Journal of Conflict & Security Law © Oxford University Press 2012; all rights reserved.
For permissions, please e-mail: journals.permissions@oup.com
doi:10.1093/jcsl/krs021

Cyber Warfare and the Notion of Direct Participation in Hostilities

David Turns*

Abstract

The domain of cyber warfare being relatively new, it is not yet matched by any comparatively novel international legal paradigm; the cyber conflicts of the present and (probably) the future therefore fall to be regulated under the existing *lex lata*. This article, assuming a scenario of international armed conflict, seeks as a specific example to apply the notion of direct participation in hostilities from Additional Protocol I (1977) to cyber war. This aspect of the topic is likely to assume particular importance in light of the contemporary tendency in many developed, Western armed forces to outsource technical specialist work (like information technology) to civilians. Whether or not such civilians can be said to be directly participating in hostilities—based on the accepted constitutive elements of threshold of harm, direct causation and belligerent nexus identified in the International Committee of the Red Cross' *Interpretive Guidance* (2005)—will also have implications for the objects and places that could lawfully be targeted in future cyber conflicts.

1. Introduction

Any discussion of the international humanitarian law (IHL) notion of direct participation in hostilities (DPH) in the context of cyber warfare (CW) is fraught with difficulties and uncertainties. Is CW in and of itself a form of armed conflict, as the term is understood in IHL? If so, what type of armed conflict is it?[1] Who are the actors in such putative armed conflicts, and what is their status according to the international law of armed conflict (LOAC)? Are they and their systems fundamentally military or civilian in nature? What do they actually do and—if they are civilians—do those actions constitute DPH? What are the consequences of such a determination, both for the civilians involved and for the physical places where they carry out their activities? While some solace might be derived from the fact that these are, as a former

* Senior Lecturer in International Laws of Armed Conflict, Defence Academy of the UK (Cranfield University). Opinions and interpretations expressed in this article are solely those of the author and do not necessarily represent those of the Armed Forces, Ministry of Defence or Government of the UK. Email: d.turns@cranfield.ac.uk

[1] These first two questions are outside the scope of this article, which focuses on aspects of the *jus in bello* as they apply to CW. See, however, M Schmitt, 'Classification of Cyber Conflict' in this volume.

US Secretary of Defence once infelicitously but famously put it, 'known unknowns',[2] this is counterbalanced by an increasing awareness on the part of technologically advanced, mostly (but not exclusively) Western States, of the reality that CW is already present as a means of warfare—and is only likely to become more widely used in future conflicts.[3] Although it is probably still the case that only a relatively small number of States in the world as yet have the current capability to engage in large-scale CW—the most frequently mentioned ones are China, India, Israel, Russia and the USA[4]—it is equally true that other nations with highly developed military establishments are scrambling to keep up.[5] It has also been noted, chillingly, that CW

> represents 'war on the cheap' for an otherwise technology starved belligerent, since cost is limited to acquisition of off-the-shelf computers and exploitation software, access to the target network, and computer expertise. Moreover, the higher-tech an opponent, the more vulnerable it is to such attacks.[6]

Technology, therefore, cuts both ways; and so does the law that applies to its use in situations of armed conflict.

This aspect of technology and hostilities is thrown into particularly sharp relief by the issue of DPH, much of the discourse on which has revolved around the (generally unspoken) assumption that it is primarily of relevance to Western

[2] Donald Rumsfeld's full quotation reads, 'Reports that say that something hasn't happened are always interesting to me, because as we know, there are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns – the ones we don't know we don't know': US Department of Defense, 'DoD News Briefing—Secretary Rumsfeld and Gen Myers' (12 February 2002) <http://www.defense.gov/transcripts/transcript.aspx?transcriptid=2636> (accessed 19 March 2012).

[3] See 'Stuxnet Worm Heralds New Era of Global Cyberwar' *The Guardian* (London, 30 September 2010) <http://www.guardian.co.uk/technology/2010/sep/30/stuxnet-worm-new-era-global-cyberwar?INTCMP=ILCNETTXT3487> (accessed 19 March 2012).

[4] See CG Billo and W Chang, 'Cyber Warfare—An Analysis of the Means and Motivations of Selected Nation States' Institute for Security Studies at Dartmouth College (November 2004) <http://www.ists.dartmouth.edu/docs/cyberwarfare.pdf> (accessed 19 March 2012).

[5] 'UK Developing Cyber-weapons Programme to Counter Cyber War Threat' *The Guardian* (London, 30 May 2011) <http://www.guardian.co.uk/uk/2011/may/30/military-cyberwar-offensive> (accessed 19 March 2012). The UK's most recent Strategic Defence and Security Review has set aside £650 m to develop the country's cyber security: 'Stuxnet Attack Forced Britain to Rethink the Cyber War' *The Guardian* (London, 30 May 2011) <http://www.guardian.co.uk/politics/2011/may/30/stuxnet-attack-cyber-war-iran> (accessed 19 March 2012).

[6] MN Schmitt, 'War, Technology, and International Humanitarian Law' Harvard University Program on Humanitarian Policy and Conflict Research, Occasional Paper Series 4 (2005) 43, <http://www.hpcrresearch.org/sites/default/files/publications/OccasionalPaper4.pdf> (accessed 21 March 2012).

