

Legal status of civilian participants in cyber operations during armed conflict

Valeriia Filinovych^{1,*†}, Assel Mukasheva^{2,†}

¹Scientific Cyber Security Association of Ukraine, Mykhaila Dontsia Str., 2A, Kyiv, 03161, Ukraine

²Kazakh-British Technical University, Tole Bi Str., 59, Almaty, 050000, Kazakhstan

Abstract

The growing involvement of civilians in cyber operations during armed conflicts raises fundamental questions regarding their legal status under International Humanitarian Law (IHL). Unlike traditional warfare, cyber operations can be conducted remotely and anonymously, outside conventional battlefields, making the distinction between civilians and combatants increasingly difficult. This paper examines the legal status of civilian participants in cyber operations through the prism of IHL, with particular attention to the concept of Direct Participation in Hostilities (DPH). The study analyzes the three constitutive elements of DPH identified by the International Committee of the Red Cross: threshold of harm, direct causation, and belligerent nexus, and evaluates their applicability in cyberspace. Special attention is paid to the challenges posed by the temporal dimension of cyber activities, including the revolving-door principle and the concept of Continuous Combat Function (CCF). The article further explores relevant provisions of the Tallinn Manual 2.0, including Rules 93, 94, 95, 97, and 102, which address distinction, civilian protection, and participation in cyber hostilities. The research concludes that existing IHL norms formally extend to cyberspace but face significant practical limitations. The anonymity of cyber actors, the dual-use nature of digital infrastructure, and the difficulty of identifying the beginning and end of participation in cyber operations underscore the need to further develop international legal regulation and clarify DPH criteria in the cyber domain.

Keywords

cyber combatant, cyber attack, International Humanitarian Law, cybersecurity law, cyber warfare, civilians, direct participation in hostilities, Tallinn Manual

1. Introduction

The world has changed, as have styles of warfare. States have moved away from conventional means – that is, using traditional weapons such as small arms and cruise missiles, excluding biological, chemical, or nuclear weapons, while observing International Humanitarian Law (IHL) [1]. Historically, confrontations took place across three domains: land, sea, and air. But the world has changed.

Today, hybrid (asymmetric, unconventional [2]) warfare is encountered daily. A key element is cyberwarfare – confrontation in cyberspace, which encompasses not only computers but also the resilience of digital infrastructure, networks, and global positioning systems. Attacks on this infrastructure can lead to physical consequences such as power outages or transport disruptions. Military and civilian cyber infrastructure increasingly overlap, so civilians often suffer from cyberattacks during armed conflict. The applicability of international humanitarian law to cyber operations during armed conflicts has also been emphasized by the International Committee of the Red Cross [3].

Currently, cyber attacks accompany almost every military operation. For example, on February 23, 2022, the eve of Russia's full-scale invasion of Ukraine, several state and banking institutions experienced DDoS attacks [4]; prior to that, there were two attack waves in January.

Thus, cyberspace is not a battlefield in the classical sense, but it is quite vulnerable due to the possibility of anonymous participation, remote access, and other factors. And, most importantly, it is

CMiGIN'26: Fifth International Workshop on Conflict Management in Global Information Networks, June 26–27, 2026, Krakow, Poland

*Corresponding author.

† These authors contributed equally.

✉ vvfilinovych@gmail.com (V. Filinovych); a.mukasheva@gmail.com (A. Mukasheva)

ORCID 0000-0001-8824-615X (V. Filinovych); 0000-0001-9890-4910 (A. Mukasheva)



© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

often not the classic military or representatives of state bodies who act here, but civilians. Therefore, the issue of the direct participation of the civilian population in hostilities in the context of cyber operations, especially under IHL, is of paramount importance.

In science, due to the frequent participation of civilians in cyber operations, the doctrinal concept of a “cyber combatant” is gradually emerging. However, this term is not officially defined and, under current IHL, a civilian’s participation in cybermilitary activities does not automatically render him a combatant.

