The Old Binding the New:
Can a cyber-attack be conducted in conformity with the principle of distinction?

Photo: Huffington Post

Master's thesis International and European Public Law
Author: Gino Silanoe LLB

Supervisor: Dr. M.E.A. Goodwin
Second Reader: Dr. A.K. Meijknecht

June 2014

Tilburg University
Law School

The Old Binding the New:
Can a cyber-attack be conducted in conformity with the principle of distinction in international armed conflict?

Table of Contents

List of abbreviations:

Geneva Conventions I, II, III, IV 1949 or GC/I, GC/II, GC/III, GC/IV	*Convention (I) for the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field, 12 August 1949. Convention (II) for the Amelioration of the Condition of Wounded, Sick and Shipwrecked Members of Armed Forces at Sea, 12 August 1949. Convention (III) relative to the Treatment of Prisoners of War, 12 August 1949. Convention (IV) relative to the Protection of Civilian Persons in Time of War, 12 August 1949.*
Additional Protocol I 1977 or AP/I	*Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts. (Protocol I), 8 June 1977.*
Additional Protocol II 1977 or AP/II	*Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of Non-International Armed Conflicts (Protocol II), 8 June 1977.*
Additional Protocol III 2005 or AP/III	*Protocol additional to the Geneva Conventions of 12 August 1949, and relating to the Adoption of an Additional Distinctive Emblem (Protocol III), 8 December 2005.*
1868 St. Petersburg Declaration	*Declaration Renouncing the Use, in Time of War, of Explosive Projectiles Under 400 Grammes Weight.*
IHL	*International Humanitarian Law*
UN Charter	*Charter of the United Nations*
ICJ	*International Court of Justice*
ICRC	*International Committee for the Red Cross*
ICS	*Industrial Control Systems*
LARs	*Lethal Autonomous Robots*
UCAVs	*Unmanned Combat Aerial Vehicles*

Introduction.

Means and methods of warfare continuously evolve. Nation-States seek to gain the upper hand when engaging the adversary. They are in search of the 'silver bullet'. Choosing the right means of warfare has to be accompanied by the correct methods of engaging the adversary. Parties to a conflict both have strengths and weaknesses. The ideology of exploiting an enemy's weaknesses is not new. Philosopher Sun Tzu and his book called the Art of War written in approximately 500 BC already elaborated on how to make optimal use of the enemy's lesser developed strengths.[1] Capitalizing on the adversary's weaknesses can still be done today to a certain extent. There is no unlimited choice of means and methods of warfare however. The permissible conduct in international armed conflict is regulated by international humanitarian law or *jus in bello*. One of the key principles in international humanitarian law is that of distinction and it can be found under article 48 of the Additional Protocol I of the Geneva Conventions. In international armed conflict, combatants or military targets are allowable objects of attack whereas civilians or civilian objects should be spared. Some objects are used in both military and civilian context alike. A prime example of such an object is that of networked systems. Current day civil society next to the military forces rely on the use of networked systems.[2] A nation's entire power supply is ultimately coordinated by the use of such systems. Nation-States are familiar with the heavy reliance on networked systems. China issued a statement indicating that it sees current military reliance of nations on networked systems as a weakness and mentioned that it will exploit that weakness in times of warfare.[3] The type of warfare primarily directed against networked systems is called cyber warfare. A definition in the laws of armed conflict is however absent. Cyber warfare differs from conventional means and methods of warfare on a multitude of levels. An attack in that context and central to this thesis is the cyber-attack. One of the characteristics of a cyber-attack is that it is non-kinetic by nature: there is no physical force applied onto the target itself. An attacker can be physically present in an entire different continent while conducting a cyber-attack at nearly the same time in another. An attack may

[1] Tzu 2002, pp. 1-101.
[2] Dijkhoff 2010, p. 202.
[3] D. Lague, 'Chinese see military dependence on computers as weakness', *New York Times* August 29 2007, http://www.nytimes.com/2007/08/29/world/asia/29iht-cyber.1.7299952.html?pagewanted=all&_r=0

well cross geographical boundaries without having to engage a nation-states' traditional defences. Cyber-attacks transcend geographical boundaries, therefore cyber-attacks pose a challenge to the doctrine of territorial jurisdiction. Moreover, a cyber-attack can be conducted without having to reveal one's identity, the threat of being retaliated afterwards is therefore virtually non-existent. Even though the means and methods used in cyber-warfare differ from conventional methods of warfare, the actors waging cyber-warfare are bound by international humanitarian law and its principles. One of the most fundamental notions of international humanitarian law is the principle of distinction. Means and methods utilized in armed conflicts are bound by this norm. A cyber-attack has not instigated an armed conflict yet, however there have been instances of non-kinetic cyber-attacks taking place in international armed conflicts in reaction to kinetic armed attacks.[4] In those cases there have been reasonable suspicions pointing towards particular Nation-States regardless of the private group carrying out the attacks and implicitly or explicitly claiming to be responsible. An example of this is the suspected involvement of the Russian government in the cyber-attacks during the war between Russia and Georgia. [5] Responsibility for the attacks were claimed by a private group of hackers called 'the Kremlin Kids'. In the summer of 2008, Georgia could not communicate with the outside world because internet was not functioning, meanwhile Russian troops invaded Ossetia.[6] Examples of attacks outside the existence of an armed conflict that are suspected to be either conducted by a State or at least had some sort of State support are numerous. Two of the best known examples are arguably the Stuxnet worm attack on the power plant in Natanz and the attack occurring in Estonia after the removal of a Russian war monument.[7] In case a Nation-State conducts an attack in times of armed conflict the principle of distinction has to be respected. The research question of this thesis is therefore going to be:

'Can a cyber-attack be conducted in conformity with the principle of distinction in international armed conflict?'

[4] Hathaway 2012, p.1.
[5] J. Carr, 'Real Cyber Warfare: Carr's Top Five Picks', *Forbes* April 2 2011 http://www.forbes.com/sites/jeffreycarr/2011/02/04/real-cyber-warfare-carrs-top-five-picks/
[6] Hathaway 2012, p.23.
[7] O'Connell 2012, pp. 188-191.

This thesis will be outlined the following way. Chapter 2 outlines why the laws of armed conflict apply in case a cyber-attack occurs in times of armed conflict. Concisely addressed are the difficulties in relation to identifying the attacker and the difficulties that cyber-attacks pose to the addendum of state responsibility and attribution of cyber-attacks. Chapter 3 will outline the legal framework of this thesis discussing the substance of the principle of distinction. That chapter will elaborate on who or what the object of an attack can be and what modalities govern a possible loss of protection from attack. Chapter 4 will define the terms cyber-attack and cyber warfare. Chapter 5 will discuss a various means of cyber warfare. This chapter will review means of cyber warfare by addressing their specific characteristics. Chapter 6 revolves around the main question whether a cyber-attack, can be conducted in conformity with the principle of distinction in times of international armed conflict. Two cases of viruses/worms will be central in determining the answer to this question. Finally, a conclusion will be reached. Definitions of cyber related notions are often non-existent in international law. In order to come to a coherent conclusion, definitions have to be established first. And although the utmost care is taken in determining a definition or categorization to be as objective as possible, some level of subjectivity is inevitable because conscious choices are inherently made with a level of subjectivity. Through motivating extensively it is aimed to keep the level of subjectivity to a minimum. This thesis departs from the existence of an armed conflict. Accountability for a cyber-attack also called attribution is a topic of academic interest and although it is an interesting phenomenon and it arguably poses new challenges to the addendum of accountability for an attack, it will not be discussed in this thesis. Neither will the challenges in determining when a cyber-attack amounts to an armed attack and whether it justifies an attack of self-defence be reviewed in this thesis.

Chapter 1. International humanitarian law applying to cyber-attacks.

The principle of distinction forms the legal framework of this research. It is a part of international humanitarian law and thus in order for it to be applicable, international humanitarian law has to govern the use of cyber-attacks when conducted in international armed conflict. This chapter motivates why international humanitarian law applies in that case and therefore the principle of distinction, establishing the relevant framework for this thesis.

§1.1 International humanitarian law.

This subsection determines why international humanitarian law applies in case a cyber-attack is conducted as a part of cyber warfare in an international armed conflict. Cyber-attacks as a means of warfare are regarded by some authors to be a novelty.[8] Cyber-attacks have not been possible for as long as a range of other means of warfare. The existing framework of provisions of international humanitarian law was setup before the possibility of conducting a cyber-attack presented itself. The Geneva Conventions were drafted in the wake of World War II and the drafters could not have foreseen that attacks could and would be carried out over a computer network.[9] This discrepancy can possibly lead to a cyber-attack falling out of the scope of the international humanitarian law doctrine. In determining the relevant framework applying to a cyber-attack in an international armed conflict, the most dominant view is that the existing framework of international law applies to a cyber-attack conducted in international armed conflict. The strongest argument for this view is based on an advisory opinion of the International Court Justice in the Nuclear Weapons case.[10]. The International Court of Justice stresses in the advisory opinion that provisions of international humanitarian law apply regardless of what weapon is used. An indicator supporting this argument is that the Geneva Conventions do not specify weapons which weapons are regulated by the Conventions and which are not.[11] Provisions apply

[8] *See*: Schmitt et al. 2013, p. 68. *See also*: Dipert 2010, p. 385. *See also:* O. Rochford, 'Putting Cyber Warfare Into Perspective', *Security Week* November 19 2012 http://www.securityweek.com/putting-cyber-warfare-perspective.
[9] Hathaway 2012, p. 26.
[10] ICJ, *Nuclear Weapons case*, Advisory Opinion, 8 July 1996, ICJ Reports 1996.
[11] *Available at*: http://www.icj-cij.org/docket/files/95/7497.pdf paras. 37-50.

therefore to every possible use of force regardless of the type of weapon being employed.[12] This view is also expressed in the Tallinn Manual in an effort to define the relevant framework of law applying to cyber warfare.[13] The International Group of Experts of the Tallinn regarded the pre-existing provisions of international law such as treaties and customary international law to be applicable to cyber-attacks. They published their findings in the Tallinn Manual and concluded herein that although cyber warfare is viewed by some as being a non-traditional means of warfare, the existing framework of international law, more in specific, international humanitarian law or *jus in bello* applies when it is conducted in international armed conflict. The International Group of Experts considered that hostilities presuppose the collective application of rules that govern the means and methods of warfare and therefore existing provision of international law are applicable regardless of cyber warfare possibly being a novelty. The International Group of Experts note that similar to other international law related cases, the applicability of particular treaty rules is determined by matters such as whether a State is a Party to the Treaty in question, its status as a party to the conflict and the type of armed conflict present (international or non-international). Next to the opinion of the International Group of Experts there are other indicators pointing out that cyber warfare is in fact governed by the pre-existing rules of international humanitarian law. Schmitt agrees to the notion of international law applying in case a cyber-attack is conducted in the context of an armed conflict. He refers to the legal obligation that Nation-States have to review weapons before they use them in armed conflict. This principle can be found in Article 22 of the 1907 Hague Regulations Respecting the Laws and Customs of War on Land, and Article 35(1) of the Additional Protocol I. If international law would not apply, there would not be any reason to have this requirement since in that case new weapons would not be subjected to this examination unless they would explicitly be mentioned to fall under its regime which renders this requirement of international law obsolete.[14] It

[12] *Ibid*.

[13] A recent development in the field of cyber warfare is a document called: the Tallinn Manual. A group of 20 independent experts have documented their findings at the invitation of NATO's Cooperative Cyber Defence Centre of Excellence. They gathered and debated on the applicability of international law on cyber warfare and their views are expressed in the Tallinn Manual. It is explicitly noted that the Tallinn Manual does not represent the views of NATO nor any other NATO allied institution. Although the Tallinn Manual is not a legal document, it helps educate, understand and clarify matters in the field of international law in relation to applicability of cyber-attacks and it might possibly shape future lawmaking by influencing on different levels of policy making.

