Evaluating the effects of cyber-attacks on critical infrastructures in the context of Tallinn Manual

Kosmas Pipyros, Lilian Mitrou, Dimitris Gritzalis

Information Security & Critical Infrastructure Protection (INFOSEC) Laboratory
Dept. of Informatics, Athens University of Economics & Business
76 Patission Ave., Athens, GR-10434, Greece

{pipyrosk,dgrit}@aueb.gr, l.mitrou@aegean.gr

Abstract

The Internet of Things (IoT), in which billions of things such as devices, applications and networks, are interconnected to each other and are capable to interact without any human intervention, is transforming the way of life. Public safety, transportation and communication, energy, healthcare, logistics, government and education are using cute-edging technologies such as intelligent sensors, wireless communication, cloud computing and data analysis techniques, with a variety of applications within different infrastructures in order to provide instant access to information, to communication and to create new economic opportunities. However, the more Critical Information Infrastructures (CII) are becoming independent the higher the vulnerability of states and the well-being of societies and citizens. The increasing number and complexity of cyber-attacks on state's CII in recent years has been transforming cyberspace into a new battlefield where "the mouse and the keyboard being the new weapons" bringing out "cyber warfare" as the "5th dimension of war". In this paper a systematic modeling methodology is presented for evaluating the effects of cyber-attacks on state's CII in the context of Tallinn Manual. The analysis is focused on the United Nations Charter's normative scheme of the "use of force" in order to define whether these attacks constitute a wrongful "use of force" under the principles of international law. In order this to be achieved the qualitative criteria for recognizing the impact of cyber-attacks as proposed by the International Group of Experts (IGE) in the Tallinn Manual has been used. Furthermore, Multi-attribute Decision Making (MADM) methods are applied. More specifically, pros and cons of the Simple Additive Weighting (SAW) method and the Weighted Product Method (WPM) lead us to present a new evaluation strategy which combines the use of the abovementioned decision making algorithms and introduces a new grouping of IGE's qualitative criteria based on their distinctive features. The methodology is been applied in real-life cyber-attacks, namely the large-scale attacks against Estonia and Iran, and cyber-attack evaluation results are been presented. The correlation of both qualitative and quantitative methods of analysis allows to achieve an improved cyber-attack evaluation assessment and as a result a more accurate and complete cyber-attack classification. The usefulness of the methodology is perceived in areas where there is uncertainty or disagreement in a number of legal analyses, and for making available a means for addressing all issues having to do with "use of force". Finally, the methodology could act as a basis for the assessment and classification of cyber-attacks that are intended towards Software-Intensive (SI) systems, component of a state's CII.

Keywords: Tallinn Manual, International Law, Use of Force, Multi-attribute Decision Making (MADM) methods, Cyber-attack.

1 Introduction

We are living in the cyber-age. In the 21th century the rapid development of Information and Communication Technologies (ICTs) have fundamentally transformed the global economy and the way of live. Public safety, transportation and communication, energy, healthcare, logistics, government and education are using cute-edging technologies such as intelligent sensors, wireless communication, cloud computing and data analysis techniques, with a variety of applications within different infrastructures in order to provide instant access to information, to communication and to create new economic opportunities. The internet of things (IoT) in which billions of things such as devices, applications and networks, are interconnected to each other and are capable to interact without any human intervention, is altering the perspective of our lives by providing a new world full of possibilities to help advance prosperity. However, the more the systems, infrastructures, societies and economies are becoming independent from human intervention the higher their vulnerability and the complexity to deal with new risks and threats that menace the sovereignty of states and the well-being of societies and citizens. In recent years, the increasing number and complexity of cyber-attacks on states' CII has been transforming cyberspace into a new battlefield where "the mouse and the keyboard being the new weapons" bringing out "cyber warfare" as the "5th dimension of war".

Haizler (2017) divided the United States' cyber warfare history in three evolutionary stages: a) the "realization phase" during the early era of the internet b) the "takeoff phase" during the interim period of pre- and post- 9/11 in which attacks were still mainly of an information-gathering nature and c) the "militarization phase" during which cyber warfare may cause similar damage to a state's strategic capabilities and critical infrastructure as a kinetic attack on a colossal level. From the "Morris worm" (Eisenberg et al, 1989) of the "realization phase" and the "Moonlight Maze" (Elkus, 2013) of the "takeoff phase" we are moving forward to the modern "militarization phase" with the "Stuxnet" worm (Langner, 2013) to be the most sophisticated malware attack publicly recorded.

In order to defend the US from cyber-attacks, former US President Barack Obama declared America's digital infrastructure as a strategic national asset (The Economist, 2010). Such decision reflect on the need to address the challenges posed with regard to cyber-attacks that could be qualified as cyberwar actions. Furthermore, Leon Panetta, former US Secretary of Defense, during his speech "Defending the nation from cyber-attacks" in 2011, pointed out that this is a pre- 9/11 moment and that a cyber-attack perpetrated by nation states or violent extremists groups could be as destructive as the terrorist attack on 9/11. The continuous increase in both the number and the intensity of cyber-attacks on states' CII renders the research on defining and evaluating these categories of cyber-attacks into a pressing need.

