyber operations against Georgia in 2008 –

> "Not knowing exactly how to sign up for a cyberwar, I started with an extensive survey of the Russian blogosphere… As I learned from this blog post, all I needed to do was to save a copy of a certain Web page to my hard drive and then open it in my browser. In less than an hour, I had become an Internet soldier. I didn't receive any calls from Kremlin operatives; nor did I have to buy a Web server or modify my computer in any significant way."[13]

The problem of civilians who directly participate in hostilities (DPH) is not particularly new. Yet even today, many find themselves unable to define the precise boundaries of the notion of DPH. Only in 2009 did the International Committee of the Red Cross (ICRC) publish the Interpretive Guidance on the Notion of Direct Participation of Hostilities ("Interpretive Guidance"), which attempted to deal with the phenomenon of civilians who decide to directly participate in hostilities and the legal

---

[11] *See* Mary O'Connell, *Cyber Security Without Cyber War*, 17 J. CONFLICT & SECURITY L. 187, 205 (2012) (comparing cyberspace to other existing treaty regimes of arms control, noting that "[t]he international community has adopted treaties in other 'dual-use' areas that are analogous to cyber space, such as the Chemical Weapons Convention and the Nuclear Non-Proliferation Treaty. Both of these treaties seek to end any use or even possession of chemical or nuclear weapons while at the same time promoting legitimate non-military uses of chemicals and nuclear power").

[12] *See* Kai Ambos, *International Criminal Responsibility in Cyberspace*, *in* RESEARCH HANDBOOK ON INTERNATIONAL LAW AND CYBERSPACE 118, 131 (Tsagourias & Buchan eds., 2015) (explaining that the "interconnectivity between civilian and military purposes" of cyberspace complicates the principle of distinction between civilians and combatants).

[13] Evgeny Morozov, *An Army of Ones and Zeroes: How I Became a Soldier in the Georgia-Russia Cyberwar*, SLATE (Aug. 14, 2008), http://www.slate.com/articles/technology/technology/2008/08/an_army_of_ones_and_zeroes.html.

and practical ramifications of such a decision.[14] The main consequences of such participation are the forfeiture of civilian protections and the legitimization of lethal targeting against civilians who DPH. The Interpretive Guidance does not directly deal with the issues raised by cyber warfare, but it does provide the general paradigm for civilians who DPH.

Additional attempts to address cyberspace and armed conflict are reflected in the 2013 Tallinn Manual on the International Law Applicable to Cyber Warfare ("Tallinn Manual"), a NATO sponsored non-binding codification of the international legal norms applicable to the wartime use of cyber-attacks.[15] This document specifically deals with cyber warfare (as opposed to the more general Interpretive Guidance by the ICRC), but it is not legally binding.[16] The recent Law of War Manual by the U.S. Department of Defense addresses the notion of DPH, but does not sufficiently elaborate on its connection to cyberspace.[17]

This paper argues that the boundary between legitimate activity in cyberspace during armed conflict (or, participation in war efforts) and hostile activities, which constitute DPH, is unclear. This resulting ambiguity has operational consequences, such as the inability to determine whether a specific individual is targetable in as a result of his or her actions in cyberspace.[18] Furthermore, this paper argues that cyber activities by

---

[14] NILS MELZER, INT'L COMM. OF THE RED CROSS, INTERPRETIVE GUIDANCE ON THE NOTION OF DIRECT PARTICIPATION IN HOSTILITIES UNDER INTERNATIONAL HUMANITARIAN LAW (2009) [hereinafter INTERPRETIVE GUIDANCE].

[15] *See generally* TALLINN MANUAL ON THE INTERNATIONAL LAW APPLICABLE TO CYBER WARFARE (Michael N. Schmitt ed., 2013) [hereinafter TALLINN MANUAL], https://ccdcoe.org/tallinn-manual.html (discussing the legal issues surrounding cyberwarfare).

[16] *Id.* at 5.

[17] *See, e.g.*, U.S. DEP'T OF DEF., LAW OF WAR MANUAL 222–27 (2015) [hereinafter LAW OF WAR MANUAL], http://www.defense.gov/Portals/1/Documents/pubs/Law-of-War-Manual-June-2015.pdf (discussing generally the complexities of civilian actions in hostilities but missing an elaboration on civilian hacking activities).

[18] This unclear boundary between direct participation in hostilities and participation in war efforts was already controversial and highly debated at the conclusion of the 1977 Additional Protocol to the Geneva Conventions, as provided in the INT'L COMM. OF THE RED CROSS, COMMENTARY ON THE ADDITIONAL PROTOCOLS OF 8 JUNE 1977 TO THE GENEVA CONVENTIONS OF 12 AUGUST 1949 at 619 (Yves Sandos et al. eds., 1987) [hereinafter ICRC COMMENTARY ON APS], http://www.loc.gov/rr/frd/Military_Law/pdf/Commentary_GC_Protocols.pdf ("There should be a clear distinction between direct participation in hostilities and participation in the war effort. The latter is

civilians belonging to the opposing party of an armed conflict should meet a certain threshold in order to be considered DPH. Although the "threshold of harm" requirement is the first prerequisite under the DPH paradigm, it is unclear what activities in cyberspace reach this threshold, as cyberspace challenges the notions of physicality, political borders, and state monopoly on force. This paper argues that the DPH framework may have to be more responsive to the myriad harms that emanate from cyberspace activities.

I. CIVILIANS AND THE NOTION OF DIRECT PARTICIPATION IN HOSTILITIES

International humanitarian law is a specific branch of international law that regulates and limits conduct in times of armed conflict.[19] It is a set of principles and norms codified in treaties, such as the Geneva Conventions of 1949 and their Additional Protocols from 1977, and customary international law.[20] The key purpose of IHL is its attempt to minimize the adverse effects associated with war, particularly the effects on uninvolved individuals and collectives, namely civilians.[21]

Determining whether an individual is a civilian or not (based on his or her status, rather than acts) is a separate and complicated issue. Article 50 of Additional Protocol I defines a civilian as "any person who does not belong to one of the categories of the persons referred to in Article

---

often required from the population as a whole to various degrees. Without such a distinction the efforts made to reaffirm and develop international humanitarian law could become meaningless. In fact, in modern conflicts, many activities of the nation contribute to the conduct of hostilities, directly or indirectly; even the morale of the population plays a role in this context.").

[19] *See* Int'l Comm. of the Red Cross, *International Humanitarian Law: Answers to Your Questions* (2002), http://www.redcross.org/images/MEDIA_CustomProductCatalog/m22303661_IHL-FAQ.pdf ("The purpose of international humanitarian law is to limit the suffering caused by war by protecting and assisting its victims as far as possible.").

[20] *See* Statute of the International Court of Justice art. 38(1)(b), June 26, 1945, 59 Stat. 1031, 33 U.N.T.S. 993, http://legal.un.org/avl/pdf/ha/sicj/icj_statute_e.pdf (stating that international custom, also known as customary international law, is "general practice accepted as law").

[21] *See* YORAM DINSTEIN, THE CONDUCT OF HOSTILITIES UNDER THE LAW OF INTERNATIONAL ARMED CONFLICT 115 (2004) ("The first and foremost inference from the basic rule is that direct—and deliberate—attacks against civilians are forbidden.").

4A (1) (2), (3) and (6)."[22]Where there is doubt, a person should be presumed a civilian."[23] Although this definition helps in understanding who qualifies as a civilian, it does not provide any guidance as to whether a civilian is DPH.

One of the fundamental principles of IHL is distinction, which provides that an attack should distinguish between civilians and combatants, and between civilian and military objectives.[24] Distinction predates modern IHL, and has been part of the law of wars jurisprudence for centuries – "*Slaughter of men armed and resisting is the law of war*… it is reasonable that they who have taken arms should be punished in battle, but that Non-combatants are not to be hurt."[25] The principle of distinction is comprised of various specific obligations. First, civilians can never be the object of direct attacks by a party to an armed conflict.[26] That is,

---

[22] Article 4(A) (1), (2), (3), and (6) of the Third Geneva Convention (Convention (III) relative to the Treatment of Prisoners of War. Geneva, 12 August 1949), reads:

> (1) Members of the armed forces of a Party to the conflict as well as members of militias or volunteer corps forming part of such armed forces.
>
> (2) Members of other militias and members of other volunteer corps, including those of organized resistance movements, belonging to a Party to the conflict and operating in or outside their own territory, even if this territory is occupied, provided that such militias or volunteer corps, including such organized resistance movements, fulfil the following conditions:
>
> > (a) that of being commanded by a person responsible for his subordinates;
> >
> > (b) that of having a fixed distinctive sign recognizable at a distance;
> >
> > (c) that of carrying arms openly;
> >
> > (d) that of conducting their operations in accordance with the laws and customs of war.
>
> (3) Members of regular armed forces who profess allegiance to a government or an authority not recognized by the Detaining Power.
>
> (6) Inhabitants of a non-occupied territory, who on the approach of the enemy spontaneously take up arms to resist the invading forces, without having had time to form themselves into regular armed units, provided they carry arms openly and respect the laws and customs of war.

[23] AP I, *supra* note 1, art. 50(1).

[24] AP I, *supra* note 1, arts. 51–52.

[25] HUGO GROTIUS, DE JURE BELLI AC PACIS LIBRI TRES 216 (William Whewell trans., London, John W. Parker 1853) (1625) (emphasis in original).