Therefore, it is worth examining existing approaches in international humanitarian law to determine the legal status of civilians participating in cyber operations. Most often, in addition to specially authorized bodies, these include the so-called hacktivists, volunteers (individuals), and specialized civilian groups. Hacktivists, usually the so-called “gray hackers”, mostly act for ideological reasons, in particular to oppose the ruling ideology or a political campaign, or for entertainment. That is, the goal pursued by such entities is not negative; nevertheless, they have not previously received official access to certain data, systems, and networks and act accordingly, “semi-legally” [5]. Volunteers and specialized groups of them are not military; they are voluntary participants in IT roles who perform specific tasks in cyber operations. Such specialists and groups have become particularly widespread in Ukraine since 2022. According to the author, the most famous of them is the IT Army of Ukraine.

And the problem that arose in connection with the activities of these entities is to establish their official status from the point of view of, in particular, international law. After all, on the one hand, they are civilians, and on the other hand, they can directly influence the course of military operations. This scientific study is devoted to solving the problem stated.

2. Summary of the primary material

Consequently, the main objective of this study is to provide clarity in assessing the actions of civilians in military cyber operations, where the classic categories of IHL are insufficient for unconditional application.

In subparagraphs 4–10, paragraph 1, section II of the Instruction on the procedure for implementing the norms of international humanitarian law in the Armed Forces of Ukraine, a clear list of IHL principles is contained, according to which responsible persons must act during military activity. Among other things, these are the following:

- the principle of legality requires subjects involved in conducting military operations to strictly follow the norms of IHL;
- the principle of proportionality prohibits attacks that may cause incidental harm to the civilian population or objects; and if such harm will be excessive in comparison with the expected military advantage;
- the principle of limitation declares the limited choice of means and methods of warfare as the right of the opposing parties;
- the principle of distinction requires that civilians and objects be clearly distinguished from combatants and military objects in order to ensure the protection of the former;
- the principle of military necessity emphasizes the use of force in the form and volume necessary to achieve a legitimate military objective, as well as the non-use of force when prohibited by the norms of IHL. At the same time, the commander must make maximum efforts to minimize unnecessary destruction and incidental losses [6].

Separately, it is necessary to dwell on the principle of distinction. Nowadays, many civilians directly participate in the fight in various ways. If we delve into the issue of cyberwarfare, a person with an ordinary laptop, access to the Internet, a VPN enabled, and specific IT knowledge can perform certain actions, and now the turning point in special operations has been reached. That is, an individual is not a military person but participates in a cyberattack or similar military actions. And here a logical

Table 1
Illustrative Examples of the Constitutive Elements of DPH

	Threshold of harm	Direct causation	Belligerent nexus
Presence of an element	DDoS attack on the servers of the Ministry of Defense during an operation (there is a direct impact on military coordination)	Transmission of coordinates for a cyberattack that was successfully carried out (there is an action and immediate concrete damage)	Hacking the enemy's logistics system to help one's own state (the action is directly related to the confrontation)
Absence of an element	Creating anti-war memes (this is entertaining content that does not pose any real harm or threat)	Selling laptops to army (no direct connection to the harm caused)	Breaking into an ATM and withdrawing money in wartime (the goal is personal gain, not support for the warring party)

question arises: who will such a person be from the point of view of IHL, and will he still be considered a civilian?

If we turn to the Additional Protocols to the Geneva Conventions of 12 August 1949 I and II, then in Articles 51(1,3) and 13(3), respectively, it is indicated, in particular, that protection against dangers caused by military operations extends to civilians if they are not directly involved in hostilities [7]. But what exactly will be considered “direct participation in hostilities” (DPH)?

The answer to this question is given in the Interpretive Guidance adopted by the International Committee of the Red Cross (ICRC) in 2009. The document identifies 3 criteria of DPH, namely:

- threshold of harm. The action either actually harms the military capabilities of the enemy or creates a risk of lethality, injury, or destruction;
- direct causation. Between what the individual did and the harm caused, there must be a direct causal connection, not distant in time and space;
- belligerent nexus. The specified action is particularly aimed at supporting one of the parties, and to the detriment of the other [8].

It is important that all 3 specified criteria exist simultaneously; if any are missing, this will no longer be considered DPH. Table 1 below provides specific examples of what will and will not be considered DPH.

For greater clarity, the cumulative nature of the three constitutive elements of DPH is illustrated by the following decision-making model (Figure 1).