[14] Schmitt 2013, pp. 176-177.

seems that a significant amount of authors agree that the laws of armed conflict apply to a cyber-attack however, there are also authors that advocate for provisions to be added to the current framework of the existing laws of armed conflict regulating cyber-attacks.[15]

§1.2 The threshold of an international armed conflict.

In this subsection, the requirements of an international armed conflict are being reviewed in a concise manner due to the presumption in this thesis of an international armed conflict being present. For the sake of completeness, the challenges surrounding the term armed conflict will be stipulated in this subsection. The research conducted in this thesis connects the principle of distinction in an international armed conflict to cyber-attacks and takes the international armed conflict as the point of departure. In order for the principle of distinction to be applicable it is necessary that the conflict one is dealing with is armed and the parties opposing one another are States. The absence of a definition of an armed conflict in international law is accompanied by the lack of requirements that establish an armed conflict. The context these challenges are presented in, is that in order for the principle of distinction to be applicable international humanitarian law has to be applicable. International humanitarian law applies as soon as the status of the conflict reaches that of an armed conflict. Common article 2 of the 1949 Geneva Conventions explicitly mentions that the provisions of the Geneva Conventions apply in case of an armed conflict by noting that:

> 'The present Convention shall apply to all cases of declared war or of any other armed conflict which may arise between two or more of the High Contracting Parties even if the state of war is not recognized by one of them. The Convention shall also apply to all cases of partial or total occupation of the territory of a High Contracting Party, even if the said occupation meet with no armed resistance.'

The central notion revolves around the notion of 'armed'. Armed as in an armed conflict, as mentioned earlier, is not specified in international law. In the Tallinn Manual the experts stress that there is no requirement of a declaration of war to be met and neither is recognition of the situation by the involved parties a requirement in order for an armed conflict to be established. Hostilities have to be present in the conflict, however.[16]An assessment has to be made based on factual circumstances on a case-by-case

[15] O'Connell 2012 p.189
[16] Schmitt et al. 2013, p.68.

approach. The ICTY in the *Tadic* case stated that: 'an armed conflict exists whenever there is a resort to armed force between States.' The ICTY further reasons and confirms that: 'International humanitarian law applies from the initiation of such armed conflicts and extends beyond the cessation of hostilities until a general conclusion of peace is reached; or, in the case of internal conflicts, a peaceful settlement is achieved. Until that moment, international humanitarian law continues to apply in the whole territory of the warring States or, in the case of internal conflicts, the whole territory under the control of a party, whether or not actual combat takes place there.'[17] Schindler reasons that: 'the existence of an armed conflict within the meaning of Article 2 common to the Geneva Conventions can always be assumed when parts of the armed forces of two States clash with each other. Any kind of use of arms between two States brings the Conventions into effect.'[18] Gasser indicates that 'any use of armed force by one State against the territory of another, triggers the applicability of the Geneva Conventions between the two States. […] It is also of no concern whether or not the party attacked resists.'[19] The International Committee of the Red Cross proposed the definition of an international armed conflict mentioning that: 'an international armed conflicts exists whenever there is a resort to armed force between two or more States.'[20] It notes that an armed international conflict exists when States resort to armed force regardless of the reasons or the intensity of this confrontation. Dominantly based on the view of the ICTY the outtake for this research is that there has to be hostilities present to establish an armed conflict and the parties conducting the hostilities need to be State parties thereby establishing an international armed conflict. These elements seem to be the common denominators in the findings of the ICTY, the theories of the authors, the opinion of International Group of Experts and the opinion of the ICRC.

[17] ICTY, *The Prosecutor v. Dusko Tadic*, Decision on the Defence Motion for Interlocutory Appeal on Jurisdiction, IT-94-1-A, 2 October 1995, para. 70.
[18] Shindler 1979, p. 131.
[19] Gasser 1993, pp. 510-511.
[20] ICRC, 'How is the Term "Armed Conflict" Defined in International Humanitarian Law?', Opinion Paper March 2008 p.5 http://www.icrc.org/eng/assets/files/other/opinion-paper-armed-conflict.pdf

§1.3 Challenges identifying the attacker.

In this subsection the challenges faced with establishing the identity of the perpetrator of the cyber-attack will be addressed. In International Law the use of force is regulated by treaty law and customary law. According to article 2 under 4 of the UN Charter which reads:

> 'Nothing in the present Charter shall impair the inherent right of individual or collective self-defence if an armed attack occurs against a Member of the United Nations, until the Security Council has taken measures necessary to maintain international peace and security. Measures taken by Members in the exercise of this right of self-defence shall be immediately reported to the Security Council and shall not in any way affect the authority and responsibility of the Security Council under the present Charter to take at any time such action as it deems necessary in order to maintain or restore international peace and security.'

There are two main exceptions to this article. The first is a (collective) armed enforcement action authorized by the Security Council, the second is perhaps more relevant for this thesis which is that of self-defence.[2122] The provision of self-defence can be found under Article 51 of the UN Charter which reads:

> 'Members shall refrain in their international relations from the threat or use of force against the territorial integrity or political independence of any state, or in any other manner inconsistent with the purposes of the United Nations.'

The right of self-defence is regarded to be customary international law being referred to by the ICJ in the *Nicaragua Case*.[23] If a state wants to react to a cyber-attack by using force by way of self-defence it first needs know who that force needs to be directed against. In the context of a cyber-attack this means that a state has to discover the attack first. The Stuxnet case reviewed in Chapter six points out that discovering a cyber-attack can be a challenge because of characteristics and built-in precautions preventing the cyber-attack to be discovered. The chapters 5 and 6 mention examples of those characteristics and precautions. These will show that the measures taken by the initiating party can be

[21] U.N. Charter art. 39.
[22] U.N. Charter art 51
[23] ICJ 27 June 1986, 14, *Nicaragua v. United States of America*.

extensive. In those cases discovering a cyber-attack might take months or years. A cyber-attack is often conducted by using a computer network, a series of computer systems that are connected together, but lack a body or institution controlling and regulating that dimension that could be of assistance in identifying who an attacker is. Next to this, a cyber-attack can be disguised which makes it even harder to establish who is behind it. It can be used to make it seem that another state is behind it. Also, objects can be used without the owner's knowledge or merely used to facilitate a cyber-attack. A counterattack in those cases could be deemed unjustifiable since a *conditio sine qua non* between the act or omission of a state's breach of its obligation will be absent. This relation is what a nation-state has to prove before a response with armed force may be issued based on international law. The chance that there is direct incriminating and identifying material linking the cyber-attack to the perpetrator is small.[24] In the few cases it does, there is a great chance that it will link a non-state actor to a cyber-attack and not a nation-state. The highest certainty that a state is behind an attack would be if a state itself claims or announces to be responsible for the cyber-attack, which at this moment, seems to be implausible based on earlier cases in which there are strong indicators pointing out state involvement in cyber-attacks. Based on those instances the trend seems to be that the state in those cases still disavows all involvement.[25] It is worth mentioning that the other way a state will be held responsible for an attack is that based on attribution of a an act carried out by a state official or a state organ or in case a non-state actor or group conducts the attack. The importance of cyber-attacks and the reliance is acknowledge by states.[26] Even though there have not been official cyber-attacks carried out by nation-states, these challenges give substance to the relevance of this thesis research on both a practical and a theoretical level. [27]

[24] Shackelford & Andres 2011 p. 976.
[25] *Ibid* p. 977.
[26] I. Drury, 'Cyber terrorists could inflict 'fatal' attack on Britain because Armed Forces rely so heavily on computers, Mps warn', *Mailonline* 9 January 2013 http://www.dailymail.co.uk/news/article-2259374/Military-cyber-attack-threat-Armed-Forces-rely-heavily-computers-MPs-warn.html *see also* D. Lague, 'Chinese see military dependence on computers as weakness', *New York Times* August 29 2007, http://www.nytimes.com/2007/08/29/world/asia/29iht-cyber.1.7299952.html?pagewanted=all&_r=0
[27] *See*: http://www.nato.int/docu/review/2013/Cyber/Cyberwar-does-it-exist/EN/index.htm

Chapter 2. What is the principle of distinction?

The research question consists of the abstract term 'the principle of distinction'. A cyber-attack as a part of cyber warfare has to be conducted in conformity with the principle of distinction, however In order to determine whether it can be conducted in conformity with it, the term principle of distinction has to be further explained. This chapter establishes a legal framework that enables a review to be done by predetermined standards. This chapter outlines the relevant aspects of the principle of distinction.

§2.1 A pillar of international humanitarian law.

This subsection determines what place the principle of distinction takes in the sphere of international law. The principle of distinction requires a distinction to be made between on the one hand combatants and on the other hand non-combatants during times of international armed conflict. At first sight this could be perceived as being a precise rule and perhaps as common sense, in practice this is not always the case. In the aftermath of World War II the UN Charter was framed which aimed to save succeeding generations from the scourge of war.[28] The principle of distinction is regarded to be a cardinal principle by the ICJ.[29] The principle of distinction also applies to objects, making it a duty for one to determine whether the object in question is military or civilian. The rationale behind the principle is that it prevents unnecessary suffering and loss to persons or objects not directly participating in hostilities, which could otherwise be widely interpreted as being unfair or unethical since they do not take part in hostilities and should therefore not become the object of attack. The principle of distinction is a part of the laws of war, but more in specific, it is a part of international humanitarian law or *jus in bello*. The principle of distinction is codified in the Geneva Conventions and is applicable in both international and non-international armed conflicts.[30]

[28] *Available at*: http://www.un.org/en/documents/charter/preamble.shtml

[29] *Legality of the Threat or Use of Nuclear Weapons (Advisory Opinion)* [1996] ICJ Rep 226, 256 [78] ('*Nuclear Weapons Opinion*') The ICJ notes that there is a second cardinal principle, which Is that of 'unnecessary suffering'. The ICJ also described this principle as one of the 'intransgressible principles of international customary law'.

[30] In Additional Protocol I, supra note 2, art. 48, this is referred to as the "basic rule." *See* Watkin 2010.

§2.2 Which documents are at the base of the principle of distinction?

In order to get a get a good understanding of the principle of the distinction, an overview of its historical background is given in this subchapter. The first time the principle of distinction was officially described as a legal norm was the Saint Petersburg Declaration of 1868. The principle of distinction was expressed in a rather formal way indicating that: 'the only legitimate object which States should endeavour to accomplish during war is to weaken the military forces of the enemy.' [31] This statement in the Declaration stresses that an attack issued by a state in conflict with another should only impact the other party's military foundation. After the Saint Petersburg Declaration the principle of distinction could be found in the Hague Conventions of 1907. The Hague Conventions deal with the conduct of military forces regulating the immediate combat zone on land.[32] One could have expected a further elaboration on the principle, this however seems not to be the case since the principle of distinction is not explicitly mentioned in the Hague Conventions of 1907. Article 25 of The Hague Conventions reads that: 'the attack or bombardment, by whatever means, of towns, villages, dwellings, or buildings which are undefended is prohibited.' It is an article derived from the notion that is the principle of distinction. Article 23 of The Hague Conventions is closely related to Article 25 which makes it a legal obligation to limit unnecessary suffering. The Hague Conventions together with the Geneva Conventions are both at the base of international humanitarian law. The Geneva Conventions of 1949 consists of four treaties and are complimented by three additional protocols. The Geneva Conventions offered a better legal protection for civilians in comparison to the Hague Conventions. The Geneva conventions have two Additional Protocols that are of particular interest. The principle of distinction is codified in Protocol I and II. Protocol I applies to international armed conflicts. Protocol II applies to non-international armed

[31] Neff 2010 p. 190.
[32] Solf 1986 p. 125.

conflicts. The principle of distinction in its purest form can be found under Article 48 of the Additional Protocol I of the Geneva Conventions. Article 48 AP/I states that:

> 'In order to ensure respect for and protection of the civilian population and civilian objects, the Parties to the conflict shall at all times distinguish between the civilian population and combatants and between civilian objects and military objectives and accordingly shall direct their operations only against military objectives.'