The threshold question relates to the adequacy and suitability of the existing "old" rules of warfare (*jus ad bellum* and *jus in bello*) that developed over generations to be applied on attacks using kinetic weapons and armed violence, to control the "brave new world of cyber warfare (Jolley, 2013). Views on the subject range from a full application of the law of armed conflict, along with the lines of the International Court of Justice's pronouncement that applies to "any use of force regardless of the weapons employed" (International Court of Justice Advisory Opinion on the Legality of the

Threat or Use of Nuclear Weapons , 1996), to strict application of the Permanent Court of International Justice's pronouncement that acts which are not forbidden in international law are generally permitted (Lotus Case, 1926).

We have to bear in mind that terms themselves, such as CII, are steadily evolving due to the impacts of the advancing domination of online communications and cyberspace on the "real world" and ubiquitous computing. The difficulties to define and to identify the effects and impacts of a cyber-attack in order to be equated to "use of force" or to an "armed attack" are obvious: if in the "traditional" *jus ad bellum* framework emphasis is given on human and/or material destruction, authors are arguing also for "unavailability" of CII as equivalent criterion (Tsagourias, 2012). Despite the progress made on regulation and research level to address the issues raised, there are still significant gaps in reaching a safe and definitive approach on when a cyber-attack constitutes "use of force" or when the right to self-defence should be recognized (Robinson et al, 2015).

The paper contributes to the development of a systematic modelling methodology for evaluating the effects of cyber-attacks on states' CII in order to answer the question of whether these attacks have risen to the level of "use of force" under the jus ad bellum, that body of international law that governs a state's resort to force as an instrument of its national policy. The threshold inquiry is crucial to assessing the level of violence between states in order to justify a lawful response. Because the UN Charter prohibits the unauthorized "use of force", a state must be able to quickly and safely assess whether a cyber-operation constitutes a "use of force" triggering the international condemnation and economic sanctions, (active) "cyber self-defense" - or an "armed attack" (with the use of conventional military weapons) as forceful response.

This is primarily achieved by adopting the "effects-based" or "consequences-based" approach, which focuses on the overall effect of a cyber-operation to the victim-state, as well as by using the qualitative criteria for recognizing the impact of cyber-attacks as proposed by the IGEs in the Tallinn Manual. Furthermore, Multi-Attribute Decision Making (MADM) methods are also applied.

The pros and cons of each MADM method are critically evaluated. The weaknesses of each MADM method lead us to present a new cyber-attack evaluation methodology that combines the use of decision making algorithms of MADM methods and introduces a new grouping of the IGEs qualitative criteria based on their distinctive features. The methodology is applied in real-life cyber-attacks, namely the large-scale attacks against Estonia and Iran, and cyber-attack evaluation results are presented. The correlation of both qualitative and quantitative methods of analysis lead us to achieve an improved cyber-attack evaluation assessment and as a result a more accurate and complete cyber-attack classification.

The paper is organized as follows: In section 2 the large-scale cyber-attack incidents against Estonia and Iran are introduced. Those cyber operations will be used for the application of the proposed cyber-attack evaluation methodology. In section 3 a useful review of the legal issues, focusing on the uncertainties when dealing with cyber-attacks using the regulatory framework of international law, is presented. Then, cyber operations are being categorized, based on their intensity and the qualitative criteria, as proposed by the IGEs` on the Tallinn Manual, are described. In section 4 the descriptions of both Simple Additive Weighting (SAW) method and Weighted Product Method (WPM) are presented. Both methods are introduced as multi-criteria decision

analysis ones for the evaluation of cyber-attacks. In section 5 the pros and cons of each method lead us to propose a new cyber-attack evaluation methodology which includes both qualitative and quantitative methods of analysis and results to a more accurate and complete cyber-attack evaluation and classification. Furthermore, the proposed methodology is been applied in real-life cyber-attacks, namely the large-scale attacks against Estonia and Iran and cyber-attack evaluation results are presented. Finally, in section 6 the indicative results of the research are critically analyzed.

2 The large-scale cyber-attack incidents against Estonia and Iran

One of the most famous wide-range attacks that ever took place was the cyber-attack in Estonia in April 2007 that lasted for almost three weeks. The cyber-attacks seemed to be the result of the Estonian Government's decision to defy Russian threats and to remove from the city center the Bronze Soldier monument, a memorial of the Soviet liberation of Estonia in World War II; a decision that caused diplomatic tension with Russia and a wave of protests in central Tallinn (Tikk et al., 2010). Reactions were transferred from the streets to cyberspace with "patriotic hackers" engaging in a coordinated large-scale cyber-attack, which was directed against Estonia's critical ICTs.

The choice to shift the "battle field" from the streets to cyberspace was not random. As early as the mid-1990s, Estonia had been characterized as an e-state, as all of its critical services were provided to its citizens through the Web: e-banking, health care and e-government services, as well as internet voting for the national elections. These attacks were meant to harm the functionality of the state, causing a number of adverse effects to the operation of public administration and the economy, leading to the destabilization of the country's financial system and threatening its national security (Tikk et al., 2010). The specific assault quickly led to the cultivation of fear among citizens and of a sense that nothing was functioning in the country. The total aim of the cyber-attacks was to undermine Estonia's social cohesion (Black, 2008).

The cyber-attacks included Denial of Service (DoS), Distributed Denial of Service (DDoS), defacement and destruction against government and commercial internet infrastructure such as banks, internet service providers, the government offices of the President and the Prime Minister, several ministries and critical public sector institutions. Most of the attacks were came from Russia, initially from private IP addresses, although experts tracked a number to Russian government institutions. There is no firm evidence that the Russian government guided or managed the attacks and as a result there is uncertainty whether the attacks were launched with the government's knowledge (Schmitt, 2011).