In his work, F. Delerue interprets the above-mentioned Interpretive Guidance and draws attention to the very complex problem of DPH in cyberspace – when exactly does a civilian lose protection and for what period of time. Thus, he notes that not only the attack itself, but also the preparation for it can be taken into account as DPH, but precisely “preparation for a specific hostile act”, and not just the preparation of the general ability to commit unspecified hostile acts. Hence, the creation of cyber weapons is usually done for a specific purpose; therefore, it will be referred to as DPH. Deployment and return from execution may also be understood as direct participation. The scientist also emphasizes that geographical distance does not affect the corresponding qualification. As for the time during which civilians participate in a specific military activity, a person loses protection during each such act of DPH [9].

We can conclude that when an IT student creates malware for a specific attack on the Ministry of Defense website, i.e., there is a specific target, then he is already directly participating in a military confrontation. But if he, as a blogger, shares universal hacking tools with subscribers, which, if desired, can be used for a similar action from the example, then we cannot talk about DPH.

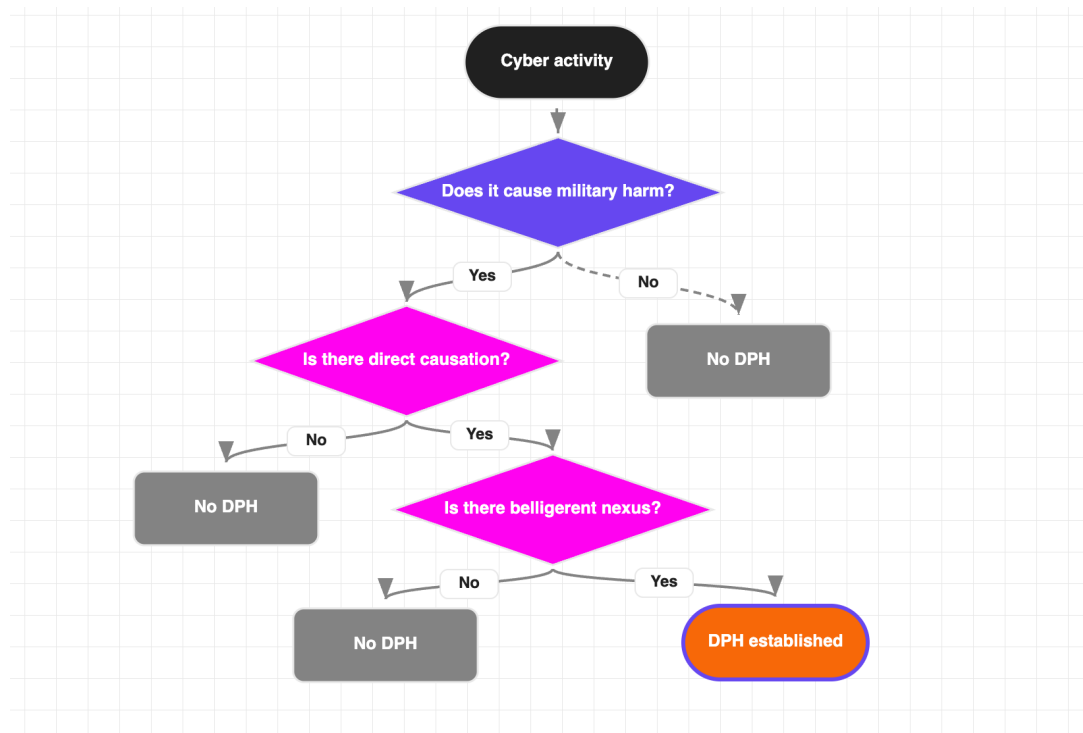


Figure 1: Decision tree for establishing DPH in cyber operations.

Regarding the physical distance from the place of hostilities, it should be noted that in a conventional war, the battlefield itself is important, but in cyberwar, a person can simply sit at home or in another country and commit controversial actions on his own laptop. But DPH can be qualified here.

There is also a nuance with the principle of revolving door, because in cyberspace an attack can last a fairly short time, for example, minutes, as well as the intervals between attacks. Also, some specific actions, for example, launching malware, can be detected much later than when the software was launched, and at the time of detection, the attacker is already under protection again, because at that moment, he is not taking direct action.

Separately, there is a need to focus on the Continuous Combat Function. Interpretive Guidance [8] directly states that the protection of civilians is based on a revolving door, while members of organized armed groups of a non-state party to the conflict lose the status of civilians, and with it the protection from direct attack for the entire time they are performing their continuous combat function. Accordingly, the protection of individuals in such a group is not restored after each separate act of participation in hostilities, but depends on the preservation or termination of their continuous combat function (Figure 2).