The principle of distinction is regarded to be one of the most fundamental pillars of International Humanitarian Law. Even more so, the principle and its concept to distinguish between civilian and combatant is regarded to be customary international law which ensures a strong legal foundation and a sound possibility to invoke its protection with success.[3334] The principle of distinction can also be found in the Articles 51(2) and 52(2) of the Additional Protocol I. A notable detail is that no reservations whatsoever have been made to those specific articles by the signatory parties.[35] Protocol I of the Geneva Conventions applies to international armed conflicts whereas Protocol II applies to non-international armed conflicts.

§2.3 The role of proportionality and necessity in relation to the principle of distinction.

This subsection takes note of additional principles that complement the principle of distinction, namely the principles of proportionality and necessity. Both principles apply in the *jus ad bellum* and *jus in bello*, however only the principles applying in *jus in bello* context will be reviewed. In this subsection it will be illustrated in what specific way these principles apply. The outcome whether an attack is illegal is not only determined by the distinction between civilian and combatant, it is also based on whether the principles of proportionality and necessity are respected. The principle of distinction is therefore complemented by these two principles. The principle of proportionality codified in Article 51(5)(b) of Additional Protocol I, and the principle of necessity which can be found in various forms in treaties, court statutes and jurisprudence such as that of ICTY and the ICC. Both principles are regarded to be

[33] *See Nuclear Weapons* Case para. 79.
[34] *See Blaškić* Case paras. 162 – 170.
[35] *See*: http://www.icrc.org/customary-ihl/eng/docs/v1_cha_chapter1_rule1 Under international armed conflicts

cardinal principles of international humanitarian law.[36] These principles are established to ensure that the attack is as severe as absolutely necessary against a military target so that the attack itself is justifiable.[37] The point of departure of the principle of necessity is that there has to be a situation that justifies extraordinary action in order to protect essential interests that are in danger of being irreparably damaged.[38] Necessity is often connected to the prerequisite of a state of emergency.[39] The principle of proportionality regulates the amount of force that is allowed to be applied. These principles therefore mark the outer lines of military conduct. An import detail is that these principles make it mandatory for one to assess the possible outcome prior to conducting an attack. A careful analysis of the situation has to be made before a subject is targeted, it has to be made *ex ante*. Dinstein in his TMC Asser lecture first illustrates the importance of the principles of proportionality and necessity by giving an example of a historical event, that of the Hiroshima bombing of 1945. Dinstein stressed that because the principle of necessity was not codified as a legal rule at the time the bombing occurred, the strike was not illegal.[40] The bombing targeted a military objective and was in 1945 therefore allowable. With the principles proportionality and necessity in mind, the bombing would not pass the test in current times, irrespective of the target being military due to the restraints on military conduct required by the principles of proportionality and necessity. There is critique to the principle. It is sometimes regarded to effectuate the opposite of what the limitation seeks to achieve, being used as an excuse to justify behaviour. [41][42] One can use the principle of necessity to argue that the behaviour was allowable because the boundaries set by the principle of necessity can be interpreted in a wide sense. The ICC however only accepts military necessity as a negation of an implicit or explicit defence and not as an excuse or justification.

[36] *See*: art. 53 GC A/D IV. *See also*: Hayashi 2010, pp. 46-52 . *See* art. 31(1)(c) ICC,
[37] Estreicher 2011, p. 6.
[38] Tsagourias 2010, p. 13.
[39] Hayashi 2010, p. 53.
[40] *See:* http://www.asser.nl/events.aspx?id=351
[41] Forrest 2007, p. 190.
[42] Hayashi 2010, p. 42.

§2.4 The legal status of the principle of distinction.

In this subsection the legal status of the principle of distinction will be reviewed. A principle's legal status ultimately determines who can invoke its protection. The principle of distinction is considered to be customary international law. It can be found in multiple military manuals reflecting state practice.[43] Because of its status, the principle is to be respected at all times, irrespective of whether the parties in armed conflict are a part of the Geneva Conventions or not.[44] In the ICJ Judgment of the Nuclear Weapons case Judge Weeramantry stressed that the principle of distinction has to be regarded a cardinal principle in international humanitarian law.[45] The ICRC published a study on rules of customary law.[46] The ICRC deemed rules to be customary based on foremost national reports as a primary source of state practice and *opinio juris*.[47] In that study, the ICRC noted 161 rules to be of customary international nature. The principle of distinction takes a prominent place: the first twenty-four rules of that database focus on the principle of distinction. The first twenty-four can be divided into rules that are subcategorized into the following six areas: distinction between civilians and combatants, distinction between civilian objects and military objectives, indiscriminate attacks, proportionality in attack, precautions in attack and precautions against the effects of attacks. All studies have their strengths and weaknesses and the ICRC study is no different in that respect. The ICRC study was criticized by publicists; one of the reasons being that the ICRC does not have the authority to declare any rule to be customary international law.[48] It should not be disregarded however that this study is one of the very few attempts to actually establish an overview of customary international law. The conclusions made by the ICRC can be verified due to the citations and references the ICRC uses in its research.

[43] *See:* http://www.icrc.org/customary-ihl/eng/docs/v2_rul_rule1
[44] *See*: http://www.asser.nl/default.aspx?site_id=9&level1=13336&level2=13374&level3=13460
[45] ICJ, *Nuclear Weapons case,* Advisory Opinion, 8 July 1996, paras. 78-79.
[46] *Available at:* http://www.icrc.org/customary-ihl/eng/docs/v1_rul
[47] McCormack 2006 p. 83.
[48] *Ibid* pp. 87-90.

§2.5 The principle of distinction in international or non-international armed conflicts.

In this subsection two types of conflicts will be mentioned since only in those type of conflicts international humanitarian law applies and therefore the principle of distinction. In international humanitarian law, the type of conflict corresponds with a specific set of rules. The definition of a combatant differs in international armed conflict from the one in non-international armed conflict. Regardless of the differences, the combatant forms a legitimate target according to the principle of distinction. In order for humanitarian law to be applicable, an international conflict has to be present. The first possibility is that there is an international armed conflict between two or more states. The second possibility is a non-international armed conflict in which one party is a governmental force, the other party a non-governmental force or armed group. This distinction can be witnessed in common Article 3 of the Geneva Conventions of 1949 and non-international armed conflicts falling within the definition provided in Art. 1 of Additional Protocol II.[49] The protection afforded in non-international armed conflict is in a considerable amount of areas less than in the international counterpart, as non-international armed conflicts are more infringing on a state's sovereignty as they in some cases can regarded to be internal affairs and at the discretion of that state. The principle of distinction applies to both armed conflicts but it differs per conflict who can indicated to be a combatant and therefore be attacked.

§2.6 The civilian and the combatant.

This subsection will determine what needs to be understood by the term civilian and combatant in the context of international humanitarian law. The principle of distinction revolves around two main concepts: the civilian and the combatant. Who exactly can regarded to be civilians for the purpose of the principle of distinction? The Brussels Declaration of 1874, the Hague Regulations and the Geneva Conventions all use a similar approach.[50] These international agreements seek to define the term combatant in a mutual exclusiveness approach. Civilians are defined by what they are not.[51] This approach creates a sort of safety net, establishing a residual category of persons who are afforded civilian

[49] *See*: http://www.icrc.org/eng/resources/documents/article/other/armed-conflict-article-170308.htm
[50] *See* the arts. 9-11 of the Brussels Declaration of 1874, the arts. 1-3 of the 1907 Hague Regulations and art. 4 of the Geneva Convention III, and the arts. 43-4 of the Additional Protocol I.
[51] Watkin 2010 p. 649.

status. In particular crucial is therefore a precise demarcation of the categories of combatants since that ultimately determines who is protected from attack and who is not. A starting point in determining who can be indicated a combatant is provided by the Geneva Conventions. Article 50 (1) of the Additional Protocol I mentions that: 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 respective Protocol. In case of doubt whether a person is a civilian, that person shall considered to be a civilian.[52] Article 50 (2) compliments the definition of a civilian by indicating that the civilian population comprises all persons who are civilian. Article 50 (3) notes that the presence within the civilian population of individuals who do not fall within the definition of civilians does not deprive the population of its civilian character. The Geneva Conventions further note that someone is not a civilian when they belong to persons mentioned in Article 43. Article 43 of the Additional Protocol I in turn refers to the organized armed forces of a Party to the conflict which are under a command responsible to that Party for the conduct of its subordinates, even if that Party is represented by a government or an authority not recognized by the other adverse Party, excluding medical personnel and chaplains.[53] The combatant is entitled to the Prisoner of War status which is accompanied by certain privileges upon being captured. An important fact in relation to the legal status is that no reservations have been made to Article 50 of the Additional Protocol I by the contracting Parties which ensures a strong legal foundation when it comes to defining civilians.[54] It has to be noted though that there are states that do not have the Additional Protocol I ratified, these include: the United States, Israel and Turkey.[55] This definition of a civilian in the context of international humanitarian law can be found in numerous military manuals and is shown by state practice.[56] The International Criminal Tribunal for the Former

[52] Article 4 (6) GC III that indicates that: 'inhabitants of a non-occupied territory, who on the approach of the enemy spontaneously take up arms to resist the invading forces, without having had time to form themselves into regular armed units, provided they carry arms openly and respect the laws and customs of war are not civilians but combatants for the purpose of the principle of distinction. The term used for this phenomenon is *levée en masse*.'

[53] Art. 43 (2) AP II.

[54] Additional Protocol I, Article 50 (adopted by consensus) (cited in Vol. II, Ch. 1, § 705).

[55] *Available at:* http://www.icrc.org/applic/ihl/ihl.nsf/States.xsp?xp_viewStates=XPages_NORMStatesParties&xp_treatySelected=470

[56] *See* http://www.icrc.org/customary-ihl/eng/docs/v1_cha_chapter1_rule3#refFn_34_2

Yugoslavia defined civilians to be: 'persons who are not, or no longer, members of the armed forces'.[57] The categories of persons who are afforded civilian status are regarded to be customary international law.[5859]

§2.7 Persons belonging to the armed forces belonging to a party to the conflict.

States oppose one another in armed conflict. Both States and their armed forces engage in battle. It is imperative to establish who can indicated to be a member of the armed forces since the principle of distinction indicates that they can be the legitimate object of attack. Article 50 of the Additional Protocol I determines that members of the armed forces to the party to the conflict are regarded to be combatants. The ICTY confirms this by noting that civilians are members who do not belong to the armed forces. Who precisely can be regarded a member of the armed forces? Additional Protocol I of the Geneva Conventions states who falls under the notion of an armed force namely: all groups, forces and units that are organized and under a chain of command. A level of organization therefore has to be present next to a military-like structure. The first requirement is that there has to be a sufficient amount organization to be found in the groups, forces or units. A second requirement is that there has to be someone accountable for the conduct shown. A strong indicator towards that one is dealing with an organized armed force is when orders are issued from one person or delegation residing in a specific city or region. This is the same for control over a particular geographical area and the ability of a group to set up a training regimen and disciplinary model.[60] The armed force has to have some sort of hierarchical relation between the persons belonging to the organized group and the commander. There has to be an internal disciplinary system present. The person in charge has to be able to issue out orders and should be responsible for the conduct displayed by his subordinates. A requirement is that these organized groups have to abide legal obligations as laid out by international law in relation to armed conflicts. Article 43 of the Additional Protocol I complements the notion of armed forces belonging to a party to the conflict. It underlines that the armed forces of a party to a conflict comprise all organized armed forces, groups and units which are under a command responsible to that Party for the conduct of

[57] Prosecutor v. Blaskic IT-95-14-A 29 July 2004 para 114.
[58] Henckaerts & Doswald-Beck 2005 pp. 384-389.
[59] Crawford 2011 p. 13.
[60] *Limaj* et al. Case, judgement para 133

its subordinates, even if that Party is represented by a government or an authority not recognized by an adverse Party.[61] There is a possibility that there are different groups involved like militias and volunteer corps, groups other than the normal armed forces. They are still recognized as being armed forces if they fulfil four elements: (a) responsible command; (b) fixed distinctive sign recognizable at a distance; (c) carrying arms openly; and (d) operating in accordance with the laws and customs of war.[62]

[61] Art. 43 [1] AP I; Customary IHL, above N 7, Vol. I, Rule 4.
[62] 14 A rt. 1 H IV R; Arts 13 [1], [2], [3] and [6] GC I and GC II ; Art. 4 A [1], [2], [3] and [6] GC III .