States seeking to determine the application of the notion of DPH *vis-à-vis* the irregular fighters who typically form their opponents in most contemporary armed conflicts: multifarious insurgents in Iraq, or Taliban fighters in Afghanistan. As Schmitt's comment quoted above indicates, however, it would not take very much in the way of resources or expertise for such non-State actors to be able to carry out cyber attacks against their Western opponents. By the same token, the question of the legal status of those Western personnel who engage in CW would become an issue in that—assuming they are civilians[7]—they would lose their protection and become subject to attack for such time as they were taking a direct part in hostilities. This would reverse the paradigmatic application of the notion of DPH as (Western) international lawyers have for the last few years been thinking about it. It would also raise the question of the legal status, as military objectives, of the facilities from which those personnel operate. Although it is most unlikely that the irregular opponents of Western States would comply with IHL in any computer network attacks (CNA) that they might carry out (mirroring their conventional operations), and thus would probably not bother to observe the principle of distinction that lies at the heart of the law in armed conflicts, a perception that Western States are employing civilians—who might appear to be directly participating in hostilities—to engage in CW could be very damaging to those States in public relations terms. The legal issues, far from being abstract and theoretical, are therefore of very great practical and operational import.

While reports of CNA are increasingly often encountered in the media[8] or through oral hearsay,[9] those incidents are probably best considered within the parameters of the *jus ad bellum*. The question of whether or not a given CNA amounts to an armed attack as an act of aggression or for the purposes of giving rise to a State's legal right to use force in self-defence is considered elsewhere in this symposium[10] and will also be the subject of detailed analysis by the present

[7] The issue of DPH does not arise in the situation of Coalition troops fighting in such theatres as Afghanistan, because as members of State armed forces, they are combatants and subject to attack in any event. On the Coalition side, the problem would arise only in relation to civilian personnel who operate computer systems that rise to the level of direct participation in hostilities.

[8] Eg 'Stuxnet worm "Targeted High-value Iranian Assets"' *BBC News* (23 September 2010) <http://www.bbc.co.uk/news/technology-11388018> (accessed 19 March 2012); 'Russia Accused of Unleashing Cyberwar to Disable Estonia' *The Guardian* (London, 17 May 2007) <http://www.guardian.co.uk/world/2007/may/17/topstories3.russia> (accessed 19 March 2012).

[9] Eg a former student of the present author at the UK Defence Academy, a cyber-operations expert in the Indian Ministry of Defence, stated that there are many daily 'pinprick' or probing attacks on the Indian defence establishment's cyber defences and firewalls, which he asserted were emanating from network users (State-sponsored or otherwise) in the People's Republic of China.

[10] See further the contributions by Russell Buchan and Nicholas Tsagourias, published elsewhere in this journal.

author in a separate work,[11] but the question of whether or not a given act of CW amounts to DPH is a question properly to be considered within the framework of the *jus in bello*, that is to say, in the context of an on-going international[12] armed conflict. Examples of such situations—where CNA has verifiably taken place during an actual contemporary armed conflict—are scanty in the extreme; to date, the only case in which CNA has coincided with a regular 'shooting war' appears to have been the South Ossetian War between Russia and Georgia in 2008.[13] Nevertheless, since it does seem clear that CNA can occur in the context of armed hostilities between States, however sparse the evidence of such instances to date, this article is written on the assumption that CNA may and should be governed by the strictures of LOAC in such situations.

There is a complete lack of international jurisprudence on CW: it is an activity that is currently developing so fast, and without a specific international legal regime to govern it, that there is no decided case law on the topic.[14] There is no treaty specific to CW and, while limited State practice does exist, there is virtually no evidence of the *opinio juris* required to make it into normative customary international law. What published authority there is for sources of international law on the subject, as defined in Article 38(1) of the Statute of the International Court of Justice, is at present restricted to the subsidiary source of academic commentary. It is therefore rather difficult to write authoritatively about international law and CW: one has a distinct feeling of the ink not yet being dry on the page (or perhaps, more appropriately, the script not yet being clear on the computer screen). Nevertheless, it is submitted that as CW exists as a matter of fact, it cannot do so in a legal vacuum: discussion of the

[11] D Turns, 'The Concept of "Attack" in Cyber Warfare' in D Saxon (ed), *International Humanitarian Law and the Changing Technology of War* (Brill 2013 forthcoming).

[12] Although the notion of DPH is not of itself limited to international armed conflicts, it is only the latter that have the formal distinction of personal status between combatants and civilians. In an attempt to achieve maximum clarity in its discussion of an essentially unclear sphere of activity, therefore, this article does not consider application of the notion to CW in non-international armed conflicts.

[13] 'Georgian Websites Forced Offline in "Cyber War"' *The Sydney Morning Herald* (Sydney, 12 August 2008) <http://www.smh.com.au/news/technology/georgian-websites-forced-offline-in-cyber-war/2008/08/12/1218306848654.html> (accessed 19 March 2012).

[14] Although Georgia did institute legal proceedings against Russia in the International Court of Justice following the 2008 conflict, for jurisdictional reasons it was forced to pursue a very narrow line of legal argumentation; issues relating to the use of CW by Russia against Georgia were not raised in the latter's Application to the Court, nor in its Memorial and Pleadings. The Court subsequently found that it did not have jurisdiction to hear the merits of the case: *Case Concerning Application of the International Convention on the Elimination of All Forms of Racial Discrimination (Georgia v Russian Federation)* (Preliminary Objections) 1 April 2011, <http://www.icj-cij.org/docket/files/140/16398.pdf> (accessed 2 April 2012).

phenomenon must take place within the parameters of the *lex lata* of IHL.[15] This article accordingly presents a 'snapshot', at the present moment in time, of a particular type of legal problem that will increasingly be encountered in this fast-evolving area of activity.