Having highlighted the main provisions of the Interpretive Guidance, it is necessary to study the Tallinn Manual, which, although not an international legal act, is a doctrinal source of law. This document is an independent research product, the creation of which was attended by approximately 20 recognized experts on the issues considered in this article.

The full name of the document is the Tallinn Manual on the International Law Applicable to Cyber Warfare; its development lasted about three years, and the official publication took place in 2013 at Chatham House. In response to the constant growth in the number and quality of cyber threats, the Tallinn 2.0 project was developed in 2017, significantly expanding on the previous version. It is worth studying it in more detail.

The essence of the Tallinn Manual is to show how IHL should be applied in the event of a cyber conflict. The document consists of 4 parts, which contain 20 chapters and 154 Rules. For the purposes of this article, it is important to familiarize yourself with the following Rules.

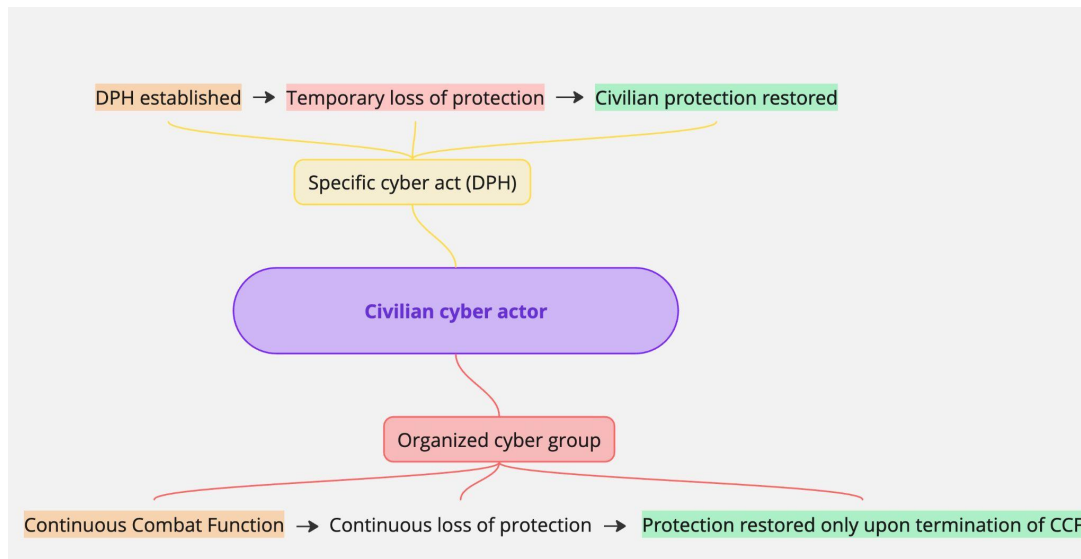


Figure 2: Distinction between DPH and CCF in cyber operations.

- Rule 93. It is devoted to Distinction. It is declared as a steady foundation that, for warring countries, the only common goal should be to weaken the enemy's forces and prevent unnecessary suffering. The importance of distinguishing civilians from combatants is again emphasized. For example, a situation is presented in which a cyberattack is directed at the residence, a civilian facility, of a civilian leader, with the aim of forcing his surrender. In this case, even if the goal is achieved, the attack will be illegal, as it does not comply with the law of war and peace [10]. The essence of the Rule is that not every IT worker is a permissible target for attack; and even more – in case of any doubt, such a person is a civilian. That is, the principle of distinction is increasingly difficult to implement in the cyber environment because infrastructure is usually dual-use, and the distinction between civilian and combatant is complicated by the former's involvement in cyber support.
- Rule 94. It prohibits cyberattacks against persons to whom Rule 91 refers those who are civilians, while an operation in cyberspace must be qualified as an attack in order for Rule 94 to apply. This rule also emphasizes once again that civilians lose protection for the time they are directly involved in hostilities. The damage caused by the attack must correspond to the level established in Rule 92. As an example, a case of a cyberattack involving the manipulation of a person's medical data in a hospital database is presented. Such a person will be considered the object of the attack, but the base itself – not, unless the achievement of the appropriate level of damage is proven [10]. The problem with implementing this Rule is that cyber operations are usually decentralized, anonymous, and continuous, making it difficult to determine the specific point in time when a civilian has been removed from and re-obtained protection.
- Rule 95. It is declared that if there is doubt about the status of a person, he or she is considered a civilian by default. But there is no specific definition of the threshold at which doubt is sufficient. It is emphasized that the use of a computer or a specific network is not conclusive evidence of military status. It is also quite difficult to physically determine who exactly is committing a cyber action [10]. The problem with the Rule in the context of cyberwar is the difficulty of proving the duration of the cyber action, its continuity, and whether the person acted independently or as part of an organized group. And there is still the problem of anonymity, decentralization, and the technical complexity of cyberspace. Nevertheless, it is this Rule that prevents the list of legitimate targets in cyberwar from expanding.
- Rule 102. This provision is very similar to Rule 95 and requires that, in the event of doubt as to whether a certain object is used for military purposes, it should be considered civilian. And to consider it a military one, a reliable assessment must be carried out; in case of doubt, the relevant