Chapter 3. Defining cyber-attack.

There is no definition of a cyber-attack in international law present. In this chapter the task of defining a cyber-attack is carried out in order to create a better understanding of what the respective term comprehends. A cyber-attack in the context of cyber warfare has to conform to the principle of distinction and a definition provides support in determining what specific attack can regarded to be a cyber-attack. A definition establishes guidelines so that it can be examined whether the conduct that amounts to a cyber-attack can be carried out while conforming to the principle of distinction.

§3.1 Definition of cyber-attack.

This subsection will outline the definition of cyber-attack that will be used for the purpose of this thesis. There is no conclusive definition determined in the laws of armed conflict of what specific act constitutes a 'cyber-attack'. The relevance of having a definition of an attack is that the notion of attack serves as a basis for a number of specific limitations and prohibitions in the law of armed conflict.[63] The terms cyber and attack are reportedly used without regard for what they are meant to include.[64] The standard in international law is the armed attack. In order for a state to invoke its right to self-defence an armed attack is what international law requires to have taken place against the state invoking its right. Article 51 of the UN Charter revolves around the notion of armed attack. The word 'attack' itself is used in different contexts. A first possible way of it being used is in military sense. An attack will indicate the targeting of a particular object or person when used in that particular context. A different way it can be used is in the context of the *jus ad bellum*. In that case, one state resorts to armed violence and an attack is what triggers a possible justifiable response of self-defence.[65] In relation to the principle of distinction the relevant version of attack is however one that applies in armed conflict or *jus in bello*. This version of an attack refers to the following: 'a particular type of military operation that involves the use of violence, whether in offense or defense.'[66] It is a definition as proposed by the International Group of

[63] *Ibid* p. 92.
[64] Hathaway 2012, p.15.
[65] *Ibid* p. 20.
[66] The Tallinn Manual mentions that 'Acts of Violence' should be understood in terms of the consequences of an attack and not its nature. p. 93

Experts in the Tallinn Manual and can also be found in the German military manual.[67] Most importantly, it is a definition based on Article 49 of the Additional Protocol I which notes: 'attacks mean acts of violence against the adversary whether in offence or defence.' In this respect, the notion of attack revolves around the usage of violence directed against a target. Defining a cyber-attack is challenging for a variety of reasons. One of the main reasons a definition of a cyber-attack is challenging to establish is that the intensity and form of the cyber-attack tends to vary. The level of a cyber-attack is difficult to determine at the time of the actual attack, since the level of the attack can shift over the course of the attack.[68] Other versions of a definition of a cyber-attack include a wide variety of acts of cyber terrorism and cyber warfare and some commentators even regard cyber-attacks to be an entire separate category.[69] A definition of a cyber-attack used by Hathaway indicates that: 'a cyber-attack consists of any action taken to undermine the functions of a computer network for a political or a purpose of national security.'[70] The International Group of Experts agreed to a different definition of a cyber-attack. The Tallinn Manual defines a cyber-attack to be: 'a cyber-operation, whether offensive or defensive, that is reasonably expected to cause injury or death to persons or damage or destruction to objects.'[71][72]The definition of the Tallinn Manual is based on Article 49 (1) of Additional Protocol I. Note the differences between the definition of the International Group of Experts and the one proposed by Hathaway. Hathaway has a fairly strong objective based definition mentioning what a cyber-attack seeks to accomplish without neglecting what sets a cyber-attack apart from a regular attack, namely that it targets a computer network thereby seeking to undermine its functioning. It stipulates that every action taken in that respect can regarded to be a cyber-attack however, the action has to target the computer network seeking to undermine its functions. The definition also mentions that the motive for attacking has to be

[67] *Ibid* p. 92. *See also*: The German military manual 'Humanitäres Völkerrecht inbewaffneten Konflikten Handbuch' that notes under 'III. Kampfmethoden 441': 'Angriffe das heißt jede offensive oder defensive Gewaltanwendung gegen den Gegner'.
[68] *Ibid*.
[69] *Ibid*.
[70] Hathaway et al. 2012, p. 10.
[71] Rule 30 – definition of Cyber-attack *see* Schmitt et al. 2013, p. 92.
[72] According to the International Group of Experts interference of functionality amounts to destruction if replacement of physical components is necessary. In case a reinstallation of an operating system restores functionality that threshold is not met. A few experts suggest that the object's loss of usability in itself already constitutes the requisite damage. Also in case an operation results in large scale adverse consequences it is definitely a type of damage that can be debated on as a few experts point p. 94.

political and based on purposes of national security. In case a state is behind a cyber-attack the political motivation is given and so is the reason of national security. The requirement of a cyber-attack undermining the computer network translates into a demand of the computer network to be comprised in such a manner that it affects the current or future ability to function.[73] More than passively observing or copying data, damage has to be done to the operating system or false misleading or unwelcome information has to be added.[74] The definition as provided by the International Group of Experts is less specified than Hathaway's. The primary difference of an attack compared to other operations is the element of violence. The definition provided by the International Group of Experts addresses this element by indicating that a cyber-attack has to be reasonably expected to cause injury or death to persons, or damage or destruction to objects. The effects, or in other words the consequences of the attacks, are central in defining an attack to be a cyber-attack. The definition sets boundaries of what might constitute an attack without requiring a specific amount. A cyber operation in offensive or defensive is stated in the definition. An example of a defensive cyber operation is for instance a countermeasure that upon being engaged strikes back, referred to as: 'electronic countermeasures designed to strike attacking computer systems and shut down cyber-attacks midstream.'[75] A distinct part of the definition is the term cyber operation. The advantage to this rather open term is that it enables multiple forms of conduct to fall under the definition. It does not specifically mention the requirement of having to undermine the functions of a computer network unlike the definition of Hathaway. Not all computers are a part of a network. A computer system can be a stand-alone unit and the definition established by the International Group of Experts prevents difficulties arising from having the network requirement. This has to be seen in perspective; a significant amount of systems nowadays is connected through a network such as the internet or a local area network. The definition as proposed by Hathaway does not focus on the term violence therefore an attack that does not meet the threshold of damage or destruction can in some cases still be classified a cyber-attack. For example, in the Tallinn Manual some Experts mention that the international community could view certain attacks that are large-scale and

[73] Hathaway et al. 2012, p. 14.
[74] *Ibid*.
[75] Carr 2011, p. 46.

adverse as cyber-attacks. In certain cases these attacks might not make the threshold of damage or destruction.[76] The absence of this requirement enables these attacks to still fall under the category of a cyber-attack. Territory seems to be left out of the definition both by the International Group of Experts and Hathaway. Territory is interlinked to the addendum of state sovereignty. The International Group of Experts discussed the topic of territory integrity in relation to the violation of State Sovereignty in the Tallinn Manual. They indicate that the main rule is that States cannot exercise sovereignty over cyberspace *per se*.[77] States may still exercise their jurisdiction based on international law in case of cybercrimes and other cyber activities. The Tallinn Manual notes three grounds for jurisdiction in case of a cyber-attack. The first ground is over persons engaging in cyber activities on a state's territory. The second ground is over cyber infrastructure located on its territory. The final ground mentioned is extraterritorially, in accordance with international law. The Tallinn Manual underlines that there are two derivative forms of jurisdiction.[78] The first is subjective territorial jurisdiction involving the application of the Law of the State exercising jurisdiction to an incident initiated within its territory but completed elsewhere. Objective jurisdiction grants jurisdiction over individuals to the state where the particular incident has effects even though the act was initiated outside its territory.[79] The definition of a cyber-attack to be used in this thesis will be the one established by the International Group of Experts mentioning a cyber-attack to be:

> 'a cyber operation, whether offensive or defensive, that is reasonably expected to cause injury or death to persons or damage or destruction to objects.'

The Resemblance to the provision in the Geneva Convention Protocol I makes it a definition close to a codified legal provision which enhances its legitimacy. That is the primary reason for choosing to use this specific definition for the purposes of this thesis.

[76] Schmitt et al 2013, p. 94.
[77] *Ibid* p. 26.
[78] Based on the opinion of the Attorney General of the European Court of Justice in the Ahlström Osakeyhtiö and Others v. Comm'm case there a two derivative
[79] Schmitt et al. 2013, p. 28.

§3.2 Cyber-attacks: kinetic and non-kinetic.

This subsection addresses the terms kinetic and non-kinetic in relation to cyber-attacks. These terms are being used in literature to point out differences between cyber-attacks and conventional attacks. There is a set of differences between non-kinetic cyber-attacks and kinetic armed attacks to be distinguished. This can be done on a variety of levels. One possible way is by looking at the cyber-attack itself. In the vast majority of cases a cyber-attack will be non-kinetic by nature. It is an attack carried out by using the power grid and the electronic connection of computer systems. It targets for instance software using a digital signal consisting out of the values 0 and 1. In comparison to a conventional attack physical matter is absent. Conventional attacks are in cases combustion driven and involve a projectile of some sort. Physical matter is used to conduct an attack by. In the context of computer attacks offensive capabilities do not necessarily translate into defensive capabilities. Being better at offensive capabilities does not make a state better defensively. That is a sharp contrast to many conventional weapon systems. It is less costly to mount cyber-attacks than conventional attacks and also more industrialized states are in general more dependent upon computer systems which leave them more vulnerable to those attacks.[80]

[80] Hathaway 2012 p. 28.

Chapter 4. Forms of cyber-attacks.

In this subsection a categorization is made of the various possibilities of conducting cyber-attacks. Based on the previously established definition of cyber-attack this subsection will cover 'cyber weapons and their associated cyber systems and cyber tactics, techniques, and procedures by which hostilities are conducted' and will further elaborate on this subject. There are three main categories of cyber-attacks distinguished. The first category is malware, the second is hacking and the third is DDoS. These forms of cyber-attacks are further explained and their characteristics are reviewed in a concise manner. The varieties chosen are the ones that frequently appear in publications by authors addressing cyber-attacks. Due to the existence of a multitude of possible ways of conducting an attack the categories as listed below, even though carefully selected, do not provide an extensive list due to the evolving nature. The specific forms mentioned are relevant when they are conducted in the context of cyber-warfare.

