

# CYBER HOSTILITIES: CIVILIAN DIRECT PARTICIPATION

Dan-Iulian VOITAȘEC\*

## Abstract

*The manner in which hostilities are being conducted has changed in recent years. The battle field has transpired beyond the physical realm and now has a virtual component. Because of this, it is now easier than ever for civilians to get involved in hostilities. International Humanitarian Law applies to all situations of armed conflict and according to the principle of distinction, the parties to the conflict must, at all times, distinguish between civilians and combatants. The problem arises when the line between combatants and civilians starts to get blurry. Direct civilian participation in hostilities has been addressed in both Additional Protocols to the Geneva Conventions of 1949 and in 2009 the International Committee of the Red Cross published the Interpretive guidance on the notion of Direct Participation in Hostilities under international humanitarian law. Another document that addresses the problem of civilian direct participation is the Tallinn Manual on the International Law Applicable to Cyber Warfare prepared by an international group of experts at the invitation of the NATO Cooperative Cyber Defence Centre of Excellence in 2013. The guide prepared by the ICRC addresses the problem of civilian direct participation during conventional situations of armed conflict, while the Tallinn Manual addresses direct participation in situations of cyber warfare. The purpose of this paper is to study the application of civilian direct participation to situations of cyber warfare.*

**Keywords:** international humanitarian law, civilian direct participation, cyber war, cyber-attacks, Tallinn Manual.

## 1. Introduction

According to a customary IHL rule, civilians are persons who are not members of the armed forces and the civilian population comprises all persons who are civilians<sup>1</sup>. A more complex definition of civilians and civilian population can be found in Article 50 of the Additional Protocol I (AP I) to the Geneva Conventions of 1949<sup>2</sup>. A fundamental principle of International Humanitarian Law (IHL), the principle of distinction, states that parties to the conflict must, at all times, distinguish between civilians and combatants and that attacks may only be directed against combatants. We see that IHL protects the civilian population but what happens when the line between civilians and combatants gets blurred. Most modern conflicts are no longer international in nature; this means that they are not fought between two or more states but between the governmental authorities of a certain state and organized armed groups or between such groups within a state. This has led to fighting in civilian populated areas and increased the number of civilians that get involved in the conduct of

hostilities. Also, the development of cyber warfare has allowed for an increasing number of civilians to get involved in hostilities.

What are the consequences that civilians face if they get involved in hostilities? The answer to that question comes in the form of direct participation in hostilities. The notion of direct participation in hostilities (DPH) has been addressed in both AP to the Geneva Conventions; in the Interpretive guidance on the notion of Direct Participation in Hostilities under international humanitarian law published by the ICRC in 2009 and in the Tallinn Manual published by the NATO Cooperative Cyber Defence Centre of Excellence in 2013. This article will focus on direct civilian participation in hostilities in situations of cyber warfare.

## 2. Content

The notion of direct participation in hostilities can be found in Article 51 (3) of Additional Protocol I to the Geneva Conventions<sup>3</sup> that focuses on international armed conflict and in Article 13 (3) of Additional Protocol II<sup>4</sup> to the Geneva Conventions

---

\* PhD candidate, Faculty of Law, "Nicolae Titulescu" University of Bucharest (e-mail: [dan.voitasec@gmail.com](mailto:dan.voitasec@gmail.com)).

<sup>1</sup> International Committee of the Red Cross – Customary IHL Database – Rule 5 – Definition of civilians accessed February 22, 2016. - [https://www.icrc.org/customary-ihl/eng/docs/v1\\_cha\\_chapter1\\_rule5](https://www.icrc.org/customary-ihl/eng/docs/v1_cha_chapter1_rule5).

<sup>2</sup> AP I to the GC of 1949 – art. 50:

1. A civilian is any person who does not belong to one of the categories of persons referred to in Article 4 A (1), (2), (3) and (6) of the Third Convention and in Article 43 of this Protocol. In case of doubt whether a person is a civilian, that person shall be considered to be a civilian.

2. The civilian population comprises all persons who are civilians.

3. The presence within the civilian population of individuals who do not come within the definition of civilians does not deprive the population of its civilian character.

<sup>3</sup> Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection Of Victims Of International Armed Conflicts Of 8 June 1977.