The case in question was clearly an unprecedented act of psychological terror. The cyber-attacks demonstrated in their full range the close interrelation that exists between cyber security and national security and the key role that the former plays in ensuring a country's social stability and the prosperity of its citizens. At the same time, it revealed the insufficiency of the European security institutions such as the EU, NATO and the Security Council to stand by Estonia while the country was under cyber-attack. Moreover, it proved extremely difficult to evaluate the impact of the cyber-attacks and to characterize the incident as a matter of law.

A few years later, in June 2010, a malicious computer worm named "Stuxnet" struck the Iranian nuclear facility at Natanz and infected over 50,000-100,000 computers and

damaged up to 2,000 of Iran's centrifuges used in the enrichment of uranium (Farwell and Rohozinski, 2011). The Symantec Security Response Team (W.32 Stuxnet Dossier, 2011) which reverse-engineered the worm and issued a detailed report on its operation, characterized Stuxnet as "the first of many milestones in malicious code history - one of the most sophisticated and unusual pieces of software ever created". This was recorded as the first cyber operation to exploit four zero-day (unpatched) vulnerabilities, compromised two digital certificates, injected code into industrial control systems and hided the code from the operator. The worm's code was approximately 500 Kbytes, fifty times as big as a typical computer worm and was written in multiple languages (Virvilis and Gritzalis, 2013). Additionally, it was the first instance of cyber operation known to cause physical damage across international boundaries. Stuxnet worm was of such complexity –requiring significant resources to develop – that few attackers would be capable of producing a similar threat. Its sophistication suggests that the creators had deep knowledge of its target and access to immense resources, perhaps with governmental support. The choice of the particular targets also reveal a political motive (Collins and McCombie, 2012).

The Stuxnet worm was highly selective not only about its targets but also in specific conditions on the targets. Unlike earlier worms which did not have physical consequences, it appeared to aiming directly to take control of critical physical infrastructure. It reportedly attacked Windows computers looking for a particular Programmable Logic Controller (PLC) made by Siemens on the vulnerable computers. Moreover, it waited for a specific program condition before it would attempt to take over control by manipulating some of the settings. Stuxnet worm was estimated to have infected 50,000–100,000 computers, mainly in Iran, India, Indonesia, and Pakistan. The Iranian government admitted that Stuxnet had set back the nuclear program but it only affected a limited number of centrifuges. Furthermore, it acknowledges that Stuxnet struck 12 industrial plants, both in and out of Iran (Karnouskos, 2011).

Ralph Langner (2011), a German security expert familiar with industrial systems security, expressed his belief that Stuxnet was the first real "cyber weapon" because it aimed to attack a physical-military target. Iran suspected U.S. and Israel involvement behind Stuxnet, although both have denied responsibility. Notwithstanding, according to the New York Times (Sanger, 2012), the worm was part of a sustained US campaign of cyber operations against the Iranian nuclear program known as "Olympic Games". The program began during the George W. Bush administration and accelerated under Obama. It featured collaboration with Israel for both operational and strategic reasons: the United States needed access to Israeli clandestine intelligence networks in Iran, and the United States wanted to dissuade Israel from launching an airstrike against Iran. The actual technical work was carried out by the US National Security Agency (NSA) and Israel's Unit 8200 and the attack was rehearsed at Israel's "Dimona" nuclear facility.

Iran downplayed the Stuxnet attack as a failure. Although the full extent of damage caused by Stuxnet is unknown, the cyber operation on Iran's uranium facilities had set back the nuclear program by at least two years, according to the Iranian government. Taking into account all the facts it can be claimed that Stuxnet was more successful than a kinetic military strike, as it produced the same results but avoided casualties and averted a full scale war.

But how these large-scale cyber-attack incidents will be treated under the prism of international law? Could it be argued that they represent a use of force, armed attack or aggression according to the UN Charter? Unfortunately, the existing legal norms do not offer a clear and comprehensive framework within which states can shape policy responses to the threat of hostile cyber operations. This is why in the following section we will focus on the legal uncertainties when dealing with cyber operations and we will examine the transformation of the international law rules to face the "brave new world of Cyber warfare".

3 The transformation of the traditional rules of international law to control the "brave new world of Cyber warfare".

Cyber warfare is inherently "international" in nature and requires an international legal response (Morth, 1998). However, the only international legal instrument regulating comprehensively cyberspace remains the "International Convention of Cybercrime" or the "Budapest Convention" (2001) which led to the creation of a reference framework aiming to address computer and internet crimes by introducing appropriate legislation and fostering international cooperation for law enforcement and exchange of respective information between government and the private sector. Nevertheless, it was not the purpose of the Convention to introduce a legislative framework for cyber warfare.

More recently, the Directive 2016/1148 of the European Parliament and of the Council "concerning measures for a high common level of security of network and information systems across the Union" laid down obligations for all Member States to adopt a national strategy on the security of network and information systems. Additionally, according to the Directive, all Members of the Union must designate national competent authorities, single points of contact and computer security incidents response teams (CSIRTs) in order to establish security and notification requirements for operators of essential services and for digital service providers. The main goal of the Directive was to contribute to the development of trust and confidence between Member States and to promote swift and effective operational cooperation so as to improve the security and functioning of the internal market.