2. Preliminary Questions

A. What is CW?

There is no formally promulgated, internationally agreed definition as such of CW, but the US Department of Defense defines a combined concept of computer network operations (CNO)[16] as including CNA, computer network defence (CND) and computer network exploitation (CNE). CNA is defined as '[a]ctions taken through the use of computer networks to disrupt, deny, degrade or destroy information resident in computers and computer networks, or the computers and networks themselves'.[17] CND is defined as '[a]ctions taken to protect, monitor, analyze, detect, and respond to unauthorized activity within the Department of Defense information systems and computer networks'.[18] CNE is defined as '[e]nabling operations and intelligence collection capabilities conducted through the use of computer networks to gather data from target or adversary automated information systems or networks'.[19] Put all those ingredients together, and the resulting mixture is CW.

An apocalyptic vision of CW and its effects has been described as leading to

> . . . a catastrophic breakdown within 15 minutes. Computer bugs bring down military e-mail systems; oil refineries and pipelines explode; air-traffic-control systems collapse; freight and metro trains derail; financial data are scrambled; the electrical grid goes down in the eastern United States; orbiting satellites spin out of control. Society soon breaks down as food becomes scarce and money runs out. Worst of all, the identity of the attacker may remain a mystery.[20]

[15] By the same token, it is suggested that the cyber domain generally falls to be regulated within the established framework of relevant parts of public international law: see E Tikk, 'Ten Rules for Cyber Security' (2011) 53 Survival 119.

[16] US Department of Defense, 'Dictionary of Military and Associated Terms' (8 November 2010 as amended through 15 February 2012) Joint Publication 1-02, 66, <http://www.dtic.mil/doctrine/new_pubs/jp1_02.pdf> (accessed 23 March 2012).

[17] ibid 65.

[18] ibid.

[19] ibid.

[20] 'War in the Fifth Domain – Are the Mouse and Keyboard the New Weapons of Conflict?' *The Economist* (1 July 2010) <http://www.economist.com/node/16478792> (accessed 23 March 2012); in this excerpt the article cites an unnamed book by author Richard Clarke, 'a former White House staffer in charge of counter-terrorism and cyber-security'.

A less lurid and more technical commentary indicates that CW requires use of a data stream (as opposed to electromagnetic pulses, for instance) to achieve such actions as, *inter alia*,

> gaining access to a computer system so as to acquire control over it, transmitting viruses to destroy or alter data, using logic bombs that sit idle in a system until triggered on the occasion of a particular occurrence or at a set time, inserting worms that reproduce themselves upon entry into a system and thereby overloading the network, and employing sniffers to monitor and/or seize data.[21]

It is clear that CW consists of the non-kinetic application of force,[22] but while in principle it may have physical or virtual effects, it is important to understand that for the purposes of a DPH analysis, the effects obtained *must* be kinetic; this follows from the definition of the first of the three constitutive elements of DPH, discussed below. In any event, an act of CW within an armed conflict certainly constitutes an 'attack' within the meaning of contemporary IHL, *viz* an act of violence against an adversary, whether in offence or defence.[23] As General James E Cartwright, Commander US Strategic Command, stated:

> History teaches us that a purely defensive posture poses significant risks; the 'Maginot Line' model of terminal defense will ultimately fail without a more aggressive offshore strategy, one that more effectively layers and integrates our cyber capabilities. If we apply the principles of warfare to the cyber domain, as we do to sea, air, and land, we realize the defense of the nation is better served by capabilities enabling us to take the fight to our adversaries, when necessary to deter actions detrimental to our interests.[24]

The violence may not be in the act itself, which may consist of an action as apparently innocuous as hitting the 'Enter' key on a computer keyboard, but is above all determined by its result. The already-classic example is that of a virus which is used to infect a computer system controlling signals on a national

[21] MN Schmitt, 'Wired warfare: Computer Network Attack and *jus in bello*' (200) 84 Intl Rev Red Cross 365, 367.

[22] See Schmitt (n 6) 44.

[23] Art 49(1), Protocol I Additional to the Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts (opened for signature 8 June 1977, entered into force 7 December 1979) 16 ILM 1391 ('AP I'). See also Schmitt (n 1) s 3.

[24] US House of Representatives, Committee on Armed Services, 110th Congress 1st Session, 'Full Hearing on Budget Request from the US Strategic Command, Northern Command, Transportation Command, and Southern Command' (21 March 2007) HASC No 110–40, 65.

railway network:[25] the non-kinetic action is the keying in of a command which causes the virus to download, which causes the signalling equipment to malfunction, which in turn causes trains to crash, thereby killing people. In this example, the malfunctioning of the signalling equipment is the virtual effect of the attack, and the destruction of railway and adjacent property and any lives lost in the consequent accident(s) is the physical effect.

B. What is DPH?

The notion of DPH seeks to deal with the fact that in modern warfare the traditional distinction between combatants and civilians, on which so much of IHL is based, is often hard to maintain. It was introduced into modern IHL in 1977 with the following formula: 'Civilians shall enjoy the protection of this Section [of the Protocol], unless and for such time as they take a direct part in hostilities'.[26] This bland statement having benefitted from relatively little elaboration in the official *Commentary* published by the International Committee of the Red Cross (ICRC),[27] an international Group of Experts at the end of a 6-year 'clarification process' published an important guidance document on how the notion of DPH is to be interpreted.[28] The *Interpretive Guidance* is far from uncontroversial: a substantial number of the Experts expressly disassociated themselves from the final product[29] and, crucially, major military powers have yet to signal in any discernibly official manner the extent to which they agree or disagree with its conclusions.[30] Although the process has undeniably been useful in enabling matters of controversy to be aired and elaborated, uncertainties remain and it is not at all clear how the guidance might be applied in practice on the physical battlefield; this is *a fortiori* the case when it comes to the virtual battlefield.