§4.1 Malware

This subsection will cover a computer-software-related cyber-attack namely that of malware. Malware is short for Malicious Software. The spreading nature of malware is one of its key characteristics. An overview of the categories of possible malware demonstrates that it comes in a variety of ways. One elementary way of categorizing the various malware threats is done by Gookin.[81] He distinguishes the following categories: phishing, spyware, Trojan horse, virus and worm. These respective categories will be addressed in a concise manner. The first form of malware is that of phishing. Phishing is trying to obtain information by misleading a target by purposely deceiving. An example of phishing is trying to make the target submit sensitive information by pretending to be a trusted institution such as a bank or store. Certain requests are then made that allows the perpetrator to collect sensitive information since the target is led to believe that the request is made by a trusted party. Malware often also monitors internet movement and sends back this information. This is then used to send specific information to the target such as offers. The most relevant in relation to the research question are the following ones. The

[81] D. Gookin, 'Know the Different Types of Malware', *For Dummies* http://www.dummies.com/how-to/content/know-the-different-types-of-malware.html

first is a Trojan horse or Trojan. A Trojan horse is a program that appears legitimate but contains hidden code allowing unauthorized collection, exploitation, falsification, or destruction of data on a host computer.[82] A Trojan horse is a program disguised as a legitimate program that has ulterior motives. Trojans can delete data, comprise security and infect a computer with a program script altering its overall functionality.[83] A different form of malware is that of a virus. A virus can take over control upon infecting a computer. It can destroy data and often searches for sensitive information such as saved passwords. This virus will likely send the information to another computer. A virus can also be used to coordinate attacks against websites on the internet.[84] A final form of malware is that of a worm. It is an independent program that replicates itself across network connections. A worm often congests networks as it spreads.[85] Worms and viruses are commonly used to describe an identical form of malware. The inherent characteristic of a worm namely that of replicating itself makes it a frequent carrier for a virus. The worm can therefore also comprehend the characteristics of a virus. As of 2012, it is estimated that there are a 75 million unique varieties of malware to be found globally.[86]

§4.2 Hacking

A different form of conducting a cyber-attack is that of hacking. A definition of hacking is provided by a website dedicated to this phenomenon. It notes that: 'hacking is the practice of modifying the features of a system, in order to accomplish a goal outside of the creator's original purpose'. The person who is consistently engaging in hacking activities, and has accepted hacking as a lifestyle and philosophy of their choice, is called a hacker.[87] The Merriam-Webster Dictionary describes the hacker to be: 'a person who illegally gains access to and in some cases tampers with information in a computer system'. In the context of cyber warfare hacking is done to gain unauthorized access to a computer system of the Nation-

[82] Vatis 2001, p. 7.
[83] *Ibid*.
[84] *Ibid*.
[85] Vatis 2001, p. 7.
[86] C. Herridge, 'NSA chief: Cyber-attacks skyrocket, account for largest 'transfer of wealth' ever', *Fox News* July 9 2012 http://www.foxnews.com/politics/2012/07/09/nsa-chief-cyber-attacks-skyrocket-account-for-largest-transfer-wealth-in/%20-%20ixzz2XJe3lSlk#ixzz2cnw5ditF http://www.foxnews.com/politics/2012/07/09/nsa-chief-cyber-attacks-skyrocket-accwount-for-largest-transfer-wealth-in/#ixzz2XJe3lSlk
[87] *See*: http://whatishacking.org/#sthash.pTv1sBdg.dpufA

State to either alter a program script or to obtain classified, strategic or sensitive information to use for their own benefit.

§4.3 DDoS

A different form of conducting cyber-warfare is by way of DDoS attack. This subsection will address its characteristics. DDoS is an abbreviation that stands for Distributed Denial of Service. It is a more sophisticated form of a Denial-of-Service attack. A definition of a DDoS attack is given by Vatis: 'Distributed Denial-of-Service attack (DDoS): action(s) by distributed computers that prevent any part of another computer system from functioning in accordance with its intended purpose'.[88] There are two types of DDoS attacks: a network-centric attack which overloads a service by using up bandwidth and an application-layer attack which overloads a service or database with application calls.[89] Both DoS and DDoS attacks are based on the same principle, the difference between them is the amount of computers used to conduct the attack by. A DDoS attack uses multiple computers carry out the attack. It seeks to hamper or even disable the accessibility of the targeted network or system so that valid users cannot make use of it. Certain countries have been suspected to have used and still use these DDoS attacks.[90] The responsibility for DDoS attacks are usually claimed by certain groups supposedly not state connected. Examples include: a series of attacks over a three week period in 2003 in which websites of Estonian organizations were targeted.[91] The youth group Nashi claimed to be responsible. Nashi is regarded to be associated with the Russian government.[92] A different case is that of the attacks against South Korea on the 63rd anniversary of the Korean War. While multiple attacks were conducted by multiple perpetrators, one of the Distributed Denial-of-Service attacks observed against South Korean government websites could be directly linked to the North Korean associated DarkSeoul gang.[93] The

[88] Vatis 2001, p. 7.
[89] M. Rouse, 'Definition distributed denial-of-service attack (DDoS)', *SearchSecurity* May 2013, http://searchsecurity.techtarget.com/definition/distributed-denial-of-service-attack
[90] D. Zhou, 'Iran Wages Cyber War Against US Banks and Arab Energy Firms', November 2012, http://www.policymic.com/articles/16555/iran-wages-cyber-war-against-us-banks-and-arab-energy-firms
[91] J. Ryan, "iWar": A New Threat, Its Convenience – and Our Increasing Vulnerability, *NATO Review* December 2007, http://www.nato.int/docu
[92] M. Schwirtz, 'Estonia bans travel for Kremlin Youth Group',*New York Times* January 2008, http://www.nytimes.com/2008/01/30/world/europe/30russia.html?_r=0
[93] *See*: http://www.symantec.com/connect/blogs/four-years-darkseoul-cyberattacks-against-south-korea-continue-anniversary-korean-war

DarkSeoul gang has also been linked to attacks on the United States on Independence Day.[94] Among the websites that were reported to be affected were the White House and The Pentagon.[95]

§4.4 What are main cyber-attack related characteristics?

This subsection outlines characteristics of cyber warfare. The most relevant of characteristics in relation to the principle of distinction will be outlined. The characteristics of cyber-attacks that will be reviewed are the following. First, the increase of distance between the attacker and target. Second, the effect it has on reciprocity on the battlefield. Finally, the targeting of dual use platforms and the non-kinetic nature of cyber-attacks. The characteristics of cyber-attacks are listed in a non-conclusive manner nor do they exclusively apply to cyber-attacks. According to a variety of authors, cyber-attacks skip the traditional battlefield on various fronts.[96] Cyber-attacks can be successfully conducted without having to traditionally engage a nation's defences. Modern technology enables the distance to be increased between weapons and the lethal force they project.[9798] One can physically be present in a certain continent and attacking in another. This then can have an impact on reciprocity between the actors of warfare. According to author Watts reciprocity can only be achieved where the actors share compatible interests and norms of obligation.[99] The increase of distance can be a factor that has effect on reciprocity because parties are not exposed in the same terms because their distance to the battlefield differs. The targeting of dual use platforms is also indicated to be a key characteristic of cyber warfare. A dual object or entity is one that serves both civilian and military purposes.[100] An example of a target that had dual purpose is the Serbian television station that was bombed during the Kosovo conflict in which 16 civilians lost their lives. NATO stated that their main reason for bombing the state-owned television station was to disrupt communications by the C3 network. NATO stressed that the object of attack was

---

[94] *Ibid*.
[95] *See:* http://www.defense.net/index.php/ddos-in-depth/ddos-timeline/
[96] Clarke 2010 p. 21.
[97] Hsia & Sperli At War Notes form the Front Lines: How Cyberwarfare and Drones Have Revolutionized Warfare *NYTIMES* June 17 2013 http://atwar.blogs.nytimes.com/2013/06/17/how-cyberwarfare-and-drones-have-revolutionized-warfare/?_r=0
[98] U.N. General Assembly, 23rd Session. Human Rights Council. *Summary Report of the Special Rapporteur on extrajudicial,*
summary or arbitrary executions, Christof Heyna. 9 April 2013 (A/HRC/23/47).
[99] Watts 2011, p. 370.
[100] Richardson 2011, p. 28.

the control system that was used to manipulate the military and security forces and not the Serbian people and Milosevic.[101] In the context of a cyber-attack it can regarded to be more elementary. The medium used to carry out an attack, the internet, is dual use by nature since it is used for civilian and military alike. Therefore there is a clear nexus between what is being used to facilitate and enable the attack and ultimately the attack itself. That inherent connection is in itself a strong indicator of why dual use is a key characteristic of a cyber-attack.[102] An example could be a cyber-attack on a power company which will affect both military and the civilian population by a power shutdown, clearly illustrating the dual use of the target. As noted earlier, a characteristic of cyber warfare is that the source of the attack cannot always be determined. For instance, when a virus is developed the program script does not necessarily give away the attackers identity. This also applies to other forms of cyber-attacks such as hacking or DDoS attacks. Another difference is that attacks are mostly non-kinetic by nature. Cyber-attacks comprise of attacks on a different level but can have the same effects their kinetic counterparts have.

[101] Final Report to the Prosecutor by the Committee Established to Review the NATO Bombing Campaign Against the Federal Republic of Yugoslavia available at: http://www.icty.org/sid/10052#IVB3.
[102] E. Kodar, 'Applying the Law of Armed Conflict to Cyber-attacks: From the Martens Clause to Additional Protocol I,*ENDC Proceeddings,* vol 15 2012.

Chapter 5. Can a cyber-attack be conducted in conformity with the principle of distinction in international armed conflicts?

Waging war in conformity and therefore respecting the principle of distinction is an explicit choice made by parties to a conflict. This chapter will review whether a cyber-attack can be conducted by a party while still respecting this fundamental principle. Hacking, malware and DDoS will be reviewed for that purpose.

§5.1 Can cyber-attacks distinguish between civilian and combatant?

In this subsection, the respective forms of cyber-attacks as outlined earlier will be reviewed. They will be connected to the element of human decision making which will be determined of key importance in respecting the principle of distinction. To illustrate the importance of human decision making when it comes to upholding the principle of distinction, the example of Lethal Autonomous Robots is given. This is done despite the fact that Lethal Autonomous Robots are not to be categorized under cyber-attacks. The importance of having human decision making is then connected to the respective categories of cyber-attacks. It will be reviewed whether they have a possibility to terminate the attack late into the attack. First, hacking followed by malware and finally DDoS attacks will be reviewed. The characteristics of the respective means of cyber warfare will be used to determine whether specific weapons can distinguish between combatant and civilian. As noted earlier, one of the key elements pointing towards the ability to distinguish between civilian and combatant is the element of having a human decision making possibility prior to the actual attack, more in specific, before the effects of the attack occur. The United Nations General Assembly at the 23rd session openly expressed their concerns to what is referred to as Lethal Autonomous Robots (LARs).[103] So did the ICRC and Human Rights Watch. LARs are weapon systems that, once activated, can select and engage targets without further human intervention.[104] Aside from the possible moral objections to this notion, the implications on requirements of international humanitarian law particularly concerns this thesis in reviewing whether

[103] *See*: Human Rights Watch, 'Losing Humanity: The Case against Killer Robots', November 2012 http://www.hrw.org/news/2012/11/19/ban-killer-robots-it-s-too-late, S. Leo, 'ICRC issues statement on robotic warfare', *AOAV* 7 August 2013, http://aoav.org.uk/2013/icrc-statement-on-robotic-warfare/

[104] U.N. General Assembly, 23rd Session. Human Rights Council. *Summary Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions, Christof Heyna.* 9 April 2013 (A/HRC/23/47).