persons must request additional information. Objects “normally intended for civilian purposes”, i.e., which are not used regularly or significantly for military purposes, include, in particular, civilian social networks and Internet services. Importantly, if an object has been converted to cyber infrastructure, and then again to civilian, to be used for military purposes in the future, then it is a military target [10]. In essence, Rules 95 and 102 are safety nets because cyber operations often traverse civilian networks, commercial platforms, and the like. The principle of distinction in cyberspace is difficult to apply because most cyber infrastructure is dual-use.

Rule 97 should be discussed individually. If a civilian is DPH in a cyber operation, he or she may be lawfully attacked, and the harm caused to him or her will not be taken into account when assessing the proportionality of the attack. It is explicitly emphasized that the responsible person must act, i.e., not an ordinary citizen whose laptop was used in a cyberattack in secret. Although the gadget itself in this situation will be a legitimate military target if it meets the criteria of Rule 100. This provision also emphasizes the importance of the presence of the three previously defined constitutive elements of DPH, while taking into account the specifics of the cyber process. Therefore, with regard to the threshold of harm, there is no requirement for physical harm; instead, there must be a negative impact on the enemy’s military power, e.g., disruption of the enemy’s command and control network following a cyberattack. Direct causation is also present in the above example, because the harm arose precisely from the cyberattack. Belligerent nexus is manifested by the fact that the attacked system was used by the enemy for military operations. Experts have agreed that, for example, developing a computer program to exploit vulnerabilities in enemy systems or collecting data on rival cyber activities and subsequently transferring them to the domestic army is an act of direct participation. At the same time, developing such software and placing it on the World Wide Web without a specific purpose, even if someone later uses it for military purposes, is no longer an act of direct participation. Doubts about direct participation arise when the developed software will logically be used in an attack, but the developer does not know the exact purpose of the attack [10].

Nevertheless, any act of direct participation of a person in hostilities turns him from a civilian into a legitimate target until he takes the appropriate action. In particular, in the context of cyber confrontation, this period may begin with assessing the system’s vulnerabilities and continue through the evaluation of the need for a second attack. The issue of delayed consequences was also raised, for example, the installation of a logic bomb that will be activated at a future time, which can then be changed by command. Therefore, experts concluded that a person’s direct participation begins with the start of planning the operation and ends with the termination of their active role in it. That is, according to the given example, the beginning of DPH is the moment when this person plans to install a logic bomb, and the end is its activation at this person’s command. Regarding repeated cyberattacks, experts’ opinions differ: some, supporting Interpretive Guidance, emphasize analyzing each individual activity, while others consider the DPH’s beginning to be the first cyberoperation, followed by subsequent ones [10]. Figure 3 demonstrates the instability of the temporal boundaries of direct participation in hostilities in cyberspace.