<sup>4</sup> Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection Of Victims Of Non-International Armed Conflicts Of 8 June 1977.

that deals with non-international armed conflicts. The notion of DPH is not defined in the Protocols, it is just stated that civilians are protected against attacks “unless and for such time as they take a direct part in hostilities”. Although the two AP are not ratified by all United Nation (UN) members, some parts of the Protocols are considered customary IHL. The fact that direct participation in hostilities leads to a loss of protection is also stated in ICRC’s Customary IHL study. Rule 6 of the study reflects the wording of Article 51 (3) and Article 13 (3). In the commentary on Rule 6 it is also stated that a precise definition of DPH does not exist<sup>5</sup>.

Due to the lack of definition and guidance on the interpretation of the concept of direct participation in hostilities, the ICRC launched, in 2003, a research project to further explore the subject. The results of the project were published by the ICRC in May 2009. The “Interpretive Guidance on the Notion of Direct Participation in Hostilities under International Humanitarian Law” is not a binding document and due to its controversial nature it expresses solely the ICRC’s views<sup>6</sup>. Although controversial, even critics of the project agreed that the Guide “clearly advances general understanding of the complex notion of “direct participation”<sup>7</sup>.

The ICRC made a number of 10 recommendations concerning the interpretation of IHL relating to the notion of DPH. These recommendations concern the following areas: the concept of civilian in international armed conflict, the concept of civilian in non-international armed conflict, private contractors and civilian employees, direct participation in hostilities as a specific act, constitutive elements of DPH, beginning and end of DPH, temporal scope of the loss of protection, precautions and presumptions in situations of doubt, restraints on the use of force in direct attack, consequences of regaining civilian protection.

In the Interpretative Guidance, the ICRC defines the term “civilian” differently for situations of international armed conflict and non-international armed conflict. The concept of civilian for the purposes of the principle of distinction in international armed conflict is defined as “*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.*”<sup>8</sup> The ICRC mimics the wording of API article 50 and 51 (3) and defines civilians negatively as not being part of certain groups.

A different approach can be found while defining civilians for the purposes of the principle of distinction in non-international armed conflict, in these situations civilians are defined as “*all persons who are not members of State armed forces or organized armed groups of a party to the conflict are civilians and, therefore, entitled to protection against direct attack unless and for such time as they take a direct part in hostilities. In non-international armed conflict, organized armed groups constitute the armed forces of a non-State party to the conflict and consist only of individuals whose continuous function it is to take a direct part in hostilities (“continuous combat function”).*”<sup>9</sup>

The ICRC introduces the concept of *continuous combat function* to distinguish members of organized armed groups that function as the armed forces of a non-State actor party to the conflict from civilians that take part in hostilities and form the civilian population. This means that a persons that serves a continuous combat function, for example “individuals whose continuous function involves the preparation, execution, or command of acts or operations amounting to direct participation in hostilities”<sup>10</sup> are subject to attack at any time.

The concept of continuous combat function does not apply to those individuals that support an organized armed group but are not directly involved in hostilities. According to the ICRC’s commentary “recruiters, trainers, financiers and propagandists may continuously contribute to the general war effort of a non-State party, but they are not members of an organized armed group belonging to that party unless their function additionally includes activities amounting to direct participation in hostilities”<sup>11</sup>. Also the concept does not apply to civilians that undertake sporadic actions for an organized armed group that amounts to direct participation. These civilians lose the protection granted to them by IHL only for such time as they directly participate in hostilities.

Regarding civilians in situations of cyber hostilities, Rule 29 of the Tallinn Manual does not define the civilian population but states that “civilians are not prohibited from directly participating in cyber operations amounting to

<sup>5</sup> ICRC – Customary IHL Database – Rule 6 Civilians’ Loss of Protection from Attack. Accessed on 24.02.2016 - [https://www.icrc.org/customary-ihl/eng/docs/v1\\_rul\\_rule6](https://www.icrc.org/customary-ihl/eng/docs/v1_rul_rule6)

<sup>6</sup> ICRC, Interpretive Guidance on the Notion of Direct Participation in Hostilities under International Humanitarian Law (Nils Melzer, 2009) p. 6 – available at <https://www.icrc.org/eng/assets/files/other/icrc-002-0990.pdf>

<sup>7</sup> Michael N. Schmitt - The Interpretive Guidance on the Notion of Direct Participation in Hostilities: A Critical Analysis (Harvard National Security Journal, vol.1, 2010) p. 2 Accessed on 24.02.2016 - <http://ssrn.com/abstract=1600243>. Michael N. Schmitt was one of the experts that took part in the research project and withdrew his name after reviewing the final draft.