However it is doubtful if the above-mentioned obligations and rules can deal adequately and effectively with the challenges posed by cyber-attacks that are to by qualified as cyberwar acts (Jougleux et al., 2016). Furthermore, there are no specific rules of international law governing the international use of cyber force (Roscini, 2014). Notwithstanding, this uncertainty surrounding cyber legislation does not mean cyber operations are taking place in a normative void. The first non-binding document that attempted to address cyber-attacks using the instrumentarium of international law and to produce a manual on the law governing cyber warfare was produced in 2013.

The "Tallinn Manual on the International Law Applicable to Cyber Warfare" or "the Tallinn Manual" (Schmitt, 2013) was a project launched by international law practitioners and scholars at the invitation of the NATO Cooperative Cyber Defence Centre of Excellence (NATO CCD COE), in an effort to examine how extant legal norms applied to this "new" form of warfare. The main goal of the Tallinn Manual was to clarify the complex legal issues surrounding cyber operations, with particular attention paid to those involving the jus ad bellum, the body of international law that governs a state's

resort to force as an instrument of its national policy, and the jus in bello, the international law regulating the conduct of armed conflict (also labelled the law of war, the law of armed conflict or international humanitarian law). In the Tallinn Manual, the International Group of Experts came to the unanimous conclusion that the general principles of international law should also apply to cyberspace. Its task was to determine how exactly this type of law can be applied and to identify any cyber-unique aspects thereof. The rules set forth in the Tallinn Manual provide specific provisions (Rules) on the topic intending to act as customary international law.

More recently, on February 2017, NATO CCD COE proceeded to the publication of the second version of the "Tallinn Manual on the International Law applicable to Cyber warfare (Tallinn Manual 2.0)", which deals with cyber-attacks that are evaluated below the threshold of the "use of force" and which are carried out during a period of peacetime. The particular handbook presents a critical review of the evolution of cyber-attacks over the past decade and extends the existing reflection of how to apply the traditional rules of international law into the cyber context by shifting the gravity center from the large-scale cyber-attacks to more limited, in terms of scale and effects, cyber-attacks that essentially constitute the majority of the attacks we face on a regular basis.

Yet, in spite of the noticeable progress at international level, toward the updating of international law rules to effectively address this "new" form of warfare, there is still confusion regarding the degree of the application of international law rules to cyber warfare. Namely, it has not yet been clarified in which cases do cyber-attacks constitute a "threat or use of force" so that the prohibition of Article 2(4) of the UN Charter can apply, neither is it clear in which cases these aggressions can be treated as an "armed-attack", making it possible for a UN Member State to respond by exercising its legitimate right of self-defense under Article 51 of the UN Charter.

3.1 Cyber operations classification from the perspective of international law.

If we try to portray cyber operations, based on their intensity, we would say that they can be categorized as follow:

(a) The lowest level of intensity of cyber operations incorporates those cyber-attacks which provoke nothing else than mere inconvenience for the state's functionality. Those cyber-attacks certainly do not consist "use or threat of force" in violation of international law since they do not have any serious impact for the victim state.
(b) The second level of intensity of cyber operations incorporates those cyber-attacks reaching the level of "use of force". As foreseen in article 2(4) of the UN Charter "all 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" which means that uses or threats of force that endanger national or international stability fall within article 2(4)'s prescriptive envelope (Schmitt, 1999).
(c) The third level of intensity of cyber operations incorporates those cyber-attacks in which the Security Council in actively involved by taking all the necessary measures in order to maintain or restore international peace and security.
(d) The highest level of intensity of cyber operations incorporates those cyber-attacks reaching the level of an "armed attack". In these cases there is an inherent right of

self-defence under Chapter VII of the UN Charter. Fig. 1 illustrates the level of intensity of cyber operations according to the provisions of the UN Charter.

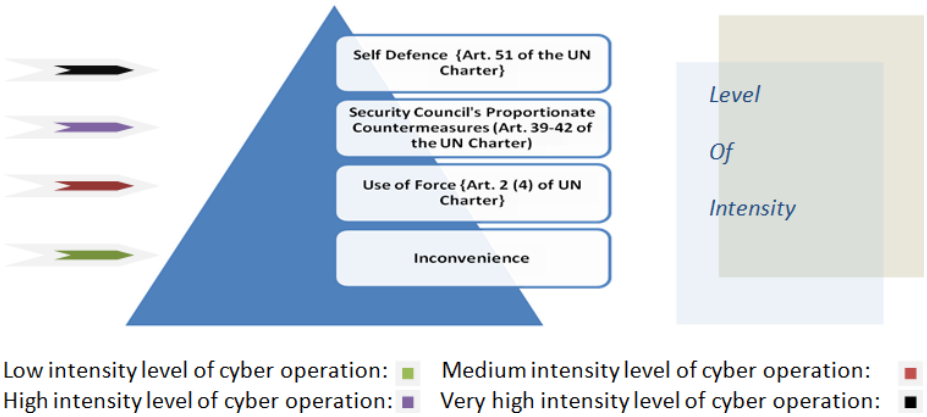

Fig. 1. *Level of intensity of cyber operation*

Nevertheless, in the cyber context, the identification of the type of conflict to which particular hostilities apply, as a matter of law, is proving extremely problematic. The difficulty in applying the traditional rules of international law, so as to deal effectively with cyber-attacks, stems from a number of factors. The most important of them is the failure to estimate properly the impact of cyber-attacks on the victim-state and on the international environment (Pipyros et al., 2016).