[26] AP I, *supra* note 1, art. 51(2).

adversaries are prohibited from aiming their weapons at civilians or civilian objects. Additionally, acts of violence, which are primarily intended to terrorize the civilian population, are prohibited.[27] Second, attacks which cannot sufficiently discriminate between civilians and combatants are prohibited.[28] Third, any attack resulting in collateral damage in the form of civilian casualties, injuries, and damage to civilian objects should be proportionate to the concrete and direct military advantage anticipated from such an attack[29] (colloquially referred to as "proportionality" and considered a distinct, fundamental principle of IHL).[30]

These broad civilian protections (also referred to as "non-combatant immunity"[31]) come with a specific exception. Namely, civilians only enjoy these protections "unless and for such time as they take a direct part in hostilities."[32] This exception represents the core principle that uninvolved civilians are protected from direct attacks, but, once they participate, they forfeit their civilian protections and become viable targets under IHL.[33] Although the exception appears in Additional Protocols, it represents customary international law and is binding upon all states regardless of ratification status.[34]

The commentary to Additional Protocol I defines direct participation in hostilities as "acts of war which are intended by their

---

[27] *Id*.
[28] AP I, *supra* note 1, art. 51(4).
[29] AP I, *supra* note 1, art. 51(1)(b).
[30] JEAN-MARIE HENCKAERTS & LOUISE DOSWALD-BECK, 1 INT'L COMM. OF THE RED CROSS, CUSTOMARY INTERNATIONAL HUMANITARIAN LAW 46–50 (2005), https://www.icrc.org/eng/assets/files/other/customary-international-humanitarian-law-i-icrc-eng.pdf (rule 14).
[31] *See, e.g.*, JUDITH GARDAM, NON-COMBATANT IMMUNITY AS A NORM OF INTERNATIONAL HUMANITARIAN LAW 116 (1993) (noting that the prevention of indiscriminate attacks on civilians is "the general rule of noncombatant immunity" in international humanitarian law).
[32] AP I, *supra* note 1, art. 51(3); AP II, *supra* note 1, art. 13(3).
[33] *See, e.g.*, Michael N. Schmitt, *Deconstructing Direct Participation in Hostilities: The Constitutive Elements*, 42 N.Y.U. J. INT'L L. & POL. 697, 702 (2010) (stating that "international humanitarian law suspends . . . civilian protections" when civilians directly participate in hostilities).
[34] *See* HCJ 769/02 Pub. Comm. Against Torture in Isr. v. Gov't of Israel 62(1) PD 507 (2006) (Isr.), http://elyon1.court.gov.il/files_eng/02/690/007/A34/02007690.a34.pdf, ¶ 12 (noting that the exception to civilian protections "reflects a customary rule of international law" and thus binds all states).

nature or their purpose to hit specifically the personnel and the *matériel*[35] of the armed forces of the adverse Party."[36] It is highly significant that the temporal scope of the targetability of civilians engaged in DPH remains unclear and continues to generate considerable controversy.[37] Moreover, civilians who DPH are not granted immunity from prosecution under domestic criminal law for their activities in the hostilities, as the immunity is associated with combatants who belong to recognized armed forces.[38] However, civilians engaging in DPH who are captured while committing their hostile activities are granted certain minimal protections under Article 75 of Additional Protocol I[39] or Common Article 3,[40] depending on the classification of the armed conflict.[41]

In this context, the ICRC Interpretive Guidance provides a framework to address the phenomenon of civilians who DPH.[42] Although the Interpretive Guidance has garnered approval and is generally well

---

[35] *See* MERRIAM-WEBSTER DICTIONARY 835 (11th ed. 2004) (defining "matériel" as "equipment and supplies used by soldiers")

[36] ICRC COMMENTARY ON APS, *supra* note 18, at 516.

[37] *Id*. at 1453 ("If a civilian participates directly in hostilities, it is clear that he will not enjoy any protection against attacks for as long as his participation lasts. Thereafter, as he no longer presents any danger for the adversary, he may not be attacked; moreover, in case of doubt regarding the status of an individual, he is presumed to be a civilian. Anyone suspected of having taken part in hostilities and deprived of his liberty for this reason will have the benefit of the provisions laid down in Articles 4 (Fundamental guarantees), 5 (Persons whose liberty has been restricted), and 6 (Penal prosecutions).").

[38] *See* Kenneth Watkin, *Warriors Without Rights? Combatants, Unprivileged Belligerents, and the Struggle over Legitimacy*, HARV. PROGRAM ON HUMANITARIAN POL'Y & CONFLICT RES. OCCASIONAL PAPER SERIES, No. 2, Winter 2005, at 12–13 (noting that combatants "have a special status . . . [which includes] the right to participate in hostilities and receive immunity from prosecution . . . for killing carried out in accordance with the law.").

[39] AP I, *supra* note 1, art. 75 (international armed conflict).

[40] Geneva Convention Relative to the Treatment of Prisoners of War art. 3, Aug. 12, 1949, 6 U.S.T. 3316, 75 U.N.T.S. 135 [hereinafter Common Article 3]. In non-international armed conflict, such detainees would be granted humane treatment, and Common Article 3 would apply, along with customary international law and human rights law, where a gap exists.

[41] *See* THE HANDBOOK OF INTERNATIONAL HUMANITARIAN LAW 44, 49 (Dieter Fleck ed., 3d ed. 2013) (stating that "[a]n international armed conflict exists if one state uses armed force against another state" and that "[a] non-international armed conflict is a situation of protracted armed violence between governmental authorities and organized armed groups or between such groups within a state").

[42] *See generally* INTERPRETIVE GUIDANCE, *supra* note 14.

received in academic and legal-military communities, it only represents the institutional view of the ICRC.[43] Due to the lack of precedential law regarding DPH and the reputation of the ICRC as having expertise in IHL, the Interpretive Guidance is often the only framework used to determine whether or not a civilian is engaging in DPH.[44] The conclusions of the Interpretive Guidance are sometimes controversial and highly contested, and there is substantial criticism of these determinations from both military and academic perspectives.[45]

The core of the DPH paradigm as provided by the Interpretive Guidance lies in three criteria that are required in order for an individual to be considered a civilian engaged in DPH.[46] First, a certain threshold of harm needs to be met.[47] Second, there needs to be a direct causal link between the act in question and the harm caused.[48] Third, the act needs to be designed to cause harm in support of one party to the conflict and to the detriment of the opposing party of the conflict (also known as a "belligerent nexus").[49]

A. Threshold of Harm

Not every hostile action by civilians will reach the threshold required for the DPH label. The Interpretive Guidance requires that the act in question reach a certain severity threshold in relation to the harm caused (or *likely* to be caused) in order for that individual to forfeit his civilian status.[50] The Interpretive Guidance provides two tests for determining whether the threshold of harm has been reached. First, an act must be *likely* to adversely affect the military operations or capacity of a party to the

---

[43] *See* INTERPRETIVE GUIDANCE, *supra* note 14, at 9 (stating that "the 10 recommendations made by the Interpretive Guidance, as well as the accompanying commentary, do not endeavour to change binding rules of customary or treaty IHL, but reflect the ICRC's institutional position as to how existing IHL should be interpreted").

[44] *See, e.g.*, Schmitt, *supra* note 33.

[45] *See* Ryan Goodman & Derek Jinks, *The ICRC Interpretive Guidance on the Notion of Direct Participation in Hostilities Under International Humanitarian Law: An Introduction to the Forum*, 42 N.Y.U. J. INT'L. L. & POL. 637 (2010).

[46] INTERPRETIVE GUIDANCE, *supra* note 14, at 46.

[47] *Id.* at 48.

[48] *Id.* at 51.

[49] *Id.* at 58.

[50] *Id.* at 47 ("The qualification of an act as direct participation does not require the materialization of harm reaching the threshold but merely the objective likelihood that the act will result in such harm").

armed conflict.[51] Second, an act must be *likely* to inflict death, injury, or destruction on persons or objects protected against direct attack.[52] The likelihood standard is evaluated objectively in each case.[53]

While the threshold of likelihood of "death, injury or destruction on persons or objects" is quite straightforward, the threshold of the likelihood to "adversely affect the military capacity or capacity of a party to the armed conflict" is somewhat obscure. An act that could result in death, injury or destruction has the same consequences as an act that is likely to affect military capacity or capacity of a party to the armed conflict. In other words, in both cases, the civilian who carries out the hostile act would be targetable, should he satisfy the remaining two requirements. In *Targeted Killings*, the Israeli Supreme Court suggested that "acts which by nature and objective are intended to cause damage to civilians" be part of the threshold of harm requirement, however that requirement may already be included in the "death, injury or destruction."[54]

Currently, it is unlikely that a cyber-attack will be objectively likely to cause death, injury, or destruction, although it might be possible with destructive cyber-attacks against critical cyber infrastructure, such as hospitals and power plants.[55] Conversely, cyber-attacks can be highly disruptive and substantially affect military capacity.[56] However, it is debatable whether a particular cyber-attack actually affects military capacity or simply represents a form of expression, propaganda, or mild nuisance. The Interpretive Guidance excluded "manipulation of computer networks" from the application of the DPH framework, even though such acts might "have a serious impact on public security, health and

---

[51] *Id.* at 46.

[52] INTERPRETIVE GUIDANCE, *supra* note 14, at 47.

[53] *Id.*

[54] HCJ 769/02 Pub. Comm. Against Torture in Isr. v. Gov't of Israel 62(1) PD 507 (2006) (Isr.), http://elyon1.court.gov.il/files_eng/02/690/007/A34/02007690.a34.pdf, ¶ 33.

[55] *See, e.g.*, Scott A. Newton, *Can Cyberterrorists Actually Kill People?*, SANS INST. INFOSEC READING ROOM (2002), https://www.sans.org/reading-room/whitepapers/warfare/cyberterrorists-kill-people-820.

[56] *See* Shane Quinlan, *Jam. Bomb. Hack? U.S. Cyber Capabilities and the Suppression of Enemy Air Defenses*, GEO. SECURITY STUD. REV. (Apr. 7, 2014), http://georgetownsecuritystudiesreview.org/2014/04/07/jam-bomb-hack-new-u-s-cyber-capabilities-and-the-suppression-of-enemy-air-defenses/ (explaining how Israel bombed a Syrian nuclear reactor and managed to avoid detection by the aerial defense systems using a cyber-attack).

commerce."[57] The Interpretive Guidance briefly addressed the possibility of cyber-attacks affecting military capacity by providing that "electronic interference with military computer networks could also suffice, whether through computer network attacks (CNA) or computer network exploitations (CNE), as well as wiretapping the adversary's high command or transmitting tactical targeting information for an attack."[58]

B. Direct Causal Link

The second requirement of the DPH framework is that the act in question directly causes the harm.[59] In a modern society, the activities and occupations that civilians pursue could contribute to the military defeat of an adversary. Examples include attacking weapons production, construction, political propaganda, production of goods, and the supply of electricity and fuel.[60] However, even though these activities could ultimately result in severe harm, reaching the threshold of harm as set in the first requirement, they may not satisfy the second requirement of a direct causal link.[61] Therefore, civilians who contribute to war efforts in such indirect ways would not be directly participating in hostilities and their protected civilian status would remain intact.[62]

The prevailing test to determine causation is determining whether the harm caused is only one causal step away from the act.[63] Therefore, all the capacity building, services, and production activities are far removed in the causal chain and will not be considered as direct causal acts. However, the Interpretive Guidance argues that acts that in isolation do not cause harm, when done in conjunction with other acts that cause the

---

[57] INTERPRETIVE GUIDANCE, *supra* note 14, at 50.