<sup>8</sup> Interpretive Guidance (2009) - p. 20.

<sup>9</sup> Idem p. 27.

<sup>10</sup> Idem p. 34.

<sup>11</sup> Idem.

hostilities but forfeit their protection from attacks for such time as they so participate.”<sup>12</sup> It is also stated in the Tallinn Manual that the principle of distinction shall apply to cyber attacks<sup>13</sup> and that the civilian population should not be subject to cyber attacks<sup>14</sup>.

Another recommendation made by the ICRC in the Interpretative Guidance refers to the constitutive elements of direct participation in hostilities. An act must meet three cumulative criteria to qualify as direct participation in hostilities: threshold of harm, direct causation and belligerent nexus.<sup>15</sup> The group of experts that worked on the Tallinn Manual “generally agreed with the three cumulative criteria set forth by the ICRC Interpretative Guidance”<sup>16</sup>

For a specific act to reach the *threshold of harm* it “*must be likely to adversely affect the military operations or military capacity of a party to an armed conflict or, alternatively, to inflict death, injury, or destruction on persons or objects protected against direct attack.*”<sup>17</sup> The definition distinguishes between two situations. Firstly for an act to reach the threshold of harm it must adversely affect the military operations of a party to the conflict. The other situation in which an act can reach the threshold of harm is when it inflicts death, injury or destruction on protected persons or objects.

For the first situation presented in the definition, to reach the threshold of harm the effects of an action *must be likely* to produce harm but the actual materialization is not required.<sup>18</sup>

The ICRC also refers to cyber operations launched by civilians against the military operations or capacity to a party to the conflict. In the Guidance it is stated that: “*Electronic interference with military computer networks could also suffice, whether through computer network attacks (CNA) or computer network exploitation (CNE), as well as wiretapping the adversary’s high command or transmitting tactical targeting information for an attack.*”<sup>19</sup> A cyber operation launched by a civilian which causes military harm will reach the threshold of harm.

In the situation where an act does not cause harm to the military capacity or military operations of a party to an armed conflict, an action must be likely to inflict death, injury or destruction on persons or objects protected against direct attack to reach the threshold of harm. This situation refers to actions directed against civilians and civilian objects. According to the ICRC *the interruption of electricity, water, or food supplies [...] the manipulation of computer networks [...] would not reach the threshold of harm in the absence of adverse military effects.*<sup>20</sup>

Although the Tallinn Manual generally agrees with the criterion put forward by the Interpretive Guidance some differences of opinion can be found. Firstly, the Tallinn Manual excludes the “likelihood” description from its interpretation of the threshold of harm and includes intention<sup>21</sup>. Also, in the Tallinn Manual it is stated that “there is no requirement for physical damage to objects or harm to individuals”<sup>22</sup> which means that actions that do not qualify as a cyber attack<sup>23</sup> will still reach the threshold of harm if they cause military harm to the enemy.<sup>24</sup>

The second constitutive element of direct participation in hostilities is called *direct causation* and is defined in the Interpretive Guidance as follows: “*In order for the requirement of direct causation to be satisfied, there must be a direct causal link between a specific act and the harm likely to result either from that act, or from a coordinated military operation of which that act constitutes an integral part.*”<sup>25</sup>

In the Guidance, the ICRC differentiates between direct and indirect participation in hostilities. For an act to be considered direct participation in hostilities “there must be a sufficiently close causal relation between the act and the resulting harm”<sup>26</sup>. The ICRC affirms that the harm caused by an action had to be “brought in one causal step”<sup>27</sup> to satisfy the direct causation criterion. Another requirement set forward by the ICRC to satisfy the direct causation criterion is that the harm has to be caused directly, it is not sufficient that the

<sup>12</sup> Michael N. Schmitt et al., Tallinn Manual on the International Law Applicable to Cyber Warfare – Cambridge University Press, Cambridge, 2013 p. 90.

<sup>13</sup> Idem Rule 31, p. 95.

<sup>14</sup> Idem Rule 32, p. 97.

<sup>15</sup> Interpretive Guidance (2009) - p. 46.

<sup>16</sup> Tallinn Manual, Rule 35, p. 102.

<sup>17</sup> Interpretive Guidance (2009) - p. 47.

<sup>18</sup> Idem.