3.2 The qualitative criteria for cyber-attack evaluation

Taking into account the total absence of an institutional framework for the evaluation of the "use of force" and "armed attack" concepts in cyberspace, the International Group of Experts in the Tallinn Manual proceeded to the adoption of an approach, following Schmitt's consequences-based approach (Schmitt, 1999), that aims objectively to identify, evaluate and characterize a cyber operation.

This approach focuses on recognizing the impact of cyber-attacks and on equating it to the corresponding impact caused by other actions (non-kinetic or kinetic) that the international community would describe as "uses of force". In these cases, the parallelism and the subsequent analogous treatment of conventional operations that verge on being characterized as "uses of force" will be the outcome of the evaluation of non-exclusive criteria (factors) based on a case-by-case assessment. Table 1 provides the qualitative criteria, as proposed by the International Group of Experts.

The above-mentioned criteria do not have a binding nature. They are predictive tools, not normative standards and shall serve as indicators that states are likely to take into consideration when making "use of force" appraisals. Moreover, as Schmitt (1999) stated "the factors must operate in concert". As an example, a highly invasive operation that causes only inconvenience, such as temporary denial of service, is unlikely to be classified as "use of force". By contrast, a number of states may categorize massive

cyber operations that cripple an economy as "use of force" even though economic or political coercion is presumptively lawful.

Severity	Is determined by the scope, duration and intensity of the caused consequences of a cyber operation.
Immediacy	Refers to the speed at which consequences manifest themselves.
Directness	Examines the chain of causation.
Invasiveness	Refers to the degree to which cyber operations intrude into the target state or its cyber systems contrary to the interests of that state.
Measurability of effects	Refers to the fact that the more quantifiable and identifiable a set of consequences, the easier it will be for a state to assess a situation when determining whether the cyber operation in question has reached the level of a use of force.
Military Character	Is a nexus between the cyber operation in question and military operations that heighten the likelihood of characterizing a cyber-attack as a use of force.
State Involvement	Refers to the fact that the clearer and closer a nexus between a state and cyber operations, the more likely it is that other states will characterize them as uses of force.
Presumptive Legality	International law is generally prohibitive in nature. Acts that are not forbidden are permitted. Absent an express treaty or accepted customary law prohibition, an act is presumptively legal.

Table 1: *Qualitative criteria*

Furthermore, Schmitt (1999) divided the spectrum of the "use of force" context into three broad bands, one each for relatively clear cases of each qualitative choice, and a central "grey area" for factually uncertain determinations as shown in Figure 2. By applying the quantitative scale to each of the eight identified factors, any given operation could be described in qualitative terms as being closer to the one end of the spectrum or the other. In other words, an action's qualitative nature (in eight more or less binary areas) could be determined by applying any fixed quantitative figure (say, a one-to-ten scale). Schmitt's contribution in translating the qualitative Charter paradigm into its quantitative components - the legal equivalent of going from analogue to digital - provides a framework for scholars and practitioners to organize analysis in something other than a quantum cloud of subjective uncertainty.

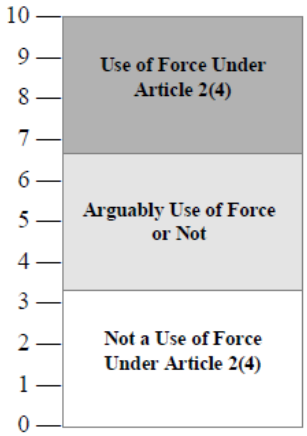

***Fig. 2**: The quantitative scale for cyber-attacks evaluation by M. N. Schmitt*

In the following section a systematic modelling methodology is presented aiming to evaluating the effects of cyber-attacks on states CII in order to answer the question of whether these attacks have risen to the level of "use of force" under the principles of international law. In order for this to be achieved two approaches are taken into consideration. First, the use of the IGE's approach, which is a transformation of the current Schmitt's consequence-based approach (Tallinn Manual, 2013). More specifically this approach has been differentiated by the replacement of the factor "Responsibility" with the factor "State Involvement", which however has similar conceptual and semantic interpretation, and by the adoption of the factor "Military Character" as a crucial factor for the determination of a cyber-attack as a "use of force". Secondly, Multiple Attribute Decision Making (MADM) methods are applied. The methodology is applied in real-life cyber-attacks, namely the large-scale attacks against Estonia and Iran, and cyber-attack evaluation results are presented. The correlation of both qualitative and quantitative methods of analysis allows to achieve an improved cyber-attack evaluation assessment and as a result a more accurate and complete cyber-attack classification.

4 Multi criteria decision analysis methods

Multiple Attribute Decision Making (MADM) involves "making preference decisions (such as evaluation, prioritization and selection) over the available alternatives that are characterized by multiple, usually conflicting, attributes (Hwang and Yoon, 1981). The problems of MADM are diverse and can be found in virtually any topic. Franklin, more than 200 years ago, recognised the presence of multiple attributes in everyday decisions, and suggested a workable solution (MacCrimmon, 1973).

By using the IGEs approach in the Tallinn Manual, three different MADM methods are applied for evaluating the effects of cyber-attacks. Furthermore, the three MADM methods are implemented in the real-life cyber-attack incidents of Estonia and Iran in order to evaluate these attacks and to answer the question of whether these cyber-attacks have risen to the level of "use of force" under the principles of international law.