certain means of cyber-attacks can conform to the principle of distinction.[105] Human control plays an important role in conforming a cyber-attack to the principle of distinction, more in specific, the possibility to abort or continue the attack right before the effects of the attack occur. UCAVs, commonly known as drones, have similar characteristics as the LARs but the most significant difference is that ultimately the decision to engage is at the discretion of the person(s) controlling the drone. The characteristic of a person controlling the drone is called having a human or man in the loop, their autonomous counterparts have therefore humans out of the loop. Proponents of the use of robots that have full autonomy in engaging the adversary indicate the arguments for the use of LARs to be multifarious. Robots have sensors that give them a degree of situational awareness.[106] LARs can determine how to react to certain findings because of processors and or artificial intelligence and effectors can then decide what action they seem fit. Robots are effective at dealing with quantitative issues, making a decision or acting a certain way based on predetermined code logic. Opponents of the use of LARs indicate that they have limited abilities to make qualitative assessments. A battlefield is a complex situation and there is a great need for interpretation when contradictory settings occur. For instance, if a soldier is wounded and is on the verge of surrendering which can be derived from past battlefield' experiences coupled with real-time analysis, a LAR would most likely engage whereas a soldier would change his strategy knowing that a surrender could be at hand. A different argument is that robots carry out the conduct they have been instructed to, without having the option to critically review the order first and in case the assigned order is illegal refusing to carry it out. Accuracy and speed are virtues that LARs possess over man, however deciding over human life and the complex issues that are interlinked to that decision call for abilities that are not or at least to a lesser degree possessed by LARs. The principle of distinction, requires: human judgment, understanding intentions behind people's actions, the appreciation of the larger picture and the anticipation of the direction in which events are unfolding.[107] Deciding who can legitimately be targeted therefore requires contextual interpretation of not only emotions but also intentions of person(s) in question and that is, at least at the moment, better

[105] *Ibid.*
[106] *Ibid* p. 14.
[107] *Ibid*.

done by humans than robots. Not disregarding shortcomings of humans that robots are not burdened by.[108] The ability of having a person ultimately decide to engage or stand-down ensures a higher probability that distinction between civilian and combatant will be respected. Researchers and scientists continue to develop algorithms that are more advanced so there could be a time that technology can reach the standard the human mind has, however it remains uncertain to indicate when technology will reach that standard. The element of a human in the loop is critical in cyber warfare for the principle of distinction to be upheld. Therefore the respective means of cyber warfare as addressed earlier will be reviewed in connection to this element. First, hacking and its characteristics of having the human decision making element present. The word hacker is commonly used to indicate a person who engages in the activity of hacking. It is a person who gains access to a computer system without permission. This already emphasizes the element of human control and therefore a man in the loop. In the context of cyber warfare hacking is done to gain access to a computer system of the opposing Nation-State. The targets can therefore be accurately selected respecting the principle of distinction by targeting only military objectives in the form of computer systems in case the target can be hit without having to engage civilian objectives in order to reach it. It has to be noted that this can be challenging because of the aforementioned characteristic of a cyber-attack namely that of dual usage of targets. The means itself is however able to make a distinction between civilian and combatant. Second, the means that is malware will be reviewed. Malware is malicious software created by a software developer. It is code written to, amongst others, effectuate a network breakdown, ensure the loss of critical data and in some cases even destroy a system entirely. An attack might be imperceptible for the users of a compromised machine.[109] The ability to distinguish between civilian and military objectives seems to differ and highly depending on what type of malware one is dealing with. Every form is therefore programmed to effectuate something specific, however as soon as it has been launched the malware carries out what it has been programmed to do. In most cases it conducts tasks the user did not intend. Every form of software may

[108] Humans have emotions that could make lead to undesirable decisions when engaging the enemy such as acting out of revenge or a because or anger. Humans are more vulnerable and that can hamper the thoroughness of collecting the necessary info at the battlefield to decide whether it is legitimate target or not getting close to a friend or foe is better done by robots.

[109] *See*: Kaspersky Securelist http://www.securelist.com/en/threats/detect?chapter=76

it be spyware, which collects information secretly and passes it on to the requesting party or a Trojan horse that sabotages the system replaces critical data, simply follows the programmed code logic as soon as the software is executed.[110] The human element as in the human in or out of the loop function is because of that characteristic not present in a significant amount of malware. In the context of cyber warfare the decision to engage and use malware is arguably given earlier in the chain leading up to the attack namely from the moment one decides to direct the malware at a certain target. In case of UCAVs, the decision in general to use a drone for a mission is given earlier, however right before 'pulling the trigger' the crucial decision to strike is given immediately before. The characteristics of malware, in particular the way it operates and the spreading nature of it, makes it more likely that this particular means of cyber warfare cannot distinguish between civilian and military as the predetermined algorithm has to have taken into account all the possible variables of which some cannot be pre-set since they shift on the battlefield in determining who can be regarded a civilian or combatant. Where spyware, phishing and the Trojan horse can be controlled to a certain extent by directing them against a military target relatively accurately, because of their reliance on a person to direct them at a specific target, the phenomena of worm and virus have a higher risk of not being able to conform to the principle of distinction due to their inherent characteristic of following the preprogramed code.[111] Finally, the cyber-attack that is DDoS will be reviewed. According to the Dutch Criminal Code article a DDoS attack is often preceded by the intrusion of an automated system.[112] A Distributed Denial-of-Service attack is one of which a multitude of compromised computer systems attack a single target, thereby causing denial of service for users of the targeted system. The definition used in the Tallinn Manual is: a technique that employs two or more computers, such as the bots of a botnet, to achieve a denial of service from a single or multiple targets. The ability of selecting a single target offers the degree of accuracy necessary to uphold the principle of distinction. A past prime example would be the attack on the Pentagon website. The main issue here is not the selection of targets, it is the use of multiple computers to carry out the attack with is what could concern in terms of upholding the principle of distinction. It is not uncommon

[110] S. Nothcutt, 'Logic Bombs, Trojan Horses, and Trap Doors', *Sans Cyber Security Grad School* 2 May 2007 http://www.sans.edu/research/security-laboratory/article/log-bmb-trp-door
[111] E. van Geest, 'Van herkenning tot aangifte: Handleiding Cyber Crime', *KLPD* Juli 2003 p. 36
[112] 138a lid 1 WvSr.

for civilian computers to be used worldwide to carry out an attack. These computers are not initially setup to carry out such an attack, however they still end up facilitating the attack regardless. These computers become a part of a network of computers called a botnet and will be used to attack a certain target by sending requests and trying to connect at the same time resulting in a loss of functionality. In a botnet multiple computers are ordered to attack a target and because of a great amount of computers these botnets usually consist of, civilian computer systems are being used. It is however not unreasonable to think that governments possess a wider range of powerful computers that are able to perform and even outperform civilian computers in their task of facilitating a DDoS attack making a DDoS a possibility in international armed conflicts. Depending on the status of the target, a DDoS attack is therefore able to be directed against targets respecting the principle of distinction.

§5.2 The Love bug and Stuxnet: 2 past cases of worms.

As noted earlier, worm and virus attacks have a higher tendency to infringe on the principle of distinction because of two inherent characteristics they possess. First, a worm replicates itself and does it in a such fashion that distinguishing between combatants and civilians could be challenging. Second, a worm conducts an attack according to its programmed script. Terminating or cancelling an attack is therefore challenging once the worm is executed. These two characteristics make it more prone for this type of means of cyber warfare to infringe on the principle of distinction. Two cases will be analysed to determine whether an infringement on the principle of distinction cannot be avoided. This subchapter outlines two cases: the first is the Love Bug case, the second is the Stuxnet case. Both cases revolve around a worm/virus. Although the attacks in question have not occurred in armed conflict, the findings are nevertheless relevant as there have been indicators of Nation-State' involvement in past cases of the targeting of military objectives. These cases will further outline and display specific characteristics of worm/virus cyber-attacks. It is strictly a qualitative and an in-depth analysis since unfortunately there is no possibility to conduct quantitative empirical research in which an analysis could be done of a greater sample of the estimated 1 million+ virus/worm outbreaks that are being carried out or have been carried

out at one point.[113] The first case is about the ILOVEYOU virus or also called the Love Bug virus. It is chosen because of several reasons. First, because of the widespread uproar the worm/virus caused at the time it was active.[114] Second, because its coverage was extensive, it reached military targets such as the Pentagon and the houses of parliament in London. Lastly, this particular virus was chosen because of the financial damage it caused which is said to have reached over 10 billion.[115][116] The second case is that of the Stuxnet worm/virus. The Stuxnet worm/virus is a cyber-attack that is primarily known for its advanced code. This code has gained recognition as being highly sophisticated and is reportedly noted to be 'complex'.[117][118] The Stuxnet worm/virus possesses unique characteristics making it an odd beast compared to other worm/viruses. The way the analyses is conducted is as follows. First, the facts of the cases will be outlined. After the facts have been discussed, the effects of the cyber-attacks will be reviewed. Since all the effects need to be taken into account when reviewing whether an attack would infringe the principle of distinction, an in-depth analysis is conducted in relation to the effects an attack has. Due to the characteristics of these forms of cyber-attacks namely that these virus/worms replicate, spread over a network and attack a target computer system by exploiting a system's weakness, a conclusion will be made whether a virus/worm infringes the principle of distinction or whether it still can be carried out respecting the obligation to refrain from attacking civilians and civilian objects. In the following subsections worm and virus are being used interchangeably and although they are not entirely the same type of means according to IT-related sources, the Love Bug and Stuxnet entities both have virus related functionalities and worm-like characteristics.[119]

---

[113] 'Computer viruses hit one million', *BBC News* April 10 2008 http://news.bbc.co.uk/2/hi/technology/7340315.stm

[114] D. Kleinbard and R. Richtmyer 'U.S. catches 'Love' virus', *CNN Money* May 5 2000 http://money.cnn.com/2000/05/05/technology/loveyou/

[115] 'Virus hits secret Pentagon network', *BBC News* May 6 2000 http://news.bbc.co.uk/2/hi/science/nature/738276.stm

[116] M. Landler, 'Filipino Linked to 'Love Bug' Talks About His License to Hack', *The New York Times* October 21 2000 http://www.nytimes.com/2000/10/21/business/a-filipino-linked-to-love-bug-talks-about-his-license-to-hack.html

[117] Falliere et al. 2011, p. 1.

[118] Richardson 2011, p. 20.

[119] *See*: http://www.cisco.com/web/about/security/intelligence/virus-worm-diffs.html, *see also* http://www.symantec.com/avcenter/reference/worm.vs.virus.pdf

§5.3 The Love Bug worm.

The first example of an attack that possibly infringes the principle of distinction, is that of the ILOVEYOU virus or commonly known as the Love Bug worm. The way the virus worked is as follows.[120] The virus enters the computer system through an email with the subject 'ILOVEYOU'. The email itself reads: 'kindly check the attached LOVELETTER coming from me.' The email is then accompanied by an attachment named: 'LOVE-LETTER-FOR-YOU.TXT.vbs'. When executing the attachment, the virus infects the system. The virus copies itself three times. It can be found in a. the WindowsSystem-folder as the Win32DLL.vbs file, b. the Windows-folder named MSKernel32.vbs and c. the TEMP-folder as the LOVE-LETTER-FOR-YOU.TXT.vbs file. When starting up the computer system, MSKernel32.vbs andWin32DLL.vbs are activated. The virus activates each time the computer starts up due to the effect the virus has on the computer registry. It changes the home page of Internet Explorer. The virus searches the computer system for passwords and emails them to mailme@super.net.ph. The virus also searches for email addresses in Microsoft Outlook to which an email is sent containing the same subject and attachment. Visual related files are replaced and audio files are hidden. Finally, the virus searches for a chat program by the name of mIRC. A HTML file is then created by the name of LOVE-LETTER-FOR-YOU.HTM. This file contains the virus and is sent to all the contacts of the account through the mIRC program thereby causing it to spread further.

§5.3.1 The Stuxnet worm.