[25] Schmitt gives examples of analogous actions such as attacks on air traffic control systems, the flow of oil pipelines, nuclear reactors or toxic chemical production/storage facilities: Schmitt (n 21) 374.

[26] Art 51(3), AP I. For the principle of distinction in cyber war see Y Dinstein, 'The Principle of Distinction and Cyber War in International Armed Conflicts' in this volume.

[27] See Y Sandoz, C Swinarski and B Zimmermann (eds), *Commentary on the Additional Protocols of 8 June 1977 to the Geneva Conventions of 12 August 1949* (ICRC 1987) paras 1942–44.

[28] ICRC, 'Interpretive Guidance on the Notion of Direct Participation in Hostilities under International Humanitarian Law' (2008) 90 Intl Rev Red Cross 991 ('Interpretive Guidance').

[29] For a flavour of the disagreements and debates, see Forum: Direct Participation in Hostilities: Perspectives on the ICRC Interpretive Guidance (2010) 42 NYU J Intl L Pol 637.

[30] Although many of the Experts were serving members of national armed forces or defence ministry officials, it was made clear from the beginning that they were involved only in their personal capacities and were not formally representing the views of their respective States.

Nevertheless, the *Interpretive Guidance* has been very useful in constructing a set of generally agreed parameters within which the debate about DPH can be conducted; in that sense at least, it may be considered authoritative. It posits three cumulative elements which together constitute the act of directly participating in hostilities:

1. the act must be likely to adversely affect the military operations of a party to an armed conflict or, alternatively, to inflict death, injury or destruction on persons or objects protected against direct attack (threshold of harm);
2. there must be a direct causal link between the 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 (direct causation);
3. the act 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 (belligerent nexus).[31]

The general aspects of these elements as to legal doctrine and military operations have been subject to detailed analysis and discussion elsewhere;[32] the present article is concerned exclusively with their interpretation in the specific context of CW.

C. The DPH Interpretive Guidance and CW

Although the Interpretive Guidance was not written specifically with CW in mind but seeks to elaborate the notion of DPH generically, there are some passages that have a direct bearing on the situation of civilians participating directly in cyber-hostilities. As to generalities, the 'military harm' required by the 'threshold of harm' criterion is explained broadly as including 'essentially any consequence adversely affecting the military operations or military capacity of a party to the conflict'.[33] Absent specific *military* harm to the adverse party, 'a specific act [constituting DPH] must be likely to cause at least death, injury or destruction';[34] thus, the causing of mere inconvenience, however unpleasant, would not suffice. With specific reference to CW, the Interpretive Guidance states: 'Electronic interference could . . . suffice [to cause military harm], whether through computer network attacks (CNA) or computer network exploitation

[31] Interpretive Guidance (n 28) 995–96.
[32] Eg MN Schmitt, 'Deconstructing Direct Participation in Hostilities: The Constitutive Elements' (2010) 42 NYU J Intl L Pol 697.
[33] Interpretive Guidance (n 28) 1017.
[34] ibid 1018.

(CNE)...'.[35] But on the other hand, 'the manipulation of computer networks... may have a serious impact on public security, health, and commerce ... However, they would not, in the absence of adverse military effects, cause the kind and degree of harm required to qualify as direct participation in hostilities'.[36] For example, disrupting the computer systems controlling a national railway network could cause the entire system to be shut down, resulting in cancellations across the network; this would be extremely inconvenient for the traveling public in the target State, but would not reach the required threshold of harm. If the system shut-down causes a signal to malfunction, leading to the derailment of a munitions train, that would constitute military harm; and if it causes on-board train computers to malfunction with the result that two passenger trains crash, that would constitute the 'death, injury or destruction' necessary to meet the required threshold of harm.

If we consider the cyber attacks on Estonia in 2007 and the cyber-elements of the Russia–Georgia conflict in 2008 in light of the above (assuming in both cases that they were perpetrated by civilians and not members of the Russian armed forces), we must conclude that in neither case would the acts in question have amounted to DPH as they would have failed to meet this criterion: the attacks on Estonian computers caused large-scale inconvenience in what is one of the most 'wired' countries in Europe, due to administrative, financial and social chaos when vital public computer systems went down, but there is no evidence that a single person died or was injured, or that any property was damaged or destroyed, as a direct result. In Georgia the impact was somewhat less, largely because the country is less computer reliant than Estonia for its public administration and banking systems; it appears to have been largely limited to propaganda effects (the website of the Georgian Presidency was defaced, for example). It would therefore have been *a fortiori* the case that in this instance there was no DPH.[37] On the other hand, the apparent use of the Stuxnet worm to target Iranian centrifuges used for the enrichment of uranium in 2010 would have amounted to DPH (had it occurred in a situation of armed conflict, which it did not), because it resulted in physical damage to the centrifuges.[38]

Turning to the requirement of direct causation, the *Interpretive Guidance* states that this, 'should be understood as meaning that the harm in question must be brought about in one causal step... it is not sufficient that the act and its

[35] ibid 1017–18. That there is no controversy that CNA is capable of amounting to DPH if it causes military harm is made clear at fn 101.

[36] ibid 1019.

[37] It should of course be remembered that in the Estonian case there was no actual situation of armed conflict under IHL, whereas in Georgia there was.