Each decision table (or decision matrix) in MADM methods has four main parts, namely: (a) alternatives, (b) attributes, (c) weight or relative importance of each attribute, and (d) measures of performance of alternatives with respect to the attributes. The decision table is shown in Table 2 and identifies alternatives as Ai (i=1,2,…,N), attributes as Bj (j=1,2,…,M), weights of attributes as wj (j=1,2,…,M) and the measures of performance of alternatives as mij (i=1,2, …,N and j=1,2,…,M). Given the decision table information to the decision-making method, the task of the decision maker is to find the best alternative and/or to rank the entire set of alternatives. Additionally, all the elements in the decision table must be normalized to the same units, so that all possible attributes in the decision problem can be considered (Rao, 2007).

Alternatives	**Attributes**			
	B_1	B_2	-	B_M
	(W_1)	(W_2)	-	(W_M)
A_1	m_{11}	m12	-	m_{1M}
-	-	-	-	-
A_N	m_{N1}	m_{N2}	-	m_{NM}

Table 2. *The decision table in MADM methods*

4.1 The Simple Additive Weighting (SAW) Method

In this section the SAW methodology is described in detail for ranking cyber-attacks on safety-critical information systems. The SAW method is probably the best known and most widely used. This method calculates the overall score of an alternative as the weighted sum of the attribute scores or utilities.

This is also called the weighted sum method (Fishburn, 1967). It is the simplest and still the widest used MADM method. Here, each attribute is given a weight, and the sum of all weights must be 1. Each alternative is assessed with regard to every attribute. The overall or composite performance score of an alternative is given by the following equation:

$$\mathbf{P_i} = \sum_{j=1}^{M} \mathbf{W_j m_{ij}}$$

Equation 1. *Overall score with SAW method*

where P_i is the overall, or composite, score of the alternatives A_i. The alternatives with the highest value of P_i are considered the best alternatives.

Table 3 demonstrates the decision matrix for the cyber-attacks against Estonia and Iran. It is important to note that the weight value of each attribute was evaluated based on the consequence analysis approach of the IGEs. Furthermore, they were normalized in a scale of 1. Additionally, the value for each attribute in each alternative was evaluated based on the real-life cyber-attack analysis of section 2.

Alternatives	Attributes							
	Severity	Immediacy	Directness	Invasiveness	Measurability of effects	Presumptive Legality	State Involvement	Military Character
	0.18	0.16	0.08	0.08	0.08	0.08	0.18	0.16
The Estonian Cyber-Attack	7	8	6	7	7	5	5	6
The Iranian Cyber-Attack	9	7	8	10	8	9	5	8

Table 3. *The decision table for the Estonian and the Iranian cyber-attacks*

In Table 4, the cyber-attacks against Estonia and Iran, which are described in the decision matrix of Table 3, are evaluated using the SAW method. It appears that the cyber-attack in Iran is more critical than the Estonia one.

Alternatives	SAW (P_i)
The Estonian Cyber-Attack	6.40
The Iranian Cyber-attack	7.72

Table 4. *Ranking using the SAW method*

Using the quantitative scale of Figure 2 and taking into account the results of Table 4, the impact of the Estonian cyber-attack can be placed on the high end of the central "grey area" on the Schmitt scale. Therefore, the cyber-attack against Estonia is arguably a "use of force". On the contrary, the impact of the Iranian cyber-attack can be placed on the low end of the high range on the Schmitt scale. Therefore, the cyber-attack on Iran's nuclear facilities was definitely a "use of force".

Nonetheless, there are still specific weaknesses using the SAW method. Let's assume a hypothetical attack where the "State Involvement" attribute is given a value of zero and the other attributes hold the same values as presented above. The SAW method for this case will place the consequences of the attack on the high end of the central "grey area" on the Schmitt scale where it cannot be identified if a "use of force" occurred or not. However, it is generally known that when the "State Involvement" attribute value of an attack is next to zero, this attack is unlikely to be classified as a "use of force". This is because the clearer and closer a nexus between a state and a cyber operation, the more likely is to be characterized as a "use of force". Absent a "State Involvement" it is unlikely that a cyber operation will be characterized as a "use of force". Therefore, it should be classified in the low range on the Schmitt scale, not in the central area. This example shows that it cannot appropriately model such kinds of attacks when applying the SAW methodology (Pipyros et al., 2016).

4.2 The Weighting Product Method (WPM)

The Weighted Product Method was introduced by Bridgeman (1922). According to Yoon and Hwang (1995) the method possesses sound logic and is computationally simple, but has not been widely utilized. Contrary to the SAW method, the different measurement units here do not have to be transformed into a dimensionless scale by a normalization process. This is because in the WPM method the attributes are connected by multiplication. The weights become exponents associated with each attribute value. In this method, the overall or composite performance score of alternatives is given by Equation 2:

$$P_i = \prod_{j=1}^{M} \left[m_{ij} \right]^{w_j}$$

Equation 2. *Overall score with WPM method*

Each value of an alternative with respect to an attribute, i.e. m_{ij}, is raised to the power of the relative weight of the corresponding attribute. The alternative with the highest P_i value is considered the best alternative.

In Table 5, using the WPM method we evaluated the Estonian and the Iranian cyber-attacks described in the decision matrix of Table 3. We observed again that the cyber-attack against Iran is more critical than the Estonian case. WPM operates on the premise that, in the absence of a conclusive definitional threshold with widespread acceptance within the international community, states must be highly sensitive to the international community's probable assessment of whether a cyber operation violates the prohibition on the "use of force".