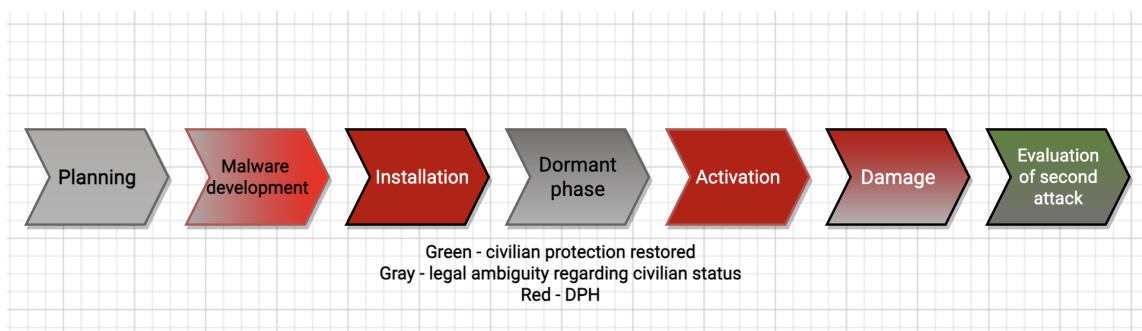


Figure 3: Temporal dimension of direct participation in hostilities in cyberspace

It is worth noting that a civilian who participates in a cyberoperation does not automatically become a combatant. This is still a civilian who becomes a legitimate target only temporarily. Rule 97 is well applied to conventional warfare, but in cyberspace, the traditional approach to IHL is difficult to apply because an individual's participation in an operation may be continuous from a technical perspective, even though it is operationally intermittent. There is a destabilization of the temporal logic of direct participation in hostilities.

In 2021, the development of the Tallinn Manual 3.0 was initiated, designed for 5 years, as version 2.0 is no longer keeping pace with technological and geopolitical developments worldwide [11].

There are three main problems that version 3.0 should solve:

- attribution of cyberattacks, i.e., establishing responsibility for a cyber operation. This problem is especially urgent in the context of artificial intelligence development, as today's operations often rely on autonomous agents trained with machine learning. And such agents, having adapted, can perform tasks without direct human intervention; accordingly, it is very difficult to determine who is responsible;
- digital sovereignty is also a concern because cyber operations are mostly aimed at the integrity of information, not at physical systems. Accordingly, it requires a rethinking from a legal point of view of the understanding of the violation of sovereignty—not only physical invasion, but also manipulation of digital systems, thanks to which the state functions;
- uncertainty of the legal threshold in the context of hybrid warfare. Typically, cyberwarfare activity is part of a hybrid war that combines traditional military actions with economic factors, psychological pressure, information weapons, and the like. Intensity, directness, and scope, as criteria for classifying a cyber operation as a manifestation of the use of force, are difficult to determine in the course of hybrid strategies [12].

Thus, the legal status of civilians participating in cyber operations during military conflicts remains unresolved, and there is no relevant international legal instrument that regulates this issue. Nevertheless, in a changing world, it is important that states not only develop laws and unite to work on international documents on cybersecurity, but also apply legal norms, taking into account the real practical conduct of cyber operations.

3. Conclusions

Summarizing the above, it should be noted that cyberspace, although not a battlefield in the traditional sense, is increasingly becoming a domain where actions directly affect the balance of power of the warring parties. That is why distinguishing the legal status of civilians participating in cyber operations during hostilities has become extremely relevant.

Based on the analysis, it is clear that, *de jure*, the norms of international humanitarian law extend to cyberspace, but *de facto*, it is difficult to apply them in such cases. First of all, difficulties arise with the application of the principle of distinction, according to which the warring parties are obliged to clearly distinguish between civilians and combatants and the corresponding objects. This is due to the dual-use nature of digital infrastructure, the anonymity of users, the decentralized nature of cyber operations, and the frequent involvement of civilians in performing cyber support functions.

The study found that the main criterion for a civilian to lose protection from an attack is direct participation in hostilities (DPH), which requires the simultaneous presence of three elements: threshold of harm, direct causation, and belligerent nexus. However, the applicability of these criteria to cyber operations is controversial due to the possibility of attacks being carried out remotely, without physical presence in the combat zone, and the potential for their consequences to appear much later. This created a problem in determining the moment of a civilian's direct participation in hostilities.

The temporal logic of DPH is also an urgent problem. If, in traditional military conflicts, a civilian loses protection only for the duration of a specific DPH action according to the revolving-door principle, then in a cyber operation, it is difficult to determine such a time window. Participation in a cyber event

may appear continuous from a technical point of view, even though it is intermittent. That is why cyberspace actually destabilizes the traditional temporal logic of DPH.