This subsection outlines the factual case of the Stuxnet worm. Stuxnet is regarded to be the first worm targeting an industrial control system. In this case, it targeted an industrial control system of a nuclear power plant.[121] The attack was initiated by the virus being introduced to the computer network of the targeted environment. It is unknown who exactly inserted the virus in the target computer network. A

[120] IBM Security Systems, 'Ahead of the threat' *available at*: http://iss.net/ seach for *iloveyou-worm*
[121] J. Fildes, 'Stuxnet worm 'targeted high-value Iranian assets', *BBC News* September 23 2010 http://www.bbc.co.uk/news/technology-11388018

plausible scenario is that a USB key containing the virus was inserted by an insider or a third party contractor.[122] Once inside the environment, Stuxnet infected not only the initial system but also other connected computer systems in its way eventually reaching the intended system. No action is necessary of the person operating the systems therefore knowledge of the virus will most likely be absent. The worm spreads itself in two ways. [123] The first was through the sharing of the infected USB key. The second way the virus spread itself was by using the local area network as a facilitator. It was done through the use of network shares and 2 specific exploits. Stuxnet was programmed to search for a computer with the task to coordinate a Programmable Logic Control, abbreviated PLC, of the brand Siemens. In particular a PLC containing SIMATIC WinCC/Step 7 controller software.[124] A PLC is a small computer with the ability to control industrial machineries such as centrifuges or pumps. One of the characteristics of the virus is that it sends information from the infected system back to its creator each time upon infecting a new system. The attacker can then decide if Stuxnet has to spread further in search of finding a computer that is able to program the target Siemens PLC or whether it has to cease its search entirely. In case Stuxnet locates its target, it collects the information stored in the so-called design documents. These are documents that contain information describing how the infected industrial control system works. After analysing the information, the attacker is then able to send a code that has a certain effect on the target. In case of the power plant it led to centrifuges spinning out of control. The virus then intercepts signals and replaces them by its own commands. These commands control the amount of hertz to which the drives spin. The Stuxnet worm targeted frequency drives spinning at somewhere between 807 and 1210 Hz revealing the ultimate goal of Stuxnet: reprogramming specific industrial controllers that could only be found in the Natanz nuclear plant. The Stuxnet worm made frequency drives spin out of control thereby destroying gas centrifuges needed for the enrichment of Uraninum.[125] The Stuxnet virus had a built-in 'kill date' that was set to June 24, 2012. The Stuxnet worm ended up infecting more than 80.000 computers worldwide and was found in countries such as

[122] Falliere et al. 2011, p. 3.
[123] *Ibid* pp. 1-55.
[124] B. Schneier, 'The Story Behind The Stuxnet Virus', October 7 2010 http://www.forbes.com/2010/10/06/iran-nuclear-computer-technology-security-stuxnet-worm.html
[125] Falliere et al. 2011, p. 1.

the United States, Indonesia and India besides Iran.[126] Over 40% of the computers attacked by Stuxnet were located outside Iran.[127]

§5.4 What are the effects of the Love Bug and Stuxnet viruses and what implication do they have on the principle of distinction?

The effects of an attack conducted by a means of cyber warfare provide factual, even more so, it provides measureable data. This subsection analyses whether the attack in the respective cases can be carried out with respect for the principle of distinction. The analysis is done through previously determined characteristics of interest that could potentially affect the way the principle of distinction can be upheld. The first variable is that of control. The central question in this section is in precisely what phase the attack can be terminated or redirected. The more dynamic a means of warfare is, the higher the possibility that unforeseen and unplanned scenarios can be taken into account thereby preventing the killings of civilians in case a situation changes last minute. In both cases the viruses will be analysed in order to see what kind of control is present and what practical implications the control has on the outcome. Second, the spreading nature of a worm will be closely examined. It will be analysed how the viruses spread and what implications it would have in case they were used as a cyber-attack, determining whether the spreading can occur without infringing the principle of distinction. First, an assessment of the control of the two means of cyber warfare will be outlined for the Love Bug virus. Second, the spreading nature and its implication on the principle of distinction will be reviewed. The Stuxnet virus will be analysed the same way afterwards.

§5.4.1 The Love Bug virus: an assessment.

In this subsection, the Love Bug virus will be analysed. It will be analysed whether there is an option to abort the attack in case the situation changes before the effects of the initiated attack occur and there is a need for it in order to prevent the targeted killings of civilians. Second, the inherent characteristic of the virus namely its spreading nature will be analysed to determine what kind of effect it has on the principle of distinction in terms of the ability to uphold it. The Love Bug virus is a fairly straight forward

[126] Richardson 2011, p.10.
[127] O'Connell 2012, p.202

virus. It has a predetermined script code that is set to do a variety of things once being executed. The main goal of the virus is overwriting image files and sending a copy of itself to contacts of the targeted and infected system through the use of Microsoft Outlook and mIRC. Software producer McAfee noted the attacks of the ILOVEYOU virus in a concise manner and it will be used as the outline for this subsection.[128] First, the virus copied itself several times and hid the copies in several folders on the victim's hard drive. Second, the virus added new files to the registry key of the computer system. Third, the virus copied itself replacing a multitude of files. Fourth, the virus sent itself through Internet Relay Chat clients as well as email by the use of Microsoft Outlook. Finally, the file WIN-BUGSFIX.EXE is downloaded.[129] In case the targeted computer downloaded the file and executed it, it would steal passwords and send them to an email address set up by the creator of the virus. When assessing the variable of control it has to be seen in what stages the initiator of the attack can terminate or redirect the conducted attack. The creator attached the virus in a document and as soon as that specific file was executed the virus replicated itself. The initial targeting of the Love Bug virus can be done with accuracy since the virus will be attached to the email that is sent to one's email address. As soon as the attached file is executed, the control over the attack is at a minimum. The program code has a predetermined script and as soon as the file is executed there is no possibility to overturn the decision. The virus will replicate and hide itself on the targeted system. It will search for Microsoft Outlook and mIRC, it will then be sent to several contacts depending on whether the targeted computer system has the software installed or not. The browser of the system will then be redirected to ensure a download of the malicious file WIN-BUGSFIX.EXE. Passwords will be stolen once the file WIN-BUGSFIX.EXE is downloaded and executed. There is no option for the attacker to intervene once the file has been executed by the targeted system. The passwords collected will be at the disposal of the initiator of the virus. He can ultimately decide what to do with the collected information. The spreading nature of the Love Bug virus will now be assessed. Upon being executed the virus will first replicate itself on the infected system and search for Microsoft Outlook and mIRC. The virus will spread itself based on whether the target

[128] J. Strickland, '10 Worst Computer Viruses Of All Time', *Howstuffworks* http://computer.howstuffworks.com/worst-computer-viruses2.htm
[129] *See*: http://www.f-secure.com/v-descs/love.shtml under Threats Email-Worm:VBS/LoveLetter

computer has an operation system running mIRC or Microsoft Outlook. Those are the deciding factors in relation to the spreading of the virus. The spreading will then occur based on those variables and differentiation is not possible. Making a distinction based on military or civilian is therefore impossible. The facts of the Love Bug Virus case state this since multiple organizations, public and private person were experiencing the effects of the virus. The characteristic of a virus being a programmed script that is set to follow a certain instruction is detrimental in this case in terms of the principle of distinction being upheld. The variables that would separate a military from a civilian computer would be differently. Setting the variables in order for the virus to target a civilian computer is a challenging task as complex variables would be necessary to effectuate the result. The virus has insufficient possibilities to distinguish a civilian from a military objective. The Love Bug virus is not capable of making a distinction between military and civilian targets and therefore infringes the principle of distinction.

§5.4.2 The Stuxnet virus: an assessment.

In this subsection the Stuxnet virus will be assessed. First, it will be determined what type of control the virus has. It will be analysed whether there is an option to abort the attack in case the situation changes right before the effects of the initiated attack occur and there is a clear need for it in order to prevent the targeted killings of civilians. Second, the inherent characteristic of the virus namely its spreading nature will be analysed to determine what kind of effect it has on the principle of distinction in terms of the ability to uphold it. Due to the complexity and the virus specific characteristics, an analysis could easily exceed the limit of pages of this thesis especially since some believe that the virus changed over time.[130] This analysis treats the attack of the virus as one consistent attack. The scenario and summary used are mainly that of software producer Symantec and is based on the technical features of the Stuxnet virus.[131] The Stuxnet virus was primarily set to target an industrial control system(ICS). Industrial control systems are being used in gas pipelines and power plants and the ultimate goal was to damage and or destruct the centrifuges enabling the creation of highly enriched nuclear uranium which is an essential component of nuclear weapons.[132] Depending on what exactly constitutes an attack, there are particular

[130] Falliere et al. 2011, p. 4.
[131] *Ibid* p. 3.
[132] Richardson 2011, p. 7

key moments of relevance. There is a chain leading up to the destruction of the centrifuges. The initial contact of the virus to a system that began the spreading is a first moment of interest. The main target(s) were not accessible from outside the network. Targeting here can occur with precision. Two ways of first contact are possible even though one is hypothetical in the case of Stuxnet. A first possibility of the system being infected is through a (remote) network and a second possibility is an infection done by a system belonging to a local network. In the case of Stuxnet, the target could not be remotely accessed so the initial contact had to come from within the closed network. In order for the virus to be able to attack the industrial control system it had to be introduced in the environment of the target.[133] This can be done in a precise manner. One can do this by direct manual input in the environment respecting the principle of distinction. In the Stuxnet case it was likely done by the use of a removable storage medium. This form of targeting can be done with accuracy and can be aborted up until the very last moment which means in the Stuxnet case cancelling inserting the storage medium. After the introduction the virus executes and spreads in search for computers used to program PLCs.[134] The Stuxnet virus has a command and control server that can establish a connection between the attacker and the virus. After the threat has installed itself, dropped its files, and collects information about the specific system it infects, it contacts the command and control server on port 80. It then transfers basic information about the compromised computer to the attacker using HTTP.[135] The attackers are notified what type of software the targeted system is running and can base their decision to attack on. Stuxnet provided a backdoor functionality making it a possibility for the attackers to download and execute additional tools next to updating the version of Stuxnet.[136] Also, the Stuxnet virus had a built-in destruction date set.[137]All these possibilities together create an environment with possibilities to manage the attack. The virus functionality can be altered and updated and a decision to abort an attack can be made late into the

[133] Falliere et al. 2011, p.3.
[134] *Ibid*.
[135] Langner 2011, 'Even though Langner mentions that the Stuxnet attack was not remotely controlled referring to an attack being conducted and being controlled in a realtime manner, there was a connection established betwen the creators of the virus and the computer system. He underlines that the command and control servers have been used with the primary reason to confirm that a system is compromised. Falliere 2011 p. mentions that a connection is unlikely since a significant amount of comprised systems did no have internet access.
[136] *Ibid.*
[137] Falliere 2011 p. 29.

attack itself because of the command and control function that connects the attacker to the virus. The spreading nature of the Stuxnet virus will now be assessed. The targeting of the PLC controllers make it a necessity for a highly sophisticated way of spreading that would be different from what is seen at the Love bug virus. The main reason is that the target is most likely not connected to a network that can be reached from outside a designated set of computer systems. This makes it more difficult for a virus to carry out its task successfully in the more common way seen at other viruses. It is likely that the characteristics in relation to the way of spreading of the virus were influenced by unique characteristics of the target, in particular the characteristic lacking an outside connection to reach it. Stuxnet copied itself onto removable drives and used tactics such as replacing the files on the drive by ones containing the virus in order to increase its chance for a successful spreading of the virus.[138] Although the worm goes about an extensive way to determine whether the targeted system is one the attackers are after by for instance, the checking of model numbers, configuration, details, and downloading program code from the controller to verify whether it is the correct program. It has to be noted however, that the worm will infect a computer if it runs the operating system of Windows upon contact. It does this regardless of the fact that this system platform was, and still is, predominantly used by civilians. A civilian computer system can either have the virus inserted into it through a local area network or by a storage medium containing the virus. When critically reviewing the Stuxnet virus one can see that there is a significant amount of built-in safeguards to ensure that the spreading keeps within limits. For instance, a USB storage unit could not spread the virus more than three times due to the virus spreading code.[139] Next to that, the virus had a built-in destruction date set to June 24 2012. It would erase itself from the infected system and disappear entirely.[140]

§5.5 A comparison between the ILOVEYOU virus and the Stuxnet virus.