[58] INTERPRETIVE GUIDANCE, *supra* note 14, at 48 & n.101 ("CNA have been tentatively defined as 'operations to disrupt, deny, degrade, or destroy information resident in computers and computer networks, or the computer and networks themselves' and may be conducted over long distances through radio waves or international communication networks. While they may not involve direct physical damage, the resulting system malfunctions can be devastating. CNE, namely 'the ability to gain access to information hosted on information systems and the ability to make use of the system itself,' though not of a direct destructive nature, could have equally significant military implications. During the expert meetings, CNA causing military harm to the adversary in a situation of armed conflict were clearly regarded as part of the hostilities.") (citations omitted).

[59] *Id.* at 52

[60] *Id.*

[61] *Id.*

[62] *Id.*

[63] *Id.*

required threshold of harm, will be seen as direct causation acts (for example – transmission of intelligence, instructing and assisting troops in executing military operations, etc.).[64]

An important point that also relates to cyber-attacks is the distinction between *causal* proximity and *temporal* or *geographic* proximity.[65] The fact that a certain hostile act is delayed or distant does not necessarily affect the *causality* of that act. An example of both a delayed and distant hostile act might be a cyber-attack carried out by State A against State B over international borders. State A collects military intelligence during the attack, and the harm from the attack materializes after a certain period of time. Although the cyber-attack was initiated by one nation-state and materially affected a state thousands of miles away after time had elapsed (i.e. no *temporal* or *geographical* proximity), it is nonetheless clear that there is *causal* proximity between the cyber-attack and the harm. In this case, the adverse effect is on the military operations of State B against State A.

Additionally, the effects of most cyber operations might be secondary or tertiary in nature, further complicating the direct causation requirement.[66] Particularly challenging are cyber operations that cause effects in multiple causal steps. This is especially the case with unaware actors being involved in the form of intermediate, compromised computer

---

[64] *Id*. at 55.

[65] *Id*.

[66] *See* TECHNOLOGY, POLICY, LAW, AND ETHICS REGARDING U.S. ACQUISITION AND USE OF CYBERATTACK CAPABILITIES 127 (William A. Owens, Kenneth W. Dam & Herbert S. Lin eds., 2009) ("One of the most difficult-to-handle aspects of a cyberattack is that in contrast to a kinetic attack that is almost always intended to destroy a physical target, the desired effects of a cyberattack are almost always indirect, which means that what are normally secondary effects are in fact of central importance. In general, the planner must develop chains of causality—do *X*, and *Y* happens, which causes *Z* to happen, which in turn causes *A* to happen. Also, many of the intervening events between initial cause and ultimate effect are human reactions (e.g., in response to an attack that does *X*, the [target] network administrator will likely respond in way *Y*, which means that *Z*—which may be preplanned—must take response *Y* into account). Moreover, the links in the causal chain may not all be of similar character—they may involve computer actions and results, or human perceptions and decisions, all of which combine into some outcome.").

systems and networks.[67] In cases like this, the direct causation requirement might be the hardest to satisfy in the cyber operation context.[68]

C. Belligerent Nexus

The belligerent nexus requirement is the third and last condition for a civilian to be considered a DPH. It requires that the hostile act in question is *specifically designed* to cause the required threshold of harm and is *specifically designed* to do so in *support* of one party to the armed conflict at the detriment of the opposing party.[69] This requirement serves two functions. First, it targets acts that happen in the context of an armed conflict with the purpose of empowering one party over the other. Second, it ignores acts that happen regardless of an ongoing armed conflict, even though the harm caused could reach the required threshold and the harm was directly caused by the act. Thus, when hostile actions are not related to the armed conflict, the civilian is not a DPH and the appropriate punishment will come from normal law enforcement. In other words, the act is not fulfilling the belligerent nexus requirement.[70] The Interpretive Guidance provides that the determination of whether a belligerent nexus exists in relation to a specific act–

> must be based on the information reasonably available to the person called on to make the determination, but they must always be deduced from objectively verifiable factors. In practice, the decisive question should be whether the conduct of a civilian, in conjunction with the circumstances prevailing at the relevant time and place, can reasonably be perceived as an act designed to support one party to the conflict by directly causing the required threshold of harm to another party.[71]

Cyber-attacks happen on a daily basis, both inside and outside of armed conflict. Therefore, the belligerent nexus requirement can be challenging to establish should a cyber-attack occur in an armed conflict context. Such cyber-attacks might fulfill the threshold of harm and direct

---

[67] *See* David Turns, *Cyber Warfare and the Notion of Direct Participation in Hostilities*, 17 J. CONFLICT & SECURITY L. 279, 288 (2012).

[68] *Id.* at 296; *see also* Vijay M. Padmanabhan, *Cyber Warriors and the* Jus in Bello, 89 INT'L L. STUD. 288, 298 (2013) ("The 'direct causation' requirement appears easier to meet in the context of cyber operations than in traditional kinetic operations.").

[69] INTERPRETIVE GUIDANCE, *supra* note 14, at 58.

[70] *Id*. at 64.

[71] *Id.* at 63–64.

causation requirements, yet the determination that the cyber-attack meets the belligerent nexus requirement may be attenuated.

D. The Notion of "Continuous Combat Function"

The notion of continuous combat function (CCF) is an essential supplement to the three requirements associated with the DPH framework.[72] It is generally believed that the DPH status is temporal, that is, it only applies to civilians *for such time* as they take direct part in hostilities,[73] so civilians who carry out hostile acts sporadically, spontaneously, or on an unorganized basis are targetable only while they are actively carrying out these acts.[74] However, there is an assumption that when a civilian assumes an integral role in an organized armed group and whose "continuous function involves the preparation, execution, or command of acts or operations amounting to direct participation in hostilities," the civilian is CCF.[75] According to the principle of CCF, such a civilian would be targetable even if he or she is not carrying out hostile acts at the moment of the targeting. In other words, a civilian who is CCF is targetable due to his or her *status*, which materialized as a result of his or her continuous and persistent *actions*. In contrast, a DPH civilian is targetable only for such time as he carries out his or her hostile acts. This means that:

> civilians lose protection against direct attack for the duration of each specific act amounting to direct participation in hostilities, whereas members of organized armed groups belonging to a non-state party to an armed conflict cease to be civilians, and lose protection against direct attack for as long as they assume their continuous combat function.[76]

Thus, a hacker who carries out cyber-attacks in an armed conflict continuously and assumes membership in an organized armed group that is a party to the armed conflict would be considered a civilian with CCF. Consequently, this hacker is targetable at all times, as long as the CCF label remains.

II. BETWEEN INTERGOVERNMENTAL AND DOMESTIC – THE TALLINN MANUAL AND THE U.S. LAW OF WAR MANUAL

The Interpretive Guidance represents the institutional view of the ICRC regarding the international humanitarian law applicable to the

---

[72] *Id.* at 33.
[73] AP I, *supra* note 1, art. 51(3).
[74] INTERPRETIVE GUIDANCE, *supra* note 14, at 34.
[75] *Id.*
[76] *Id.* at 70.

notion of DPH. While it is an authoritative and compelling document, it does not represent binding international law. There are more concrete implementations of DPH in the intergovernmental context, such as The Tallinn Manual, and in the domestic context, the U.S. Law of War Manual. Thus, it is useful to analyze the Tallinn Manual and U.S. Law as they have slightly differing views on the notion of DPH as it applies to cyber-attacks.

A. The Tallinn Manual on the International Law Applicable to Cyber Warfare

The Tallinn Manual is an initiative of the NATO Cooperative Cyber Defense Centre of Excellence, based in Tallinn, Estonia. The Tallinn Manual was published in 2012, and it is a non-binding set of rules, based on the rules pertaining to the use of force (in the UN Charter sense) and IHL. It was unanimously agreed upon by the group of experts assigned to draft the Tallinn Manual. The Tallinn Manual consists of black letter rules and a commentary for each rule presented. According to the Manual, the rules reflect *lex lata* (existing law) and not *lex ferenda* (the law that should be).[77]

Rule 35 of the Tallinn Manual, which is precisely the same as Article 51(3) of Additional Protocol I and Article 13(3) of Additional Protocol II, reads "civilians enjoy protection against attack unless and for such time as they directly participate in hostilities."[78] The commentary that accompanies this rule sets forth some of the differences between the Interpretive Guidance and the Tallinn Manual.

With regard to the threshold of harm, while the Interpretive Guidance requires the materialization of harm or if harm did not materialize, the objective likelihood standard, the Tallinn Manual uses the term "*intended* or actual effect."[79] In other words, objective likelihood is not required for the threshold of harm requirement, but instead actual harm or individual intent to cause that harm is used. This is a lower standard than the one suggested by the Interpretive Guidance. Therefore, a civilian with intent to cause sufficient harm, who carries out a sloppy cyber-attack with no chance of affecting the targeted adversary, will still lose his protection from direct attack since the cyber-attack would still be seen as

---

[77] *See* Ido Kilovaty, *Cyber Warfare and the* Jus ad Bellum *Challenges: Evaluation in the Light of the* Tallinn Manual on the International Law Applicable to Cyber Warfare, 5 AM. U. NAT'L. SECURITY L. BRIEF, no. 1, 2014, at 91, 96.
[78] TALLINN MANUAL, *supra* note 15, at 118 (rule 35).
[79] *Id.* at 119 (rule 35, cmt. 4) (emphasis added).