Alternatives	**WPM (P_i)**
The Estonian Cyber-Attack	6.31
The Iranian Cyber-Attack	7.55

Table 5. *Ranking using the WPM method*

Assuming again the hypothetical attack of the previous section, where the "State Involvement" attribute is given with a value of zero while keeping the same values for other attributes, it is easily understood that the overall performance score (which is a product) becomes zero now. This is because in the WPM method, the attributes are connected by multiplication. Thus, the WPM method for this case will place the consequences of the attack as not a "use of force" whichever quantitative scale someone decides to use. Although applying WPM in some kind of attacks gives better results than SAW, the lack of a definitional threshold for the appropriate ranking and classification of them seems to be a major drawback. Moreover, the nonlinear relationship between attributes and overall score in WPM makes more difficult the definition of a

quantitative scale for the classification of attacks than using the SAW method (where a linear relationship exists).

For the above reasons, in the following section we present a new strategy for cyber-attacks evaluation that combines the use of the first two methodologies and introduces a new grouping of the International Group of Experts criteria for achieving a more accurate and complete cyber-attack modelling assessment.

5. A new cyber-attack evaluation methodology

In this section we continue our analysis by presenting a new modelling methodology that introduces a new calculation procedure and a new usage of the IGEs qualitative criteria for the better evaluation and classification of cyber-attacks. This new strategy combines the use of the previous two decision making algorithms and introduces a new grouping of IGEs criteria based on their properties for achieving a better modelling of cyber-attacks. Figure 3 is a schematic diagram of this new strategy for cyber operations evaluation. The next paragraphs describe our methodology.

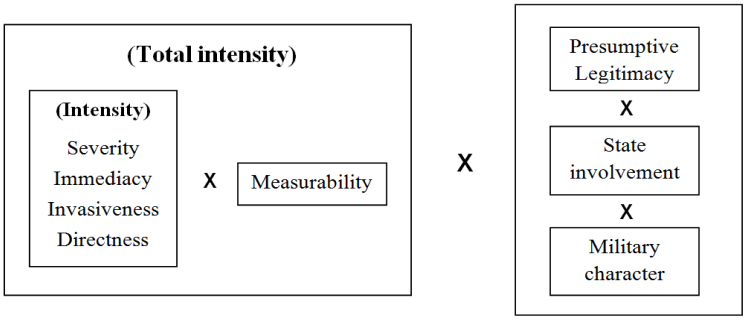

Fig. 3. *The schematic diagram of the new strategy for cyber-attack evaluation*

Firstly, as shown in Figure 3, "Severity", "Immediacy", "Invasiveness" and "Directness" are grouped together giving a new group named "Intensity". "Severity" refers to the degree of destruction of critical infrastructure or loss of human lives. It is, self-evidently, the most significant factor of the analysis. "Immediacy" focuses on the temporal aspects of the consequences in question whereas "Directness" examines the chain of causation (the indirect causal connection between the initial act and its effects). Furthermore "Invasiveness" refers to the degree to which cyber-attacks intrude into the target state or its cyber systems contrary to the interests of that state. The more secure a targeted cyber system, the greater the concern as to its penetration.

The four criteria are grouped together for two reasons: a) they are referred to the magnitude (intensity) of a cyber-attack, and b) they can be quantified by using the same quantitative scale (say, a one-to-ten scale). These attributes are the base of our calculation procedure and by applying the SAW method we can calculate the "Intensity" group score of a cyber-attack. Table 6 demonstrates the decision matrix for the above mentioned kinetic and cyber-attack on Supervisory Control and Data Acquisition (SCADA) systems so as to calculate the "Intensity" score of such attacks. For doing so, we should

use in the decision matrix the four IGE's criteria: "Severity", "Immediacy", "Directness" and "Invasiveness". The weights of these attributes need to be redefined again such that they could meet the requirement that the sum of all weights must be 1.

Alternatives	Attributes			
	Severity	**Immediacy**	**Directness**	**Invasiveness**
	0.38	0.30	0.16	0.16
The Estonian Cyber-Attack	7	8	6	7
The Iranian Cyber- Attack	9	7	8	10

Table 6. *The decision table for the cyber-attacks against Estonia and Iran for "Intensity" score calculation*

In Table 7 and by using the SAW method, we calculate the "Intensity" score of the Estonian and Iranian cyber-attacks described in the decision matrix of Table 6. It appears that the "Intensity" score of the cyber-attack in Iran is higher than the Estonian incident.

Alternatives	**Intensity (P_i)**
The Estonian Cyber-Attack	7.14
The Iranian Cyber-Attack	8.40

Table 7. *Ranking using the SAW method*

Next, as presented in Figure 3, we multiply the "Intensity" score of an attack by the "Measurability" attribute to calculate the "Total Intensity" score. The more quantifiable and identifiable a set of consequences, the easier it will be for a state to assess the situation when determining whether the cyber operation in question has reached the level of a "use of force". The "Measurability" attribute can be quantified by using the quantitative scale from 0 to 1. By using a value of 1 it means that a complete and accurate (100%) measurement of the effects of an attack can be achieved. By using zero it means that the effects of an attack are not measurable. In Table 8, we calculate the "Total Intensity" score of the cyber-attacks against Estonia and Iran described in the decision matrix of Table 6.