[38] See D Albright, P Brannan and C Walrond, 'Did Stuxnet Take Out 1,000 Centrifuges at the Natanz Enrichment Plant?' (Institute for Science and International Security Report 22 December 2010) <http://isis-online.org/uploads/isis-reports/documents/stuxnet_FEP_22Dec2010.pdf> (accessed 29 April 2012).

consequences be connected through an uninterrupted causal chain of events'.[39] This approach has serious implications for the ability to apply the DPH paradigm to CW in light of the probable reality that most cyber attacks will be indirect in effect. Consider the following description:

> 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 (eg, in response to an attack that does X, the [target] network's 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.[40]

The implications of this aspect of DPH for CW appear uncertain. On the one hand, the Interpretive Guidance indicates that indirect effect is not enough and that the harm resulting from an act of DPH must be *objectively likely* ('harm which may reasonably be expected to result from an act in the prevailing circumstances').[41] On the other hand, it seems that both intended and unintended consequences of cyber-actions are likely to occur over several causal steps; for instance, a cyber attack can be routed to its target system through an intermediate, compromised computer or network[42] (something which incidentally could lead to a revival of reference to the traditional law of neutrality, largely considered redundant since 1945). In these circumstances, it appears doubtful that CW could ever meet the requirement of direct causation for DPH, which suggests that civilians could engage in CW with impunity.

The Interpretive Guidance does not specifically comment on CW in respect of the third constitutive element of DPH, belligerent nexus. However, if an act needs to be 'specifically designed' to cause harm directly to the detriment of a party to the conflict, the implications of the indirect consequences of CW and the requirement of reasonable foreseeability of the harm, noted above, need to be considered. If the indirect consequences are in fact the intention behind the attack, then the resulting harm would be objectively likely and would have

[39] Interpretive Guidance (n 28) 1021–22.
[40] WA Owens, KW Dam and HS Lin (eds), *Technology, Policy, Law, and Ethics Regarding US Acquisition and Use of Cyberattack Capabilities* (The National Academies Press 2009) ('NRC Report') 127.
[41] Interpretive Guidance (n 28) 1017.
[42] NRC Report (n 40) 268–70.

the requisite belligerent nexus, but it would fail the direct causation test. If the indirect consequences were neither intended nor foreseen, the harm might still be objectively likely but belligerent nexus, as well as direct causation, would be lacking. On the other hand, belligerent nexus could become relevant at an earlier stage in the CW process—for example, if a program is being designed and written specifically to disable certain weapons systems in the target State. It may seem illogical, but belligerent nexus would be more likely to apply before the hostile act is actually committed than during or after its commission.

D. What is the Specific Relevance of DPH in the Context of CW?

As with other aspects of this topic, considerable uncertainty surrounds the identity and legal status of the individuals involved in CW, as well as a fundamental question about the nature of the systems which they operate. The persons involved fall, it is suggested, into three main categories according to their functions:

1. those who design and write the programs used for offensive or defensive CW operations;
2. those who install these programs on the computer systems, act as service administrators ('webmasters') and provide technical maintenance for them; and
3. those who actually operate the computer programs in a CW scenario.

Any of these personnel could, conceivably, be actual military personnel. In 2009, the US Secretary of Defence ordered the establishment of United States Cyber Command (USCYBERCOM), with the following mission statement:

> USCYBERCOM plans, coordinates, integrates, synchronizes, and conducts activities to: direct the operations and defense of specified Department of Defense information networks and; prepare to, and when directed, conduct full-spectrum military cyberspace operations in order to enable actions in all domains, ensure US/Allied freedom of action in cyberspace and deny the same to our adversaries.[43]

USCYBERCOM became operational on 21 May 2010. It is subordinate to US Strategic Command, is located in the military facility at Fort Meade, Maryland, and is headed by a senior US Army officer, General Keith B Alexander.[44] The

[43] US Department of Defense, 'US Cyber Command Fact Sheet' (25 May 2010) <http://www.defense.gov/home/features/2010/0410_cybersec/docs/CYberFactSheet%20UPDATED%20replaces%20May%2021%20Fact%20Sheet.pdf> (accessed 5 April 2012).

[44] ibid.

little information that is available about USCYBERCOM in the public domain suggests that it is a primarily military, or at least heavily *militarized*, organization. A report published by the US National Research Council in the same year that USCYBERCOM's creation was mandated suggested that

> the systems used to launch cyberattacks are legitimate military targets, and civilians who qualify for the narrow category of 'civilians accompanying the armed forces' (presumably those who operate and maintain those systems)—even if they do not actually press the button that launches a cyberattack—are both eligible for prisoner-of-war status and also legitimate military targets for the enemy.[45]

To the extent that any personnel engaged in CW within the context of an armed conflict are actually card-carrying members of the armed forces or of militias forming part of the armed forces, as defined in the Hague Regulations 1907[46] and Geneva Convention III (1949),[47] they are clearly combatants under the LOAC and the paradigm of DPH is not relevant.

It would be equally inapplicable if such persons are considered to be spies or saboteurs (if members of the armed forces); thus, for example, uniformed military personnel involved in gaining access to adversary computer systems for the purpose of obtaining information, or those who are involved in those systems' *physical* destruction (eg by placing explosives in a computer laboratory), will be considered combatants. The archaic approach to espionage in the Hague Regulations, which requires the spy's presence in 'the zone of operations of the hostile army' while wearing uniform,[48] virtually ensures that this particular rubric will not be applicable in CW. Physical sabotage is more likely to be relevant to CW, although the development of technology enabling the non-kinetic disablement of computer systems from a great distance will conceivably render it redundant. The same rule under LOAC applies in relation to

[45] NRC Report (n 40) 266, fn 25.