<sup>19</sup> Idem – p. 48.

<sup>20</sup> Idem – p.50.

<sup>21</sup> Collin Allan - Direct Participation in Hostilities from Cyberspace – p. 181 (Virginia Journal of International Law, Vol. 54, No. 1, 2013) Accessed on 24.02.2016 - <http://ssrn.com/abstract=2617867>

<sup>22</sup> Tallinn Manual – p. 102.

<sup>23</sup> Tallinn Manual – Rule 30 Definition of Cyber Attack: A cyber attack is a cyber operation, whether offensive or defensive, that is reasonably expected to cause injury or death to persons or damage or destruction to objects.

<sup>24</sup> Tallinn Manual – p. 102.

<sup>25</sup> Interpretive Guidance – p. 51.

<sup>26</sup> Idem – p. 52.

<sup>27</sup> Idem – p. 53.

act and its consequences to be connected through an uninterrupted causal chain of events<sup>28</sup>. For example acts such as providing financial assistance or certain services (electricity, fuel) to a party to an armed conflict are considered indirect participation.

In the case of collective operations an act that does not reach the threshold of harm could still satisfy the direct causation criterion if it “constitutes an integral part of a concrete and coordinated tactical operation that directly causes such harm.”<sup>29</sup>

No commentary on the direct causation criterion is found in the Tallinn Manual. It is only stated that “a direct causal link between the act in question and the harm intended or inflicted must exist”<sup>30</sup> for the action of a civilian to qualify as direct participation.

The last of the three cumulative criteria needed to qualify an act of a civilian as direct participation in hostilities is the *belligerent nexus*. In the Guidance the belligerent nexus is defined as 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.*”<sup>31</sup>

The concept of direct participation in hostilities is restricted only to those acts that are an integral part of the hostilities<sup>32</sup>. Not all acts that cause harm in a situation of armed conflict can be linked to a party to the conflict. Even in situations of armed conflict some civilian acts are driven by private gain and they will engage in acts that cause harm. One of the examples given by the ICRC is the exchange of fire between civilians and the police during a bank robbery.<sup>33</sup> The belligerent nexus does not exist because the action was not realized in support of a party to the conflict. The action of a civilian to resort to violence in self-defense lacks belligerent nexus<sup>34</sup>.

Even though, out of the three criteria, the belligerent nexus is the least controversial, Michael Schmitt considers that the criterion should be defined differently, instead of *in support of a party to the conflict and to the detriment of another to an act in support or to the detriment of a party*<sup>35</sup>. A similar approach can be found in the Tallinn Manual where the belligerent nexus is interpreted as an act directly related to the hostilities<sup>36</sup>. The belligerent nexus of a cyber operation launched by a civilian exists when it is directly related to the hostilities, the requirements that

the act should be in support of a party to the conflict and to the detriment of another is not taken into consideration by the Tallinn Manual.

The group of experts involved in writing the Tallinn Manual agreed with the ICRC that acts of purely criminal or private nature are ruled out of the belligerent nexus criterion.<sup>37</sup>

Another problem arising from the concept is the beginning and the end of direct participation in hostilities. According to the sixth recommendation made by the ICRC in the Interpretive Guidance “*measures 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.*”<sup>38</sup>

Preparatory measures amount to direct participation in hostilities when they are an integral part of a hostile act. Preparatory measures “*aiming to carry out a specific hostile act qualify as direct participation in hostilities, whereas preparatory measures aiming to establish the general capacity to carry out unspecified hostile acts do not*”<sup>39</sup> The ICRC also states that the temporary proximity or close geographical proximity is not required for the preparatory measures to be an integral part of the hostile act.<sup>40</sup>

In the Guidance it is stated that for acts that do not require geographical displacement as is the case for certain cyber operations “*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 act*”<sup>41</sup> In the case of cyber operations launched by civilians only the execution of the act will be considered direct participation in hostilities according to the ICRC. The civilians will not be targetable for the duration of the deployment and return.

This view is not shared by the Tallinn Manual expert’s that state: “*Any act of direct participation in hostilities by a civilian renders that person targetable for such time as he or she is engaged in the qualifying act of direct participation. All of the Experts agreed that this would at least include actions immediately preceding or subsequent to the qualifying act. For instance, travelling to and from the location where a computer used to mount an operation is based would be encompassed in the notion.*”<sup>42</sup>

<sup>28</sup> Idem – p. 54.