DPH under the Tallinn Manual approach.[80] Such civilians who might not be the savviest hackers, might be more easily targeted due to the failure of their cyber operation.

The Tallinn Manual applies the intent standard to the second requirement of direct causation as well. It requires a "direct causal link between the act in question and the harm *intended* or inflicted."[81]

The Tallinn Manual also describes the third requirement of belligerent nexus as "acts [that] must be directly related to the hostilities."[82] At first glance, it is a more expansive approach than the one suggested by the Interpretive Guidance. However, the footnote to that requirement is a reference to the Interpretive Guidance definition of "an act [that] must be specifically designed to directly cause the required threshold of harm in support of a party to the conflict and to the detriment of another."[83] This suggests that the Tallinn Manual is actually in line with the Interpretive Guidance, and thus, this paper tends to disagree with the view that the Tallinn Manual's perspective of belligerent nexus is more expansive than the one in the Interpretive Guidance.[84]

B. U.S. Department of Defense Law of War Manual

The recent U.S. Department of Defense Law of War Manual represents, to some extent, a set of rules to govern the conduct of hostilities in armed conflicts. In a sense, it suggests desirable norms for cyber warfare as far as the U.S. is concerned.[85] Although, naturally, the Law of War Manual binds the U.S. alone, its understanding of the DPH framework in cyberspace is analyzed in this section in an attempt to compare that understanding to the DPH norms advanced by the ICRC and NATO.

Rule 16.5.5 of the Law of War Manual entitled "Use of Civilian Personnel to Support Cyber Operations" provides that States are not prohibited under the law of war to employ civilian personnel to carry out cyber operations.[86] This proposition, according to the Manual, is also true if the cyber operation amounts to DPH.[87] Later in this Rule, the Manual

---

[80] *See* Collin Allan, Note, *Direct Participation in Hostilities from Cyberspace*, 54 VA. J. INT'L. L. 173, 182 (2013).
[81] TALLINN MANUAL, *supra* note 15, at 119 (rule 35 cmt. 4).
[82] *Id.*
[83] *Id.* at 119–20 n.65.
[84] *See* Allan, *supra* note 75, at 189.
[85] *See* LAW OF WAR MANUAL, *supra* note 17, at 1007.
[86] *Id.*
[87] *Id.*

provides that "Civilians who take a direct part in [cyber] hostilities forfeit protection from being made the object of attack."[88] However, the Manual does not give any guidance as to the notion of direct participation in *cyber* hostilities. Nonetheless, there is a separate chapter dealing with direct participation in *non-cyber* hostilities.[89]

Interestingly, the Manual clarifies that the United States did not adopt the definition of direct participation in hostilities as provided by Additional Protocol I[90] or the ICRC Interpretive Guidance.[91] Although this is the case, the Manual agrees that certain parts of the Interpretive Guidance are reflective of customary international law, while other parts are not.[92] The general DPH framework, which the Manual suggests is a departure from the three-tiered structure, was adopted by both the Interpretive Guidance and the Tallinn Manual.[93]

First, the Manual focuses on the threshold requirement, by requiring either "proximate or "but for" cause of death, injury, or damage to persons or objects" or an adverse effect on military operations or capacity.[94] Already in the description of the required harm, the Manual sets the standard of causation by demanding "proximate or "but for" cause."[95] The causation test provided by the Manual is somewhat looser than the one advanced by the Interpretive Guidance, since it could be satisfied with proximity and could apply to a broader set of cyber operations.

Second, the Manual sets forth the grounds for the belligerent nexus requirement, providing that hostile acts should be evaluated by the degree to which they are connected to military operations *or* by "the degree to which the act is temporally or geographically near the fighting."[96] The first condition is somewhat similar to the ICRC's

---

[88] *Id*. at 1008.
[89] *See id*. (describing rule 5.9).
[90] *Id*. at 222
[91] *See* Stephen Pomper, *Toward a Limited Consensus on the Loss of Civilian Immunity in Non-International Armed Conflict: Making Progress Through Practice*, 88 INT'L L. STUD. 181, 186 (2012) ("From the operational perspective, the feedback [on the ICRC's INTERPRETIVE GUIDANCE] was that the report was too rigid and complex, and did not give an accurate picture of State practice or (in some respects) of a practice to which States could realistically aspire.").
[92] LAW OF WAR MANUAL, *supra* note 17, at 223.
[93] *See id*. at 226–27 (discussing the full structure of the LAW OF WAR MANUAL).
[94] *Id*. at 226.
[95] *Id*.
[96] *Id*.

belligerent nexus interpretation. However, the broader term "military operations" is used in place of hostilities. Moreover, the condition of temporal and geographical proximity is again a departure from how belligerent nexus is generally understood, and in any case, represents a complication to cyber hostilities that might be temporally and geographically distant from the fighting.[97] The following consideration put forth by the Manual is actually closer to the Interpretive Guidance understanding of the belligerent nexus requirement, although the state of mind required is slightly different. The Manual requires that the act be "*intended* to advance the war aims of one party to the conflict to the detriment of the opposing party."[98] However, the Interpretive Guidance requires an act that is "*specifically designed* to directly cause the required threshold of harm in support of a party to the conflict and to the detriment of another."[99] Therefore, the Manual provides a standard that is lower and only requires intent, as opposed to the more objective standard of specific design.

Third, the Manual sets forth the military significance consideration, which focuses on the degree of contribution of the hostile act to the military action against the opposing party. It seeks to determine whether the value of contribution of the hostile act is comparable or greater than the more common DPH acts, or whether the hostile act poses significant risk to the targeted party.[100] These considerations are not part of the Interpretive Guidance or Tallinn Manuals' frameworks. However, it is reasonable to find them in a military manual, which provides a more specified set of considerations to support military operations.

Fourth, the Manual suggests evaluating whether the act is traditionally military, that is, whether the hostile act is performed traditionally by military forces (such as "combat, combat support, and combat service support functions"[101]) or whether there is a decision-making process as to the use or application of combat power.[102] However, it is unclear how this consideration could be applied to cyber-attacks, since cyber-attacks represent an emerging phenomenon that cannot at this point qualify as traditional, and there is insufficient data pertaining to the military application of cyberspace capabilities in times of armed conflict.

It is critical to understand that the aforementioned factors are considerations according to the Manual, and they provide some guidance

---

[97] ROSCINI, *supra* note 8, at 207.
[98] LAW OF WAR MANUAL, *supra* note 17, at 226 (emphasis added).
[99] INTERPRETIVE GUIDANCE, *supra* note 14, at 58 (emphasis added).
[100] LAW OF WAR MANUAL, *supra* note 17.
[101] *Id*. at 227.
[102] *Id.*

for evaluation of whether a specific hostile act qualifies as DPH. However, the factors differ fundamentally from the views presented by the Interpretive Guidance and the Tallinn Manual, which have three basic requirements to be fulfilled in order for a civilian to qualify as a direct participant in hostilities.

III. DIRECT PARTICIPATION IN '*HACKTIVITIES*' – SHORTCOMINGS

Cyber-attacks in modern armed conflicts are only one factor that exacerbates the gaps and ambiguities within the DPH framework. As demonstrated in earlier sections, the DPH framework can be indeterminate, and there is a deep and fundamental disagreement as to the interpretation and application of critical points within the DPH analysis. These shortcomings will be discussed and analyzed in this section, with a focus on the difficulty regarding the characterization of a hostile cyber act as DPH and the temporal scope challenge.

A. Boundaries of the Harm Threshold Cyber Activities

The DPH threshold requirements are particularly opaque and amorphous in the context of cyber operations. Nevertheless, they contain at their core certain general notions that may be helpful in relation to certain cyber operations. Specific examples of such cyber operations exist, and they will be discussed in detail in the following sub-section.

1. Passive v. Active Cyber Defenses

The DPH framework is inadequate when it comes to cyber operations intended to enhance the cyber defense of one party, while not directly targeting the opposing party to the armed conflict.[103] Moreover, when a cyber operation only incidentally affects military operations of the opposing party, or when there is no direct causal link, the only possible conclusion is that such activity would not qualify as DPH. However, certain authors believe that the DPH framework applies to such acts.[104]

The Tallinn Manual explicitly addressed passive cyber defense by providing that "*[s]ome* members of the International Group of Experts took the position that acts that enhance one's own military capacity are included, as they necessarily weaken an adversary's relative position. An example is maintaining *passive* cyber defenses of military cyber assets."[105] However, before analyzing the position of the Tallinn Manual, it is essential to define active as opposed to passive cyber defenses, as often the two concepts are conflated.

---

[103] ROSCINI, *supra* note 8, at 205.
[104] *Id.*
[105] TALLINN MANUAL, *supra* note 15, at 119 (emphasis added).

Jeffrey Carr, in his landmark book "Inside Cyber Warfare," defined active cyber defenses as

> …electronic countermeasures designed to strike attacking computer systems and shut down cyber attacks midstream. Security professionals can set up active defenses to automatically respond to attacks against critical systems, or they can carry them out manually. For the most part, active defenses are classified, though programs that send destructive viruses back to the perpetrator's machine or packet-flood the intruder's machine have entered the public domain.[106]

According to Carr's definition, active cyber defenses can be manual, meaning that professionals carry them out as needed in response to actual or anticipated cyber-attacks, or automatic, meaning that the targeted computer system is set to respond using certain cyber measures if a cyber-attack occurs.[107]

Carr goes on to define passive defenses by limiting them to "traditional forms of computer security used to defend computer networks, such as system access controls, data access controls, security administration, and secure system design."[108] The focus in passive cyber defenses is that there is no hawkish response to cyber-attacks, but there are defensive layers resident in the computer systems that are meant to restrict or at least make it harder to access and manipulate the computer system.[109]

According to this analysis, it seems that if civilians are engaged in *active* cyber defense measures, such as actually countering cyber-attacks that reach the threshold of harm, they would be considered directly participating in hostilities. In contrast, if civilians are only involved in *passive* cyber defense measures, such as installing anti-virus software, encrypting critical data and securing the system, they would fulfill neither the direct causal link requirement nor the requisite threshold of harm.[110]

---

[106] JEFFREY CARR, INSIDE CYBER WARFARE 46 (2nd ed. 2011).