A detailed analysis of the Tallinn Manual 2.0, a fundamental doctrinal source for the application of international law to cyber operations, shows that it can be applied to such processes, but many gaps remain. Thus, the problems of establishing the moment of loss of protection by a civilian, defining the boundaries of DPH, assessing the role of organized civilian cyber groups, and the like remain unresolved.

It should be noted separately that a civilian participating in a cyber operation does not automatically acquire the status of a combatant. He still remains a civilian, but temporarily, for the period of DPH, loses protection from direct attack. At the same time, in the case of members of organized groups, the concept of continuous combat function can be applied, according to which protection is not renewed after each separate act of participation in hostilities. This concept may become one of the key ones for the further development of legal regulation of civilian participation in cyber operations.

In general, the research conducted allows us to conclude that modern international humanitarian law is at an early stage of adaptation to the realities of cyberspace. Existing legal categories can be formally applied, but the rapid growth in the quantity and quality of cyber operations reveals their functional limitations. Further improvement of international legal regulation should aim to clarify the criteria for direct participation in hostilities in cyberspace, develop clearer approaches to the attribution of cyberattacks, determine the legal status of organized civilian cyber groups, and adapt the principle of distinction to the conditions of the digital environment. The development of the Tallinn Manual 3.0 confirms that the international community is aware of the inadequacy of existing legal regulation and the need to further rethink the application of international humanitarian law in cyberwar.

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

- [1] M. Kalnyk, *Conventional War: Essence, History and Modern Realities*, 2026. URL: <https://nfront.org.ua/konvencziyna-vijna-sut-istoriya-ta-suchasni-realiyi/>.
- [2] D. Viedienieiev, O. Semeniuk, *Essential Features and Functional Components of Modern Unconventional (Hybrid) Conflict*, *Strategic Panorama* (2021) 30–47. doi:10.53679/2616-9460.1-2.2021.02.
- [3] International Committee of the Red Cross, *International Humanitarian Law and Cyber Operations During Armed Conflicts*, Technical Report, International Committee of the Red Cross, 2019. URL: https://www.icrc.org/sites/default/files/document/file_list/icrc_ihl-and-cyber-operations-during-armed-conflicts.pdf.
- [4] State Service for Special Communication and Information Protection of Ukraine, *Another Cyberattack on Government Websites and Banks*, 2022. URL: <https://www.cip.gov.ua/ua/news/chergova-kiberataka-na-saiti-derzhavnikh-organiv-ta-banki>.
- [5] V. Filinovich, A. Mukasheva, *Can Hacking Be Ethical: A Legal Review*, in: *CH&CMiGIN'25: Fourth International Conference on Cyber Hygiene & Conflict Management in Global Information Networks*, CEUR Workshop Proceedings, Kyiv, Ukraine, 2025, pp. 163–169. URL: <https://ceur-ws.org/Vol-4024/short04.pdf>.
- [6] Ministry of Defense of Ukraine, *On Approval of the Instruction on the Procedure for Implementing the Norms of International Humanitarian Law in the Armed Forces of Ukraine*: Order No. 164, 2017. URL: <https://zakon.rada.gov.ua/laws/show/z0704-17#Text>.
- [7] United Nations, *Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts (Protocol I)*, 1977. URL: <https://treaties.un.org/pages/showdetails.aspx?objid=08000002800f3586>.

- [8] N. Melzer, Interpretive Guidance on the Notion of Direct Participation in Hostilities Under International Humanitarian Law, International Committee of the Red Cross, Geneva, 2009. URL: <https://www.refworld.org/policy/legalguidance/icrc/2009/68382>.
- [9] F. Delerue, Civilian Direct Participation in Cyber Hostilities, IDP. Revista de Internet, Derecho y Politica (2014) 3. doi:10.7238/idp.v0i19.2425.
- [10] M. N. Schmitt (Ed.), Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations, Cambridge University Press, Cambridge, 2017. doi:10.1017/9781316822524.
- [11] NATO Cooperative Cyber Defence Centre of Excellence, The Tallinn Manual, 2026. URL: <https://ccdcoe.org/research/tallinn-manual/>.
- [12] Journal of Law & Cyber Warfare, Tallinn Manual 3.0: Sovereignty and Attribution in 2025, 2025. URL: <https://www.jlcw.org/2025/07/22/tallinn-manual-3-0-sovereignty-and-attribution-in-2025-cyber-warfare/>.