In this subsection the main focus will lie on the differences between and similarities of the two cyber weapons in relation to the ability to uphold the principle of distinction. The conclusion whether a cyber-

[138] *Ibid* p. 32.
[139] M.J. Gross, 'A Declaration of Cyber War', *Vanity Fair* April 2011 p. 4. http://www.vanityfair.com/culture/features/2011/04/stuxnet-201104
[140] ***Ibid*.**

attack can be conducted while respecting the principle of distinction will be primarily based on the outcome of the analysis. At the elementary level both cyber-attacks use similar tactics because of their nature. The initial attack can be directed against a target in a precise manner. Both viruses focus on different aspects to prevent discovery after compromising the target. The ILOVEYOU virus hides itself on the computer system, replaces certain files and uses a social application to spread itself making it seem that the recipient opens a personal email. Stuxnet exploits a vulnerability in the system gaining entry undetected using a Windows rootkit that hides Stuxnet's binaries.[141] The Stuxnet virus places emphasis on ensuring that the PLC controlling the centrifuges will continue to carry out the rogue code which makes them spin in an uncontrolled and destructive manner. So another similarity is that both cyber weapons deceive the users of the compromised systems and hide on the computer system even though they do this clearly in different ways. The most notable similarity in relation to upholding the principle of distinction is however that they both spread indiscriminately. Both weapons of cyber warfare seem to infect a computer system regardless of the status whether this is civilian or military. The preprogramed code logic does not prevent civilian computer systems to be targeted. Based on these varieties of the cyber weapon worm/virus it cannot be denied that the spreading nature can be a threat to the principle of distinction. There is simply an insufficient amount of safeguards built-in to prevent civilian computers from being targeted. There are important differences that have to be noted however. The first difference relates to the spreading nature of both viruses. The creators of the Stuxnet virus seemed to have a higher focus on keeping the spreading of the virus from getting out of hand. Stuxnet could only spread three times from a storage device, up until a certain date after that the virus would terminate itself. There was a command and control function that allows for termination of the virus by the attackers in case there was a need for it. Possibly the most notable difference is that the Stuxnet virus upon infecting a system searches for a setup that is solely found in the computer system that manages the PLC controller, ultimately determining the frequency to which the centrifuges of the nuclear plant of Natanz spin. This does not prevent a civilian computer to be infected by the Stuxnet virus. It does strongly mitigate its effects however. Apart from some minor altercations on the infected system and

[141] Falliere 2011, p. 2.

their respective files there would be no irreparable damage done to those systems unless they were ultimately connected to the centrifuges by way of a program logic controller or PLC in a unique configuration. Falliere compares Stuxnet to a self-directed stealth drone. He states that Stuxnet is the first known virus that, released into the wild, can seek out a specific target, sabotage it, and hide both its existence and its effects until after the damage is done.[142] This statement acknowledges that Stuxnet has the ability to find and be directed against a specific target.

§5.6 The Stuxnet virus in the light of the principle of distinction.

The principle of distinction makes it mandatory for parties to an armed conflict to distinguish between the civilian population and combatants and between civilian objects and military objectives. It does this in close conjunction with the principle of subsidiarity and proportionality. This subsection connects the Stuxnet virus to the principle of distinction to establish an overview of its characteristics to motivate why Stuxnet does not infringe the principle of distinction. An abstract interpretation of the principle of distinction could therefore be that the principle of distinction is limited to making a distinction between the two categories. This interpretation is however, after researching the principle of distinction, not a conclusive one. The aim that the principle of distinction seeks to achieve, and ultimately the foundation the principle of distinction is built on, is about sparing civilians and the civilian population from hostilities and their effects. It therefore transcends a minimalistic interpretation of the principle of distinction limiting only to distinction itself. Hostilities and their effects are therefore as crucial as the distinction by itself in determining whether a cyber-attack can be conducted while still respecting the principle. The ICJ in the *Nuclear Weapons* advisory opinion reiterates that States must never make civilians the object of attack and States must consequently never use weapons that are incapable of distinguishing between civilian and military targets.[143] While it is true that civilian objects were ultimately involved in the search of Stuxnet to allocate the centrifuges in the Natanz nuclear plant, the effects of the attack occurred at the intended target: the Natanz nuclear plant. Moreover, it set back the nuclear enrichment program of the Iranian government. After closely examining the Stuxnet virus one

[142] *Ibid* p.4.
[143] ICJ Nuclear Weapons para. 78.

can conclude that this version had certain measures built-in indicating that civilian objects were not the intended target. In case the target was not identified as being the PLC system, the virus limited its spreading to three times. After the three times, the specific branch of infection stopped spreading and terminated itself. Also, the attacker could control the virus due to a command and control function that connected the attack to the virus and moreover, the attacker could decide to abort the attack. Another measure was that of a termination date. After June 24, 2012, the Stuxnet virus ceased to exist and the effects were limited to the intended target thereby sparing the civilian objects it encountered on the way to the target. It is because of those indications, or measures even, that neither civilians nor civilian objects were the intended object of attack. The object of attack was the Natanz nuclear plant, more in specific its centrifuges. Regardless of it being a virus or worm that due to its characteristics potentially infringes the principle of distinction per se, the Stuxnet virus was directed against a target in accordance with the principle of distinction provided that one classifies the Natanz security plant to be a military target. The Stuxnet virus therefore displays that the cyber-attack that is a worm/virus can be conducted while respecting the principle of distinction. A cyber-attack can therefore be conducted in conformity with the principle of distinction.

Conclusion.

There are different cyber-attacks possible. The worm and virus have been the focus of this thesis due to their inherent characteristics that pose a greater challenge to the principle of distinction than other forms of cyber-attacks. The worm and virus are traditionally less capable to distinguish between civilian and military objectives. Their indiscriminate spreading leads to the targeting of military and civilian objectives. Control of an attack ensures a greater possibility that the principle of distinction is respected. The possibility to abort an attack in a late stage before the effects of an attack occur gives the type of control that is important to prevent violations of the principle of distinction from occurring. The Love Bug virus may not have been conducted as a part of an international armed conflict, it does illustrate the characteristics of a worm/virus and the damage it can cause to military and civilian objectives making it a viable and potential means in an armed conflict. Neither has the Stuxnet virus been officially conducted as a cyber-attack as a part of an international armed conflict but the target and the context makes it relevant for this research. The Stuxnet virus displayed a precision not seen at other viruses and or worms. The Stuxnet virus spread in a traditional way by using civilian and military computer systems, however there was virtually no damage done to civilian objectives. The target was the Natanz nuclear plant in Iran and the damage was limited to the centrifuges setting back the Iranian nuclear enrichment program. This displays that it is possible to conduct a virus and or worm attack that at first sight might violate the principle of distinction since it affects civilian and military objectives alike, the damage of such an cyber-attack is however limited to a specific military objective. If such an attack would have been conducted in an international armed conflict there would be no actual damage to civilian objectives. A cyber-attack can therefore be conducted in conformity with the principle of distinction in armed international armed conflict.

Literature list:

Sun Tzu 2002
S. Tzu**,** *The Art of War***,** Mineola: Courier Dover Publications.

Clarke & Knake 2010
R.A. Clarke & R.K., *Cyber War: The Next Threat to National Security and What to Do About It*, New York: Harper Collins Publishers.

O'Connell 2012
M.E. O'Connell, International Law and the Use of Force, Eagan: Foundation Press.

Shackelford & Andres 2011
Shackelford & Andres, 'State Responsibility for Cyber-attacks: Competing Standards for a Growing Problem', *Georgetown Journal of International Law* 4, no. 4.

Gasser 2013
H.P. Gasser, *International Humanitarian Law: an Introduction*, in: *Humanity for All: the International Red Cross and Red Crescent Movement*, Berne: Paul Haupt Publishers.

Hathaway et al. 2012
A. Hathaway et al. ,'The law of Cyber-Attack', *Califonrnia Law Review* vol.100 no.4.

Carr 2011
J. Carr, *Inside Cyber Warfare: Mapping the Cyber Underworld*, Sebastopol: O'Reilly Media Inc.

Watkin 2010
K. Watkin, Opportunity Lost: Organised Armed Groups and the ICRC "Direct Participation in Hostilities" Interpretive Guidance, 42 *N.Y.U. J. INT'L L. & POL.* 641, 643-44.

Neff 2010
S.C. Neff, *War and the Law of Nations*, Cambridge: Cambridge University Press 2008

Solf 1986
W.A. Solf,' Protection of Civilians Against the Effects of Hostilities Under Customary International Law and Under Protocol I', *American University International Law Review*, vol. 1 issue 1, 117-135.

Falliere et al. 2011
N. Falliere et al., 'W.32 Stuxnet Dossier', *Symantec Security Response*, version 1.4.

Hayashi 2010
N. Hayashi, 'Requirements of Military Necessity in International Humanitarian Law and International Criminal Law', *Boston University International Law Journal* vol. 28:39.

Vatis 2001
M. A. Vatis, 'The war one terrorism: a predictive analysis', *Institute for Security Technology Studies at Darthmouth College* 2001.

Forrest 2007
C.J.S. Forrest, 'The Doctrine of Military Necessity and the Protection of Cultural Property during Armed Conflict', *California Western International Law Journal v*ol. 37 no. 2.

Dijkhoff 2010
K. Dijkhoff, *War, Law, and Technology*, Oisterwijk: Wolf Legal Publishers.

Schmitt 2013
M.N. Schmitt, 'Cyberspace and International Law: The Penumbral Mist of uncertainty', *Harvard Law Review Forum*, Vol. 126:176.

Schmitt et al. 2013
M.N. Schmitt et al., *Tallinn Manual on The International Law applicable to Cyber Warfare*, New York: Cambridge University Press.

Schindler 1979
D. Schindler, 'The different Types of Armed Conflicts According to the Geneva Conventions and Protocols', *RCADI*, Vol. 163, 1979-II.

Henckaerts & Doswald-Beck 2005
J.M. Henckaerts & L. Doswald-Beck , *Customary International Humanitarian Law*, Cambridge: Cambridge University Press Volume 2 Volume I.

Richardson 2011
J.C. Richardson, 'Stuxnet as Cyberwarfare: applying the Law of War to the Virtual Battlefield', *Social Science Research Network Working Paper* 2011.

Estreicher 2011
S. Estreicher, 'Privileging Asymmetric Warfare: The Proportionality Principle Under International Humanitarian Law', *New York University Public Law and Legal Theory Working Papers* 2011.

Tsagourias 2010
N. Tsagourias, 'Necessity and the Use of Force: A Special Regime', *Netherlands Yearbook of Intl Law* vol. 41.

Watts 2011
S. Watts, 'Recprocity and the Law of War', *Harvard International Law Journal* Vol. 5.

McCormack 2006
T. McCormack, 'An Australian perspective on the ICRC customary international humanitarian law study' in A. Helm (ed), *The Law of War in the 21st Century: Weaponry and the Use of Force* 2006.

Dipert 2010
R.R. Dipert, 'The Ethics of Cyberwarfare', *Journal of Military Ethics*, vol. 9, No. 4.

Crawford 2011
E. Crawford, 'Levée En Masse – A Nineteenth Century Concept in a Twenty-First Century World', Legal Studies Research Paper Sydney: University of Sydney No. 11 31.