[107] *Id.*

[108] *Id.*

[109] *See* Matthew J. Sklerov, *Solving the Dilemma of State Responses to Cyberattacks: A Justification for the Use of Active Defense Against States Who Neglect Their Duty to Prevent*, 201 MIL. L. REV. 1, 8 (2009); s*ee also* Oona Hathaway et al., *The Law of Cyber-Attack*, 100 CAL. L. REV. 817, 858 (2012) (discussing that "countermeasures might go beyond 'passive defenses' that aim to repel cyber-attacks (such as firewalls), and constitute 'active defenses,' which attempt to disable the source of an attack.").

[110] *See* Emily Crawford, *Virtual Battlegrounds: Direct Participation in Cyber Warfare*, 9 J.L. & POL'Y FOR INFO. SOC'Y 1, 15 (2012) ("[W]hile civilians

The Law of War Manual recognizes that some civilians might be authorized to support cyber operations, and it provides that "[a]s with non-cyber operations, the law of war does not prohibit States from using civilian personnel to support their cyber operations."[111] However, the Law of War Manual itself does not make the distinction between support that constitutes DPH and support that does not.[112]

The Tallinn Manual defines "passive cyber defense" as "a measure for detecting and mitigating cyber intrusions and the effects of cyber attacks that does not involve launching a preventive, pre-emptive or countering operation against the source. Examples of passive cyber defense measures are firewalls, patches, anti-virus software, and digital forensic tools."[113] Even given this definition, the Tallinn Manual concludes that such measures would qualify as DPH if carried out by civilians.[114] This assertion, following the analysis provided in this section, is counterintuitive, and interestingly, Michael Schmitt himself (the General Editor of the Tallinn Manual) was less confident about that assertion. In a recent article, he asked, "Is passive cyber defense of enemy systems an act of direct participation such that contractors who perform the task lose their immunity from attack?"[115] The fact that passive cyber defenses alone are becoming less efficient than active cyber defenses will increase the involvement of civilians in active defensive cyber operations and might expose them to the risk of being labeled as DPH.[116] The

---

employed to generally maintain computer networks for an armed force (in the capacity of general IT services such as email, websites, etc.) would likely not be considered as taking direct part in hostilities, any employee or contractor who was specifically employed to conduct hostile CNA/CNE would, in theory, be considered as taking direct part in hostilities.").

[111] LAW OF WAR MANUAL, *supra* note 17, at 1007.

[112] *Id.*

[113] TALLINN MANUAL, *supra* note 15, at 261.

[114] *Id*. at 119.

[115] Michael N. Schmitt & Sean Wattset, *The Decline of International Humanitarian Law* Opinio Juris *and the Law of Cyber Warfare*, 50 TEX. INT'L. L.J. 189, 228 (2015).

[116] *See* Pascal Brangetto, Tomáš Minárik & Jan Stinissen, *From Active Cyber Defence to Responsive Cyber Defence: A Way for State to Defend Themselves – Legal Implications*, NATO LEGAL GAZETTE, Dec. 2014, at 16–17 ("The mere fact of labelling current defensive tools as passive is a call for a more empowering definition of cyber defence. The use of only passive measures is no longer sufficient to protect networks in the face of rising threat levels. As a way to overcome this, and to be able to hold the high ground, a concept was developed to enable the defending party to play an active part in its own cyber defence.")

recommended approach is to distinguish between acts designed to enhance "general capacity to carry out unspecified hostile acts" and acts that are "aimed to carry out a specific hostile act."[117]

2. Data Destruction and Alternation

Cyber operations carried out by civilians that directly alter military data required for upcoming military operations will most likely qualify as adversely affecting military operations. According to some, such cyber operations would be equivalent to cyber operations that directly disrupt unmanned aerial vehicle systems, as well as radars and weapons that are operated by computer systems.[118]

While destruction of sensitive military data that directly and adversely affects military operations is a clear-cut case of direct participation in hostilities, there are far more difficult cases of data destruction such as destruction of medical data belonging to certain individuals.[119] The immediate question then is whether such destruction would qualify as "destruction to objects" under the threshold of harm requirement.[120]

The Interpretive Guidance excludes "manipulation of computer networks" from the application of the DPH framework,[121] while the

---

[117] INTERPRETIVE GUIDANCE, *supra* note 14, at 66; *see also* Allan, *supra* note 80, at 191.

[118] *See* Nils Melzer, *Cyberwarfare and International Law*, UNITED NATIONS INSTITUTE FOR DISARMAMENT RESEARCH, 2011, at 28, http://www.unidir.org/files/publications/pdfs/cyberwarfare-and-international-law-382.pdf ("In line with this interpretation, cyber operations aiming to disrupt or incapacitate an adversary's computer-controlled radar or weapons systems, logistic supply or communication networks may not directly cause any physical damage, but would certainly qualify as part of the hostilities and, therefore, would have to comply with the rules and principles of IHL governing the conduct of hostilities. The same would apply to cyber operations intruding into the adversary's computer network to delete targeting data, manipulate military orders, or change, encrypt, exploit, or render useless any other sensitive data with a direct (adverse) impact on the belligerent party's capacity to conduct hostilities.") (citations omitted).

[119] François Delerue, *Civilian Direct Participation in Cyber Hostilities*, Revista de Internet, Derecho y Política, Oct. 2014, at 9, http://journals.uoc.edu/index.php/idp/article/download/n19-delerue/n19-delerue-en.

[120] *Compare* INTERPRETIVE GUIDANCE, *supra* note 14, at 47 (discussing "destruction on . . . objects protected against direct attack") *with* AP I, *supra* note 1, art. 52(1) ("Civilian objects shall not be the object of attack or of reprisals. Civilian objects are all objects which are not military objectives.").

[121] INTERPRETIVE GUIDANCE, *supra* note 14, at 50.

Tallinn Manual does not explicitly address destruction and alteration of data in the sense of DPH. By analogy to Tallinn Manual's definition of "Cyber Attack"[122] it can be inferred that the Tallinn Manual sees the destruction of data as a cyber-attack only if it results in physical effects, such as "injury or death of individuals or damage or destruction to physical objects."[123] Such an approach fails to acknowledge the importance of data in a modern, interconnected, and cyber-reliant society. An example demonstrating the inadequacy of this approach compares a cyber operation resulting in the complete deletion of data belonging to an entire State's banking system with the physical destruction of a single data center. The data deletion is absent of injury, death or physical damage, would not qualify as a "cyber-attack," and would unlikely be considered as DPH, whereas the destruction of a data center would fall under the "death, injury or destruction to objects" and would be considered as DPH.[124]

This paper's key recommendation is to overcome the obsolete confines of physicality and to understand that in today's world, data can be just as crucial to the overall wellbeing of a society as other physical objects. In discussing whether civilian data should be defined as an "object" under IHL, thereby prohibiting attacks on civilian data because such attack would be a violation of the prohibition on direct attacks against civilians and civilian objects,[125] Michael Schmitt observed that while there is no definitive answer to the question, "data should not be characterized as an object in itself."[126] Even though not all targeted data would lead to severe consequences, some targeted data would nonetheless constitute a potential "civilian object" that enables the cyber operation to reach the threshold of harm because the operation would constitute an infliction of "destruction on . . . objects protected against direct attack."[127] As Schmitt suggested, there are two instances in which data should be considered an "object" under IHL. First, if the data is "*directly transferable* into tangible

---

[122] Which, similarly to the threshold of harm requirement, the TALLINN MANUAL, at 107, defines as "cyber operation, whether offensive or defensive, that is reasonably expected to cause injury or death to persons or damage or destruction to objects."

[123] *See* Rain Liivoja & Tim McCormack, *Law in Virtual Battlespace: The Tallinn Manual and the* Jus in Bello, 15 Y.B. INT'L HUMANITARIAN L. 45, 53 (2012).

[124] *Id.*

[125] For a comprehensive discussion of data as civilian objects, see Michael N. Schmitt, *The Notion of 'Objects' During Cyber Operations: A Riposte in Defence of Interpretive and Applicative Precision*, 48 ISR. L. REV. 81 (2015).

[126] Michael N. Schmitt, *Cyber Operations and the* Jus in Bello*: Key Issues*, 87 INT'L L. STUD. 89, 96 (2011).

[127] Melzer, *supra* note 113, at 49.

objects."[128] Banking data that represents data pertaining to actual physical objects (money), if deleted or altered, would therefore constitute destruction to physical objects. Second, data with intrinsic value, such as digital art, should also be considered an "object."[129]

While protecting digital art is important, the focus should be on protecting data that is not necessarily directly transferable to physical objects, but nevertheless constitutes an integral part of societal structure and order. Examples include police records, maps, and academic research, to name a few. These examples represent data that is not directly transferable to physical objects, but is nonetheless essential for every society in its day-to-day affairs. Moreover, it is immaterial whether data is digital or physical, since in both cases its destruction would cause harm, and claiming that only physical destruction counts for the purpose of DPH is counterintuitive and detrimental to the protection of the civil society. Although the DPH paradigm focuses on very physical aspects of destruction, when it comes to cyberspace, it is time to appreciate the importance and centrality of data and to protect it from exploitation.