[46] The armed forces as such are not defined, but 'militia and volunteer corps' fulfilling the conditions of being under responsible command, having a fixed distinctive emblem recognizable at a distance, carrying arms openly, and conducting operations in accordance with the LOAC, are considered equally to be combatants: Art 1, Annex to The Hague Convention (IV), Regulations Respecting the Laws and Customs of War on Land (18 October 1907) 2 AJIL Supp 90.

[47] Prisoners of war are defined as, *inter alia*, members of the armed forces or of militias or volunteer corps forming part of the armed forces, and members of other militias (including organized resistance movements) that satisfy the requirements of Art 1 of The Hague Regulations: Art 4(A)(1)–(2), Geneva Convention (III) Relative to the Treatment of Prisoners of War (opened for signature 12 August 1949, entered into force 21 October 1950) 47 AJIL Supp 119 ('GC III').

[48] Art 29, The Hague Regulations. Rather confusingly, persons satisfying these requirements 'are not considered spies'; it would be more correct to say that they are spies by conduct, but as combatants by legal status, who are not engaged in an illegal activity under the LOAC, they cannot be punished under that body of law. See also Art 46, AP I.

sabotage as in cases of espionage: to benefit from combatant and POW status, the saboteur must be a uniformed member of the armed forces at the point of capture.[49]

There is some ambivalence on the subject of 'civilian scientists and weapons experts', in that these are viewed as *indirectly* participating only, and are conceded to be normally entitled to protection from direct attack, save that

> some doubts were expressed [in the Group of Experts on DPH] as to whether this assessment could be upheld in extreme situations, namely where the expertise of a particular civilian was of very exceptional and potentially decisive value for the outcome of an armed conflict, such as the case of nuclear weapons experts during the Second World War.[50]

The significance of this is that, first, the computer programmers who research and write attack programs could be likened to 'scientists' in the Interpretive Guidance's parlance, while operators who execute cyber attacks could certainly be considered 'weapons experts', although one may wonder how credible this would be if all that would be required to launch a cyber attack would be the pressing of a button on the computer keyboard. Secondly, it is quite conceivable that a cyber-strike against a conventionally more powerful adversary could herald a decisive turning point in a conflict (for example, if a virus incapacitates the network controlling the launch of weapons systems such as missiles), in which case the participation of such civilians would shift from indirect to direct.

The point has been well made that the armed forces traditionally emphasize

> skills such as marksmanship, physical strength, and the ability to jump out of airplanes and lead combat units under enemy fire. Accolades are heaped upon those who excel in these areas. Unfortunately, these skills are irrelevant in cyberwarfare . . . Absent [from the armed forces] is recognition for technical expertise.[51]

In the same vein, Schmitt notes that, 'Some technologies, such as computer network or space operations, require education that the average member of the armed forces lacks.'[52] The trend towards increasing civilianization of certain

[49] *Ex parte Quirin* [1942] 317 US 1. The case uses the confusing terminology of 'unlawful combatants' to describe saboteurs captured out of uniform (even if they are members of the armed forces), which is a conflation of conduct and status under the law and as such is best avoided.

[50] Interpretive Guidance (n 28) 1021 at fn 122.

[51] Lt Col G Conti and Col J 'Buck' Surdu, 'Army, Navy, Air Force, and Cyber—Is it Time for a Cyberwarfare Branch of Military?' (2009) 12 Information Assurance Newsletter 14 at 16 <http://www.rumint.org/gregconti/publications/2009_IAN_12-1_conti-surdu.pdf> (accessed 5 April 2012). The authors argue for the creation of a specific branch of the *armed forces* dedicated to CW.

[52] Schmitt (n 6) 28.

functions in the armed forces has also been emphasized: 'Today, there is a growing dependence of the modern military on civilians and civilian-provided services and expertise that blurs traditional distinctions between military and civilian activity and personnel'.[53] The maintenance of these 'traditional distinctions' is of the utmost importance to the effective application of IHL in hostilities:

> Arguably, the armed forces should establish separate networks for targets the enemy would find especially attractive in order to minimize the risk of collateral damage or incidental injury. Similarly, it might be argued that the military should avoid using dual-use assets, such as air traffic management systems, that are particularly vulnerable to computer attack. The reality, however, is that the trend is in precisely the opposite direction, as most militaries seek to save money by outsourcing functions performed traditionally by the military and purchasing 'off-the-shelf' equipment and services. The extensive use of civilian internet services and commercial software is illustrative.[54]

In light of these trends, there is a high probability that many, if not most, of the personnel substantively involved in cyber operations may actually be civilians;[55] it is in these circumstances that the paradigm of DPH is engaged.

E. Possible Classifications of CW Personnel Excluding the DPH Paradigm

Before proceeding to an analysis of how the notion of DPH might apply to civilians engaged in CW, however, it is worth briefly considering three other possible LOAC characterizations of personnel engaged in CW:

1. Computer technicians, technical maintenance personnel and others who perform similar tasks could be assimilated to 'supply contractors [or] members of labour units' under GC III, as suggested by the earlier quotation from the *NRC Report*.[56] They could thus qualify for inclusion in the limited category of 'persons who accompany the armed forces without actually being members thereof', which means that, 'provided

[53] NRC Report (n 40) 266–67.