<sup>29</sup> Idem – p. 55.

<sup>30</sup> Tallinn Manual – p. 102.

<sup>31</sup> Interpretative Guidance – p. 58.

<sup>32</sup> Idem.

<sup>33</sup> Idem – p. 60.

<sup>34</sup> Idem – p. 61.

<sup>35</sup> Michael N. Schmitt (2010) – p. 34.

<sup>36</sup> Tallinn Manual – p. 102.

<sup>37</sup> Idem – p. 103.

<sup>38</sup> Interpretative Guidance – p. 65.

<sup>39</sup> Idem – p. 66.

<sup>40</sup> Idem.

<sup>41</sup> Idem – p. 68.

<sup>42</sup> Tallinn Manual – p. 103.

The seventh recommendation made by the ICRC refers to the temporal scope of the loss of protection. In the Guidance the ICRC differentiates between civilians who sporadically participate in hostilities and those who serve a continuous combat function: *“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.”*<sup>43</sup>

The ICRC based this definition on the two AP to the Geneva Conventions that state that civilians will be protected against attack “unless and for such time as they take a direct part in hostilities”<sup>44</sup>. The part of the definition found in the Guidance regarding the loss of protection of civilians who sporadically take part in hostilities is called “the revolving door” due to the loss and regaining of protection<sup>45</sup>. The revolving door concept is highly controversial, even the experts that worked on the Tallinn Manual could not reach a consensus regarding this issue. Some experts accepted the view presented in the Interpretive Guidance and stated that civilians who launched repeated cyber operations that qualify as direct attack should lose protection against attack for the duration of each specific acts while others argued that direct participation begins with the first cyber operation launched and continues “throughout the period of intermittent activity”<sup>46</sup>. In the case of cyber operations some problems may arise due to the short time span in which they take place, some cyber

attacks last minutes or less while others will have delayed effects. In the opinion of Michael N. Schmitt if the revolving door concept is applied to cyber operations “there would appear to be no window of opportunity for the victim of an attack to respond”<sup>47</sup>. Schmitt considers that the reasonable interpretation of “for such time” in a situation of cyber conflict is to encompass “the entire period during which the direct cyber participant is engaging in repeated cyber operations”<sup>48</sup>

### 3. Conclusions

The ICRC is the leading humanitarian organization and the States party to the Geneva Conventions have given it a mandate to protect victims of international and non-international armed conflict thus the views of the ICRC will always be geared towards protecting the victims. On the other hand, IHL experts who have a different professional background may not always share the views of the ICRC. This situation can be observed in the different approaches found in the Interpretive Guidance and the Tallinn Manual. Although the Interpretive Guidance and the Tallinn Manual share the same view regarding the three cumulative requirements needed for direct participation in hostilities, the interpretation of certain criteria may differ. This article has highlighted that the ICRC’s interpretation of direct participation in hostilities offers a higher degree of protection to the civilian population while the Tallinn Manual increases the window of opportunity for civilians who take part in cyber hostilities to be targeted.

### References:

- Michael N. Schmitt et al., Tallinn Manual on the International Law Applicable to Cyber Warfare (Cambridge: Cambridge University Press, 2013);
- Michael N. Schmitt - Cyber Operations and the Jus in Bello: Key Issues (2011);
- Michael N. Schmitt - The Interpretive Guidance on the Notion of Direct Participation in Hostilities: A Critical Analysis (Harvard National Security Journal, vol.1, 2010);
- ICRC, Interpretive Guidance on the Notion of Direct Participation in Hostilities under International Humanitarian Law (Nils Melzer, 2009);
- ICRC – Customary IHL Database;
- Collin Allan - Direct Participation in Hostilities from Cyberspace (Virginia Journal of International Law, Vol. 54, No. 1, 2013);
- Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts of 8 June 1977;
- Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection Of Victims Of Non-International Armed Conflicts Of 8 June 1977.

<sup>43</sup> Interpretative Guidance – p.70.

<sup>44</sup> AP I to GC – article 51 (3), AP II to GC – article 13 (3).

<sup>45</sup> Interpretative Guidance – p. 70.

<sup>46</sup> Tallinn Manual – p. 104.

<sup>47</sup> Michael N. Schmitt - Cyber Operations and the Jus in Bello: Key Issues (2011) – p. 14. Accessed on 02.03.2016 - [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1801176](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1801176)

<sup>48</sup> Idem.