3. Critical Infrastructure Disruption & Civilian Nuisance and Inconvenience

Another problem arising from the threshold of harm requirement is the relationship to harm to civilians and civilian infrastructure. While it is clear that civilian harm (other than death, injury or destruction) accompanied by adverse military effect will reach the threshold, it is unclear whether harm to civilian life that results in major inconveniences, or even terrorization of civilians,[130] could cross the threshold alone.[131] This question relates to "destruction" in the non-physical sense[132] and

---

[128] Schmitt, *supra* note 121, at 96 (emphasis added).
[129] *Id.*
[130] *See* ICRC COMMENTARY ON APS, *supra* note 18, at 618 ("[T]here is no doubt that acts of violence related to a state of war almost always give rise to some degree of terror among the population and sometimes also among the armed forces. It also happens that attacks on armed forces are purposely conducted brutally in order to intimidate the enemy soldiers and persuade them to surrender. This is not the sort of terror envisaged here. This provision is intended to prohibit acts of violence the primary purpose of which is to spread terror among the civilian population without offering substantial military advantage. It is interesting to note that threats of such acts are also prohibited. This calls to mind some of the proclamations made in the past threatening the annihilation of civilian populations.") (citations omitted).
[131] ROSCINI, *supra* note 8, at 206.
[132] Melzer, *supra* note 113, at 28.

highlights one of the shortcomings of the DPH framework: it lacks adequate protections for civilians from hostilities carried out by other civilians. There is a double standard when it comes to non-physical destruction or harm. Such harm would reach the threshold if it adversely affects military operations. However, if it only adversely affects civilian wellbeing, the answer is unclear.

Two cyber-attacks illustrate this point. First, the cyber-attacks on Estonia[133] in 2007 that caused the large-scale disruption of services in different areas, such as banking, administration, and media. Second, the cyber-attacks on Georgia[134] in the wake of the 2008 Russian Georgian War that caused even less non-physical harm outside of the defacement of certain websites, such as the Georgian President's website. Neither attack would reach the threshold of harm required by the DPH framework if it was established that they occurred in an armed conflict context, but carried out by civilians.[135]

Under this narrow interpretation, a cyber operation that targets a civilian power plant will only reach the threshold of harm if it adversely affects military operations *or* directly causes death, injury or (physical) destruction. However, under a broader interpretation even cyber operations that impose "mere harassment or inconvenience" to civilians would reach the threshold and qualify as hostilities.[136] The Interpretive Guidance itself provides that "the interruption of electricity, water, or food supplies . . . . would not, in the absence of adverse military effects, cause the kind and degree of harm required to qualify as direct participation in hostilities."[137] However, while the Interpretive Guidance threshold of harm will not be crossed in instances of civilian harm, it is possible that the Tallinn Manual interpretation will be applicable. This is the case if the *intent* of the person carrying out the cyber operation was to cause death, injury or destruction, but the cyber operation ended up only causing non-physical civilian harm.

---

[133] For in-depth analysis of the Estonia cyber-attacks, see Stephen Herzog, *Revisiting the Estonian Cyber Attacks: Digital Threats and Multinational Responses*, 4 J. STRATEGIC SECURITY, no. 2, 2011, at 49.

[134] *See* David Hollis, *Cyberwar Case Study: Georgia 2008*, SMALL WARS J., Jan. 6, 2011, http://smallwarsjournal.com/blog/journal/docs-temp/639-hollis.pdf.

[135] Turns, *supra* note 62, at 287.

[136] Melzer, *supra* note 113, at 29.

[137] INTERPRETIVE GUIDANCE, *supra* note 14, at 50.

However, many scholars believe the threshold of harm is "under-inclusive" when it comes to possible harms to civilians.[138] The result is that "[c]yber warriors are free to engage in cyber operations that could exact a significant toll on the civilian population of the enemy State without risk of being targeted, a consequence seemingly at odds with the goal of protecting civilians from the consequences of armed conflict."[139] Therefore, in order to meet the new threats posed by cyber operations, the DPH framework must recognize new potential harms, especially when it comes to civilians and civilian objects.

4. Cyber Espionage on Military Intelligence Targets

Cyber operations that aim to gather data and military intelligence present another difficult issue under the DPH framework. Cyber espionage operations cannot generally be considered to reach the threshold of harm, as simply collecting intelligence does not directly cause physical damage, bodily harm, or death, and it is unlikely on its own to affect military operations or military capacity. However, these operations could be considered as reaching the threshold of harm if they adversely affect the military operations or capacity of the targeted party. Such a scenario is not farfetched, as sensitive intelligence on upcoming military operations may well thwart and prevent a military operation, thus adversely affecting it.

The Interpretive Guidance adopts an approach that distinguishes two types of intelligence gathering. First, the Interpretive Guidance clarifies that "individuals whose function is limited to the purchasing, smuggling, manufacturing and maintaining of weapons and other equipment outside specific military operations or to the *collection of intelligence other than of a tactical nature*" would not be considered DPH.[140] This is a rational approach as these activities contribute to the general war efforts, and the lack of operational-tactical value to the intelligence gathering activity would not adversely affect the military operations of the adversary. Second, even if the intelligence gathered is of a tactical nature, the Interpretive Guidance requires that it constitute an "integral part of a concrete and coordinated tactical operation that directly causes such [above the threshold] harm."[141]

---

[138] *See* Schmitt, *supra* note 33, at 719.
[139] Padmanabhan, *supra* note 63, at 299.
[140] INTERPRETIVE GUIDANCE, *supra* note 14, at 34–35 (emphasis added).
[141] *Id.* at 54–55. The example that is provided by the INTERPRETIVE GUIDANCE on the matter is that "an unarmed civilian sitting in a restaurant using a radio or mobile phone to transmit tactical targeting intelligence to an attacking air force

Interestingly, the Tallinn Manual provides that "[c]yber espionage and other forms of information gathering directed at an adversary during an armed conflict do not violate the law of armed conflict."[142] However, the Manual later clarifies that in some instances, cyber espionage could constitute direct participation in hostilities, although no examples besides cyber espionage activities are given.[143]

Cyber espionage activities challenge both the threshold and temporal questions discussed *infra*. First, the impact of cyber operations that merely collect intelligence is not always visible or apparent, making it difficult to respond to with real-time targeting.[144] Second, while cyber espionage activities might be detected while they are still ongoing, they can also be detected after intelligence collecting has ceased and the individual in question has stopped participating in the activity.[145] To a lesser degree, similar challenges accompany cyber operations that actually produce physical effects that sometimes go undetected until well after the operations are completed. This was the case in Iran when Stuxnet infected its nuclear plants, causing massive damage to centrifuges.[146]

To overcome the challenges of cyber espionage carried out by civilians, contextual analysis is required. First, the focus needs to be on cyber operations collecting intelligence that is essential and integral to a military operation or the military capacity of an adversary. Second, the intelligence collected by the opposing party should be evaluated to determine whether it is objectively likely to reach the threshold of harm in a reasonable temporal proximity to the act of cyber espionage itself. Thus, not every intelligence collection operation would qualify as DPH, as it depends on whether such operation fulfills the threshold of harm requirement.

---

would probably have to be regarded as directly participating in hostilities." *Id.* at 81.

[142] TALLINN MANUAL, *supra* note 15, at 192 (rule 66(a)).

[143] TALLINN MANUAL, *supra* note 15, at 194 (rule 66 cmt. 4).

[144] *See* WILLIAM BOOTHBY, CONFLICT LAW: THE INFLUENCE OF NEW WEAPONS TECHNOLOGY, HUMAN RIGHTS AND EMERGING ACTORS 134 (2014).

[145] *Id.* (explaining that "if undertaken by a civilian, remote cyber information gathering and close access cyber espionage are likely to constitute direct participation in hostilities and would render the person concerned liable to attack *while so engaged*." (emphasis added)).

[146] David E. Sanger, *Obama Order Sped Up Wave of Cyberattacks Against Iran*, N.Y. TIMES (June 1, 2012), http://www.nytimes.com/2012/06/01/world/middleeast/obama-ordered-wave-of-cyberattacks-against-iran.html?.

B. The Temporal Challenge

The temporal idea behind the DPH framework is that civilians are targetable "*for such time* as they take direct part in hostilities."[147] This means that if they cease to participate in hostilities, they are no longer targetable with lethal force.[148] The Interpretive Guidance adds that the temporal scope includes "[m]easures preparatory to the execution of a specific act of direct participation in hostilities, as well as the deployment to and the return from the location of its execution. . . ."[149] When it comes to cyber operations, the equivalent of "preparatory measures" and "deployment and return" is sometimes nonexistent, as cyber operations can be carried out instantaneously without any preparation or travel to and from the location of the attack. Consequently, there are two questions in regards to the temporal aspect. First, at what point in time do the activities undertaken by a civilian qualify as direct participation in hostilities? Second, at what point in time and under what conditions does a civilian cease being targetable under the DPH framework? The answer to both of these questions is far from obvious.

In a meeting on "Direct Participation in Hostilities under International Humanitarian Law," Nils Melzer, ICRC's legal adviser, summarized the positions of current experts on the temporal scope issue:

> At one end of the spectrum were experts who preferred narrowly defining temporal scope and favoured strictly limiting loss of protection to the period where DPH is actually being carried out. At the other end were experts who said that, once a person had undertaken an act constituting DPH, that person must clearly express a will to definitively disengage and offer assurances that he or she will not resume hostilities in order to regain protection against direct attack. However, opinions varied greatly and could not easily be divided into two groups supporting distinct positions.[150]

In that meeting, opinions varied from a narrow scope of loss of immunity from an attack with the focus on the duration of the specific act

---

[147] Bill Boothby, *"And For Such Time As": The Time Dimension to Direct Participation in Hostilities*, 42 N.Y.U. J. INT'L L. & POL. 741, 742 (2010) (emphasis added).
[148] ROSCINI, *supra* note 8, at 209.
[149] INTERPRETIVE GUIDANCE, *supra* note 14, at 65.
[150] *See* Nils Melzer, Direct Participation in Hostilities Under International Humanitarian Law 34 (Oct. 25–26, 2004) (unpublished background paper), https://www.icrc.org/eng/assets/files/other/2004-03-background-doc-dph-icrc.pdf.

in question, to a broader scope that legitimizes targeting of civilians as long as the armed conflict takes place.[151]

The Interpretive Guidance adopted a somewhat more ambiguous and flexible scope by providing that "civilians lose protection against direct attack *for the duration* of *each specific act* amounting to direct participation in hostilities . . . ."[152]

1. Beginning and Cessation of DPH Status

The Interpretive Guidance clarifies that "[m]easures preparatory to the execution of a specific act of direct participation in hostilities, as well as the deployment to and the return from the location of its execution, constitute an integral part of that act."[153] In addition, the Interpretive Guidance addresses cyber operations specifically, by providing that –

> Where the execution of a hostile act does not require geographic displacement, as may be the case with computer network attacks or remote-controlled weapons systems, the duration of direct participation in hostilities will be restricted to *the immediate execution of the act and preparatory measures forming an integral part of that attack*.[154]

The Tallinn Manual took a broader stance by including "actions immediately preceding or subsequent to the qualifying act. For instance, traveling to and from the location where a computer used to mount an operation is based…" in the temporal scope of the DPH framework.[155]

The Law of War, on the other hand, is not decisive on the temporal scope issue and focuses on the cessation rather than the beginning of the act as constituting DPH, by providing that:

> In the U.S. approach, civilians who have taken a direct part in hostilities must not be made the object of attack *after* they have permanently ceased their participation because there would be no military necessity for attacking them. Persons who take a direct part in hostilities, however, do not benefit from a "revolving door" of protection. There may be difficult cases not clearly falling into either

---

[151] *Id.* at 34–35.
[152] INTERPRETIVE GUIDANCE, *supra* note 14, at 70 (emphasis added).
[153] *Id.* at 65.
[154] *Id.* at 68 (emphasis added).
[155] TALLINN MANUAL, *supra* note 15, at 121 (rule 35 cmt. 7).