[54] MN Schmitt, HA Harrison Dinniss and TC Wingfield, 'Computers and War: The Legal Battlespace' Background Paper prepared for Informal High-Level Expert Meeting on Current Challenges to International Humanitarian Law (Cambridge 25–27 June 2004) 10 <http://www.hpcrresearch.org/sites/default/files/publications/schmittetal.pdf> (accessed 18 April 2012).

[55] Civilians as such are not defined in IHL. Instead, they are subject to a negative definition, inasmuch as anyone who is not a combatant is a civilian: Art 50(1), AP I.

[56] See n 45.

that they have received authorization from the armed forces which they accompany', they would qualify for POW status if captured.[57] But they would not be combatants and therefore could not lawfully attack others or be attacked. It has been suggested, almost certainly correctly, that this category would apply only to 'persons . . . more analogous to computer technicians that keep the machines in order, and not ones that actually undertake . . . attacks'.[58] Nevertheless, it could reasonably cover persons contracted to provide such services to defence ministries or armed forces, by analogy with private military/security companies.

2. So-called 'patriotic hackers' or 'hacktivists' (that is, people who are not part of their State's armed forces but on their own initiative carry out attacks against perceived 'enemy' computer systems, without the authority and outside the control of their government but in pursuance of common political ends)[59] could arguably be assimilated to the rarely used category of *levée en masse*, in which case they would be entitled to POW status if captured, and could also be subject to attack while acting as such a body. However, in order to meet the legal conditions of a *levée en masse* they would have to 'take up arms' spontaneously 'on the approach of the enemy'; they would need to lack organization and 'carry arms openly'.[60] It is rather difficult to see precisely how these criteria could be applied in a cyber-conflict. Conceivably the taking up of arms could be loosely compared with the execution of CW commands against hostile systems, but how would they 'carry arms openly' in this context? Indeed, what would the 'arms' in question be—laptop computers, perhaps? If on the other hand the 'arms' are considered to be the software that executes the cyber attacks, how can they by nature ever be deemed to be carried 'openly'? What would constitute 'the approach of the enemy' in CW? Even if it is accepted that it could be the initiation of a cyber-attack, the window within which a *levée en masse* could then legitimately constitute itself (given that it is by definition a very temporary status)[61] would be impossibly small: a matter of minutes at most, probably far less. Finally, 'hacktivists' generally—far from not having time to organize themselves—tend to be very well organized; indeed, it is the very concentration and intensity of their attacks that usually make them so effective.

[57] Art 4(A)(4), GC III.

[58] L Doswald-Beck, 'Some Thoughts on Computer Network Attack and the International Law of Armed Conflict' (2002) 76 Intl L Stud 163, 172.

[59] NRC Report (n 40) 276. For specific examples of such 'hacktivism', see ibid 278–79. In particular, it is suggested that the Russians who hacked into Estonian and Georgian websites in 2007–08 were 'hacktivists', as the Russian Government denied that they acted on its instructions.

[60] Art 4(A)(6), GC III.

[61] See Y Dinstein, *The Conduct of Hostilities under the Law of International Armed Conflict* (CUP 2010) 48.

3. In certain very limited circumstances, contracted technical personnel might notionally be assimilated to mercenaries, in which case they would not be combatants and would not benefit from POW status in the event of capture. In order for this to be the case, they would have to satisfy the cumulative criteria of (a) being specially recruited to take part in an armed conflict, (b) actually taking direct part in hostilities, (c) being motivated by the desire for private gain and receiving material compensation substantially in excess of that paid to members of the armed forces, (d) being a non-national and non-resident of a State party to the conflict, (e) not being a member of the armed forces of a State party to the conflict and (f) not representing another State as a member of its armed forces.[62] Although it would not be impossible to imagine a technology-poor State, or one with a low level of technical education, recruiting computer technicians from abroad to engage in CW on its behalf, the ease with which the most complex computer technology spreads around the world, and the ubiquity of information technology courses in further and higher education internationally, suggests that this scenario would be uncommon. Note that the reference to DPH in criterion (b) is not meant in the same sense as DPH under Article 51(3) of the Protocol, since the latter refers to civilians, whereas mercenaries are considered to be essentially unprivileged combatants;[63] the recruitment of such foreign technical specialists would therefore have to be with a view to their incorporation into the State armed forces, which seems highly unlikely.

3. Specific CW Activities as DPH

Ultimately, it is submitted that the best way to 'visualize' the concept of DPH in the context of CW lies in the tabulation of a spectrum of possible or likely CW activities that might be undertaken by civilian technical experts, so that their systematic compliance or otherwise with the constitutive elements of DPH can be presented. The table below offers a suggested (though illustrative and non-exhaustive) paradigm to that effect.[64]

[62] Art 47, AP I.

[63] The ICRC *Commentary* to the Protocol specifies that, '... this condition excludes foreign advisers and military technicians, who are found in numerous countries nowadays ... The increasingly perfected character of modern weapons ... requires the presence of such specialists ... As long as these experts do not take any direct part in the hostilities, they are neither combatants nor mercenaries, but civilians who do not participate in combat' (n 27) para 1806.

[64] An original version of this proposed spectrum of activities (without the specific DPH analysis) appears in the NRC Report (n 40) at 268. Four of the activities (nos 5–7 and 9) on the present spectrum were the original examples used in that source; the other six were added by the present author.