> of these categories, and in such situations a case-by-case analysis of the specific facts would be needed.[156]

The temporal aspect is exacerbated in cyberspace due to the instantaneous nature of cyber operations through the initiation, effects materialization, and termination of the cyber operation.[157] In addition to these difficulties, the effects of some cyber operations are only realized long after the perpetrator regained his or her civilian status.[158] Some argue that the narrow opportunity to respond to cyber operations carried out by civilians in an armed conflict is too restrictive and would *de facto* eliminate the right to strike back.[159] However, it is important to note that the DPH framework is intended to provide a tool to stop ongoing hostile acts, rather than punish the perpetrators *ex post facto*.[160] The perpetrator may still face criminal prosecution as a consequence of his or her violations of international or domestic criminal law.[161] The main difficultly remains with the prolonged effects of cyber operations, such as "logic bombs."[162] In this case, a major gap lies between civilians who cease carrying out cyber operations and the effects of these cyber operations still experienced by the victims. However, it is important to distinguish between two continued effects of cyber operations. First, there are cyber operations in which the attack code is actively running, causing continued effects.[163] In this case, the individual carrying out the cyber

---

[156] LAW OF WAR MANUAL, *supra* note 17, at 230 (emphasis added).
[157] *See* Schmitt, *supra* note 121.
[158] Delerue, *supra* note 114, at 11.
[159] Schmitt, *supra* note 121, at 102 ("The restrictive interpretation of the for such time criterion would suggest that the direct participant can only be attacked while actually launching the operation. This is problematic in that many cyber operations last mere minutes, perhaps only seconds. Such a requirement would effectively extinguish the right to strike at direct participants.").
[160] Delerue, *supra* note 114, at 11.
[161] THE HANDBOOK OF INTERNATIONAL HUMANITARIAN LAW, *supra* note 41, at 255–57.
[162] *See* Stephen Northcutt, *Logic Bombs, Trojan Horses, and Trap Doors*, SANS TECH. INST.: SECURITY LABORATORY: METHODS ATTACK SERIES (May 2, 2007), http://www.sans.edu/research/security-laboratory/article/log-bmb-trp-door ("Logic bombs are small programs or sections of a program triggered by some event such as a certain date or time, a certain percentage of disk space filled, the removal of a file, and so on. For example, a programmer could establish a logic bomb to delete critical sections of code if she is terminated from the company. Logic bombs are most commonly installed by insiders with access to the system.").
[163] Dinniss, *supra* note 4, at 276.

operation is targetable during the time in which the code is running, as long as the DPH requirements are fulfilled.[164] Second, there are cyber operations that cause continued effects. Those effects are not due to the ongoing cyber operation but instead to secondary and tertiary effects on the target's computer systems and networks. In this situation, the individual would be targetable only during the duration of the cyber operation.[165]

2. Concept of "Revolving Door"

While civilians who are members of non-state organized armed groups (OAGs) and assume a continuous combat function are targetable as long as they are members of that group,[166] civilians who are not members of OAGs are generally targetable at the time their hostile act takes place. Civilians who directly participate in hostilities sometimes engage in repeated hostile acts on different occasions. Such action gave rise to the concept of a "revolving door,"[167] referring to their statuses constantly changing between civilians and civilians who utilize DPH.[168]

While the Interpretive Guidance limits the application of the DPH framework to each hostile act,[169] there was major disagreement among experts as the concept of "revolving door" in the Tallinn Manual.[170] The Tallinn Manual posited that when an individual mounts repeated cyber operations, it is unclear (the experts were split) whether each act should be evaluated separately. Therefore, the duration of targetability is limited to

---

[164] *Id.*
[165] *Id.*
[166] INTERPRETIVE GUIDANCE, *supra* note 14, at 71.
[167] *See* W. Hays Parks, *Air War and the Law of War*, 32 A.F. L. REV. 1, 118 (1990).
[168] *See* HCJ 769/02 Pub. Comm. Against Torture in Isr. v. Gov't of Israel 62(1) PD 507 (2006) (Isr.), http://elyon1.court.gov.il/files_eng/02/690/007/A34/02007690.a34.pdf, ¶ 40 ("These examples point out the dilemma which the 'for such time' requirement presents before us. On the one hand, a civilian who took a direct part in hostilities once, or sporadically, but detached himself from them (entirely, or for a long period) is not to be harmed. On the other hand, the 'revolving door' phenomenon, by which each terrorist has 'horns of the alter' to grasp or a 'city of refuge' to flee to, to which he turns in order to rest and prepare while they grant him immunity from attack, is to be avoided. In the wide area between those two possibilities, one finds the 'gray' cases, about which customary international law has not yet crystallized. There is thus no escaping examination of each and every case.").
[169] INTERPRETIVE GUIDANCE, *supra* note 14, at 44–45.
[170] TALLINN MANUAL, *supra* note 15, at 121–22.

that act, or whether targetability "continues throughout the period of intermittent activity."[171]

Conversely, the Law of War Manual is quite explicit on the issue, stating that "[t]he law of war, as applied by the United States, gives no 'revolving door' protection; that is, the off-and-on protection in a case where a civilian repeatedly forfeits and regains his or her protection from being made the object of attack depending on whether or not the person is taking a direct part in hostilities at that exact time."[172]

Realizing that most cyber-attacks today are carried out by organized groups is key to the assessment of the targetability of these individuals in an armed conflict context. For example, in April 2007, Estonia decided to relocate a memorial commemorating the Soviet liberation of Estonia from the Nazis from a central location in its capital to a less central area, which set off an intense period of riots and protests among the Estonian Russian-speaking community.[173] In the aftermath, Estonia was attacked by a massive wave of cyber-attacks that targeted mostly its commercial banking and media websites.[174] Although the cyber-attack did not occur in an armed conflict context, the cyber-attacks were led by an organized group called "Nashi" (Russian for "Ours"), which in the past conducted operations on Moscow's behalf.[175] This demonstrates that a massive cyber-attack, such as the Estonian attack, is often carried out by an organized group playing a repeat role in similar operations. These repeated acts should then be analyzed cohesively, transcending the evaluation of individual sporadic acts.

(a) Characterization as Organized Armed Group?

The gap within IHL with regard to OAGs is apparent when applied to hacking groups. The broader question is whether a cyber operation is carried out by an individual civilian, whose DPH status will be evaluated individually, by civilians as part of a hacking group, who will be evaluated as members of organized armed groups, or by civilians with continuous combat function, who are targetable at all times due to their affiliation. More specifically, two issues arise from this concept. First, whether

---

[171] *Id.* at 122.
[172] LAW OF WAR MANUAL, *supra* note 17, at 231.
[173] Herzog, *supra* note 128, at 50–51.
[174] *Estonia Hit by 'Moscow Cyber War'*, BBC NEWS, http://news.bbc.co.uk/2/hi/europe/6665145.stm (last updated May 17, 2007).
[175] Noah Shachtman, *Kremlin Kids: We Launched the Estonian Cyber War*, WIRED (Mar. 11, 2009, 12:45 PM), http://www.wired.com/2009/03/pro-kremlin-gro/.

"virtually" organized groups are "organized" and qualify as OAGs.[176] Second, whether hacking groups, who reach the threshold of harm, are "armed."[177]

The Commentary to Additional Protocol I provides that "the term 'organized' is obviously rather flexible, as there are a large number of degrees of organization."[178] It further provides that the fighting character of a group is "collective," meaning that it is "conducted under proper control and according to rules," rather than "individuals operating in isolation with no corresponding preparation or training."[179] In that sense, hacking groups can be acting under a direct command structure and carrying out their cyber operations collectively by sharing intelligence and hacking tools and by identifying exploitable vulnerabilities.[180] Hacking groups that do not have a clear command structure but operate for a common cause are more difficult to analyze and should be evaluated on a case-by-case basis. However, the assumption is that States will still deem them to be organized armed groups.[181]

The requirement of "armed" is closely related to a collective of individuals who can carry out "attacks" under IHL, meaning "acts of violence against the adversary, whether in offence or in defense."[182] The Tallinn Manual defines cyber attacks as "cyber operation[s], whether offensive or defensive, that [are] reasonably expected to cause injury or death to persons or damage or destruction to objects."[183] The main issue with the notion of "armed" is that not all hacking groups are causing the required threshold of harm to be considered "armed" organized groups. However, there is a clear distinction between harm in the context of "armed," which necessarily requires violent acts, and harm in the context of DPH, which is a more lenient standard that can be satisfied by adverse military effects, even in the absence of violent acts. Given that gap, individuals who are members of hacking groups might be considered DPH. However, these hacking groups will not be considered "organized armed groups" since the threshold for "armed" was not reached. That disparity challenges targeting decisions since each individual is evaluated in isolation from the whole group in question. In this regard, the DPH

---

[176] Schmitt, *supra* note 121, at 98.
[177] *Id.*
[178] ICRC COMMENTARY ON APS, *supra* note 18, at 512 (citations omitted).
[179] *Id.*
[180] Schmitt, *supra* note 121, at 98.
[181] *Id.* at 99.
[182] AP I, *supra* note 1, art. 49(1).
[183] TALLINN MANUAL, *supra* note 15, at 106 (rule 30).

framework is actually more permissive, and as the Israeli Supreme Court put it in the *Targeted Killings* case, "it is possible to take part in hostilities without using weapons at all."[184]

Another major flaw in the context of OAGs is that the OAG is required to "belong to a party to the conflict."[185] This requirement might be fulfilled if the OAG is engaged in an armed conflict, whether international or non-international. One example is the Islamic State. It is engaged in an armed conflict against the United States while also carrying out concurrent cyber operations in relation to that armed conflict.[186] However, hacking groups that solely employ cyber operations are unlikely to trigger a separate armed conflict,[187] as the threshold for such armed conflict would not be reached solely based on cyber operations.[188] As the ICRC stated, "whether CNA [computer network attacks] alone will ever be seen as amounting to an armed conflict will probably be determined in

---

[184] HCJ 769/02 Pub. Comm. Against Torture in Isr. v. Gov't of Israel 62(1) PD 507 (2006) (Isr.), http://elyon1.court.gov.il/files_eng/02/690/007/A34/02007690.a34.pdf, ¶ 33.

[185] INTERPRETIVE GUIDANCE, *supra* note 14, at 23.

[186] *See, e.g.*, Michael Safi, *ISIS 'Hacking Division' Releases Details of 1,400 Americans and Urges Attacks*, THE GUARDIAN (Aug. 12, 2015, 23:55 EDT), http://www.theguardian.com/world/2015/aug/13/isis-hacking-division-releases-details-of-1400-americans-and-urges-attacks.

[187] "International armed conflict" is defined in the COMMENTARY TO THE GENEVA CONVENTIONS as "[a]ny difference arising between two States and leading to the intervention of armed forces is an armed conflict . . . . It makes no difference how long the conflict lasts, or how much slaughter takes place." INT'L COMM. OF THE RED CROSS, I COMMENTARY ON THE GENEVA CONVENTION FOR THE AMELIORATION OF THE CONDITION OF THE WOUNDED AND SICK IN ARMED FORCES IN THE FIELD (Jean S. Pictet ed. 1952).

[188] *See What Limits Does the Law of War Impose on Cyber Attacks?*, INT'L COMM. OF THE RED CROSS (June 28, 2013), https://www.icrc.org/eng/resources/documents/faq/130628-cyber-warfare-q-and-a-eng.htm ("Most cyber operations are not linked to an armed conflict, so IHL does not even apply. Even in armed conflict, most hackers would be civilians who remain protected by IHL against direct attack – although they would remain subject to law enforcement and possible criminal prosecution depending on whether their activities violated other bodies of law."); *see also* Asim Rizvanovic, *Cyber-Attacks Will Not Result in Armed Conflicts in the 21st Century*, E-INT'L. REL. STUDENTS (May 7, 2013), http://www.e-ir.info/2013/05/07/cyber-attacks-will-result-in-armed-conflicts-in-the-21st-century/.

a definite manner only through future state practice."[189] In that regard, the Tallinn Manual clarifies that a non-international armed conflict using cyber terms "exists whenever there is *protracted armed violence*, which may *include* or *be limited* to cyber operations."[190] This introduces the possibility that cyber operations alone could qualify as non-international armed conflict if the operations reached "a minimum level of intensity" and a "minimum degree of organization."[191] This restatement represents a widely accepted definition, as stipulated by the International Criminal Tribunal in the Former Yugoslavia during the *Tadic* judgment.[192]

The Interpretive Guidance provides that such groups, without belonging to a party to the armed conflict, "cannot be regarded as members of the armed forces of a party to the conflict" and are therefore considered civilians.[193] But the Guidance further provides that such groups "could still be regarded as parties to a separate non-international armed conflict provided that [their] violence reach[ed] the required threshold."[194] Consequently, unaffiliated hacking groups pose a great challenge to the notion of OAGs given that they might be acting on their own behalf, isolated from the armed conflict politics, and not reaching the threshold of armed conflict.[195] However, as noted before, the DPH paradigm might still apply to these groups in relation to their cyber operation in an ongoing armed conflict, even if such groups are not parties to the conflict. This is because the DPH framework applies to civilians who directly participate in hostilities, rather than existing parties to armed conflict.[196]

---

[189] Knut Dörmann, *Applicability of the Additional Protocols to Computer Network Attacks*, INT'L COMM. OF THE RED CROSS (Nov. 19, 2004) https://www.icrc.org/eng/assets/files/other/applicabilityofihltocna.pdf.

[190] TALLINN MANUAL, *supra* note 15, at 84 (rule 23) (emphasis added).

[191] *Id.*

[192] Prosecutor v. Tadić, Case No. IT-94-1-I, Decision on Defence Motion for Interlocutory Appeal on Jurisdiction, ¶ 70 (Int'l Crim. Trib. for the Former Yugoslavia Oct. 2, 1995), http://www.icty.org/x/cases/tadic/acdec/en/51002.htm ("[A]n armed conflict exists whenever there is a resort to armed force between States or protracted armed violence between governmental authorities and organized armed groups or between such groups within a State.").

[193] INTERPRETIVE GUIDANCE, *supra* note 14, at 23.

[194] *Id.* at 24.

[195] *See* Schmitt, *supra* note 121, at 100 ("[T]he criterion would also exclude patriotic hacker groups unaffiliated with one of the belligerent parties, even if conducting cyber attacks for its benefit, because the group's activities would lack the 'agreement' of that party and its actions would in no other way be attributable to the party under the law of State responsibility.").

[196] d armed groups or between such groups within a State.").

Overall, when a civilian engages in cyber operations sporadically and without a continuous pattern, he or she shall be protected after the cyber operation has ceased. However, evidence that points towards membership in hacking groups could cause a civilian to lose protection. In the rare cases where a hacking group is considered an OAG, members who carry out cyber operations on behalf of that OAG would not be protected after the cyber operation has ceased unless they permanently become unaffiliated with that OAG.

CONCLUSION

The emergence of cyber operations as an integral part of modern armed conflicts introduces a myriad of challenges for legal experts and policymakers given the conspicuous absence of consistent state practices and the dearth of substantive norms needed to reform and ultimately govern cyberspace conduct. However, there are some key considerations that might assist in the norm formation process in relation to cyber operations in armed conflicts.

First, the threshold of harm needs to be refined to encompass new types of harms, which would trigger the forfeiture of civilian status. It must encompass effects that, although inconvenient, will not be considered direct participation in hostilities. For example, terrorizing civilians through cyber operations, alteration of critical civilian data, and incapacitating civilian services through cyber operations should be considered sufficient harm to trigger the DPH framework. However, effects and activities such as propaganda, commercial and non-tactical espionage, free expression relating to the armed conflict, and online advocacy, even if highly inconvenient, should not be considered DPH. The applicability of the DPH framework to a civilian does not necessarily mean that he will be targeted, but that the DPH framework will, in a way, assist in creating binding norms for behavior in cyberspace, particularly pertaining to civilians, in the context of an armed conflict.

Second, the temporal aspect of the DPH framework is another major issue. Cyber operations might be ongoing, in which case the perpetrators are targetable. However, operations are more commonly instantaneous and distant. The DPH framework was intended mainly to address "hot battlefield" issues, such as civilians physically participating in hostilities with geographical and temporal proximity to the battlefield.

---

[196] INTERPRETIVE GUIDANCE, *supra* note 14, at 26 ("All persons who are neither members of the armed forces of a party to the conflict nor participants in a levée en masse are civilians and, therefore, entitled to protection against direct attack unless and for such time as they take a direct part in hostilities.").

We are witnessing the civilianization of the cyber-battlefield,[197] that is, cyberspace is becoming a "civilian-occupied battlespace."[198] States, on the other hand, find it increasingly challenging to retain their monopoly over cyber-force, given the ease of use and accessibility of computer systems for civilians.[199] Two phenomena accompany that assertion. First, civilians will become more involved in cyber operations during armed conflicts. Second, these civilians will be increasingly classified based on their status, either as civilians with continuous combat functions or contractors who participate in combat.[200]

At this point, only State practice, actual materialization of these predictions, and more serious and nuanced harms due to cyber operations can point towards the expected changes in the temporal concept as it relates to the DPH framework. The Interpretive Guidance, Tallinn Manual, and U.S. Law of War Manual's uses of ambiguous and broad terms and concepts to address cyber warfare are inadequate, and reveal that even the most advanced instruments and organizations cannot solve these threats in one fell swoop. Thus, more thought and normative development, especially for international rules of conduct in cyberspace, are desperately needed.

---

[197] *See* Andreas Wenger & Simon J.A. Mason, *The Civilianization of Armed Conflict: Trends and Implications*, INT'L COMM. OF THE RED CROSS (Dec. 31, 2008), https://www.icrc.org/eng/assets/files/other/irrc-872-wenger-mason.pdf.

[198] Karine Bannelier-Christakis, *Is The Principle of Distinction Still Relevant in Cyberwarfare?*, *in* RESEARCH HANDBOOK ON INTERNATIONAL LAW AND CYBERSPACE 343, 358 (Tsagourias & Buchan eds., 2015).

[199] *See generally* Susan W. Brenner with Leo L. Clarke, *Civilians in Cyberwarfare: Conscripts*, 43 VAND. J. TRANSNAT'L L. 1011 (2010).

[200] E.L. Gaston, Note, *Mercenarism 2.0? The Rise of the Modern Private Security Industry and Its Implications for International Humanitarian Law Enforcement*, 49 HARV. INT'L. L.J. 221, 225 (2008) (explaining that "private military firms . . . offer combat capabilities, tactical analysis, and other direct military support").