CW activity undertaken by a civilian	Threshold of harm?	Direct causation?	Belligerent nexus?	DPH?
1. Research for the development of CW programs generally	No—the research is *in abstracto*	No—no harm is actually caused	No—the research is not tied to any particular conflict	No
2. Design/writing of a specific CW program	Yes, if the program is designed to cause the harm specified	No—any eventual harm that might result is too remote	Yes, potentially, if research takes place with a specific future target or conflict in mind	No
3. Installation of a CW program on a computer system	Yes—the program cannot be used to cause the harm unless it is installed	No—any eventual harm is too remote from the installation	Yes, if imminent use is intended	No
4. Provision of regular/routine operational maintenance for the CW-equipped system	No—any harm is too remote from mere maintenance	No—routine maintenance does not in itself cause any direct harm	No—system would require routine maintenance irrespective of its use in conflict	No
5. Identification of a vulnerability on a system in a target State	No—in itself would not cause any harm	No—other positive action would still be required to exploit the vulnerability	Yes, if imminent exploitation is intended	No
6. Posting of a vulnerability notice for a system in a target State (which a cyber attack conducted by other persons can subsequently exploit)	Yes, if harm would not have occurred but for the posting of the notice	No—one step removed from the action that would directly cause the harm	Yes, if intention was specifically to enable CNA to occur	No
7. Exploitation of a vulnerability on a target State system by introduction of a hostile agent that does not damage it immediately but that can be directed to cause damage subsequently	Yes—introduction of the hostile agent is what eventually causes the harm; time lapse irrelevant	No, if separate autonomous action is required to activate the agent; yes, if activation is pre-programmed by the same person; time lapse relevant	Yes—intention is clearly hostile	Yes or No
8. Exploitation of a vulnerability on a target State system by introduction of a hostile agent that damages it directly	Yes—introduction of the hostile agent is what causes the harm	Yes—there is no intermediary between introduction of the agent and its activation	Yes—intention is clearly hostile	Yes
9. Dictation or written provision, to a combatant, of the precise set of commands needed to activate the hostile agent	Yes—harm would not occur but for provision of the commands	Yes—activation is caused directly by the input of the commands	Yes—intention is clearly hostile	Yes
10. Personal entry, by the civilian, of the precise set of commands needed to activate the hostile agent	Yes—activation of the agent causes the harm	Yes—activation is caused directly by the input of the commands	Yes—intention is clearly hostile	Yes

4. Conclusions

The final page on cyber-conflicts and their legal regulation is not yet written; indeed, arguably the first page is barely complete. In the absence of any visible appetite for a new IHL treaty to regulate the conduct of military hostilities in cyberspace, however, it is clear that the *lex lata* must be adapted and applied, insofar as it is possible to do so. In this spirit, there are certain observations that may be made about the application of the DPH paradigm to CW activities performed by civilians, as tabulated above, and the implications thereof.

1. Because the three constitutive elements of DPH are cumulative, the threshold for reaching all three is relatively high; thus, only 3 out of the 10 suggested activities on the spectrum are *unequivocally* classifiable as DPH. One might conclude from this that the application of DPH to CW will be rather difficult *in concreto*. Nevertheless, one or two individual constitutive elements of DPH are satisfied in all but two of the activities—general research and routine maintenance.

2. The easiest of the constitutive elements to satisfy is belligerent nexus, which is unsurprising since CW is such a specialized activity that almost any act it naturally entails would be intended to support an intended CNA or CNE.

3. Equally unsurprisingly, the hardest to satisfy is direct causation. The number of steps in the typical chain of causation in CW, along with the high probability of unintended consequences, makes this inevitable.

4. The most variable element, and therefore in practice likely to be the crucial one, is the threshold of harm. There is a fine line between cyber attacks that cause inconvenience on a massive scale and those that actually lead directly to death or destruction or have an adverse effect on *military* operations. This is a consequence of the high degree of integration and interoperation between civilian networks and actual or potential military effects.

5. The last point just made, in conjunction with the current military doctrinal emphasis on effects-based targeting,[65] suggests that in practice most public computer systems that might be targeted in cyber-conflicts are best viewed as dual-use objects; that is, objects that are primarily civilian in character but the actual or potential uses of which in wartime can convert them into legitimate military objectives.[66] Whether it will be

[65] See PM Carpenter and WF Andrews, 'Effects-based Operations: Combat Proven' (2009) 52 Joint Force Q 78; Major RB Herndon and others, 'Effects-Based Operations in Afghanistan' (2004) Field Artillery 26.

[66] International Criminal Tribunal for the Former Yugoslavia, 'Final Report to the Prosecutor by the Committee Established to Review the NATO Bombing Campaign Against the Federal Republic of Yugoslavia' (2000) 39 ILM 1257, 1266–67.

lawful to target them will then depend on whether they amount to, 'objects which by their nature, location, purpose or use make an effective contribution to military action and whose total or partial destruction, capture or neutralization, in the circumstances ruling at the time, offers a definite military advantage'.[67]

6. A final, collateral point may be made in this context about whether the premises or installations from which CW are conducted will be considered military objectives, and the consequences of that determination for any civilian personnel working there. Clearly the answer on the first point is affirmative if the premises or installations are actually military by nature; USCYBERCOM, based as it is in a US military facility, would clearly be subject to entirely lawful attack by an adversary in any cyber-conflict. In such situations, civilian technical staff killed or wounded would (depending on their function) either be DPH—that is, legitimate military targets in themselves—or collateral damage.[68] Civilian objects from which cyber-operations are conducted could be considered military objectives by use, although the fact that viruses can be uploaded and activated on hostile systems from a laptop means that conceivably any location at all could be targeted.

[67] Art 52(2), AP I.
[68] Art 51(5)(b), AP I.