Alternatives	**Intensity (P_i)**	**Measurability**	**Total Intensity**
The Estonian Cyber-Attack	7.14	0.7	4.99
The Iranian Cyber-Attack	8.40	0.8	6.72

Table 8. *Calculating the "Total Intensity" score*

Last but not least, “State Involvement”, “Military Character” and “Presumptive Legitimacy” are some of the most valuable factors for the characterization of a cyber operation as a “use of force” or not. The extent of “State Involvement” in a cyber operation lies along a continuum from operations conducted by a state itself to those in which its involvement is peripheral. The clearer and closer a nexus between a state and cyber operations, the more likely is that other states will characterize them as uses of force by that state. Furthermore, a nexus between the cyber operation in question and military operations heightens the likelihood of characterization as a “use of force”. The “use of force” has traditionally been understood to imply force employed by the military or other armed forces. This contention supported by the fact that the UN Charter is particularly concerned with military actions. Finally, absent an express treaty or accepted customary international law prohibition, an act is presumptively legal. This being so, acts like propaganda, espionage, psychological operations are less likely to be considered by states as uses of force. Only if the criteria of “State Involvement”, “Military Character” and “Presumptive Legitimacy” are met, a state can characterize a cyber-attack as a ‘use of force’. For this reason, in order to quantify them we use binary logic assigning to them the values 0 or 1 (false or true).

In Figure 3, the attributes “Total Intensity”, “Military Character”, “State Involvement” and “Presumptive Legitimacy” are connected with multiplication. Thus, the last three should be “true” in order to have a non-zero overall score as a final result in the evaluation procedure. If one of them is zero, the overall score will be also zero. Therefore, for the evaluation of cyber-attacks by using this methodology it is of fundamental importance to be able to decide if these three criteria are met or not.

In Table 9, we calculate the overall score of the cyber-attacks against Estonia and Iran, described in the decision matrix of Table 6, by using our methodology. It is observed that the cyber-attack in Iran is more critical than the Estonia one.

Alternatives	**Total Intensity**	**State Involvement**	**Military Character**	**Presumptive Legitimacy**	**Overall Score**
The Estonian Cyber-Attack	4.99	1	1	1	4.99
The Iranian Cyber-Attack	6.77	1	1	1	6.77

Table 9. *Calculating the overall score*

Thus, using again the quantitative scale of Figure 2 and taking into account the results of Table 9, the consequences of the cyber-attack against Estonia are placed in the mean of the central “grey area” on the Schmitt scale which denotes that the cyber-attack against Estonia was “arguably a use of force”. On the other hand, the consequences of the cyber-attack against Iran are placed on the low range of the high range on the Schmitt scale which denotes that the cyber-attack against Iran was actually a “use of force”

In conclusion the existing legal norms do not offer a comprehensive framework in the way that states can shape policy to the threat of hostile cyber operations. Furthermore, state practice is lacking in characterizing a cyber operation as a “use of force” or

not. Even though there were many cyber operations that could be reach the level of a "use of force", in none of these cases states have been identified as the initiator of the cyber operation which might amount to a "use of force". The threshold of a "use of force" must be balanced between on the one hand state's willingness to avoid any harmful consequences caused by the actions of others states and one the other hand its motivation to preserve their freedom of action. The evaluation criteria proposed by the IGEs in the Tallinn Manual seek to balance these conflicting objectives through consideration. However, as Schmitt admitted (2011) "the criteria are admittedly imprecise, thereby permitting states significant latitude in characterizing a cyber operation as a 'use of force' or not". Furthermore, a state, depending on the attendant circumstances, may look also to other factors such as the prevailing political environment, whether the operation portends the future "use of force", the identity of the attacker and the nature of the target. In fact, a finding that a cyber operation is a "use of force" is not a legal but a political decision, as it shows the states willingness to involve itself in a particular matter.

For the above mentioned reasons the authors of this paper have chosen to present this work in a manner to provide clear structure for discussion. It was not their intent to provide an absolute algorithm for producing the "right answer" given any input. The proposed systematic methodology is applied in order to portray a better modelling evaluation of cyber-attacks. It contributes in areas where there is uncertainty or disagreement in a number of a legal analysis and for making available means for addressing all issues related to the "use of force" concept.

6. Conclusions

In this paper, the aim was to present a new systematic modelling methodology for evaluating the effects of cyber-attacks on states' CII in order to define whether these attacks constitute a wrongful "use of force" under the *jus ad bellum*, that body of international law that governs a state's resort to force as an instrument of its national policy. We have adopted the "effects-based" or "consequences-based" approach, which focuses on the overall effect of a cyber operation to the victim-state, as well as by using the qualitative criteria for recognizing the impact of cyber-attacks as proposed by the IGEs in the Tallinn Manual. Furthermore, MADM methods and more specifically the SAW method and the WPM were applied.

Evidently, the characterization and classification of cyber-attacks on state's CII depends largely on the extent of their consequences. In other words, the categorization of the type of attack lies heavily on its impact level both in terms of the loss of human lives and in terms of the destruction of critical infrastructures. Consequently, the degree of the immediate as well as of the long-term effects of a cyber-attack constitutes a critical factor for its categorization. Furthermore, the greater the degree of impact of a cyber-attack, the greater the chances are that it will be characterized as "use of force", or even worse, as "armed attack" when its magnitude is so great as to cause loss of human lives. Thus, the main issue of investigation is to define the method of measurability of the impact of a cyber-attack.

The main contribution of the paper relies on the development of a new cyber-attack evaluation methodology that combines the IGEs qualitative criteria with MADM methods. The pros and cons of each MADM method lead us to present a new cyber-attack

evaluation strategy that combines the use of decision making algorithms of MADM methods and introduces a new grouping of the IGEs criteria based on their distinctive features. The methodology has been applied in real-life cyber-attacks, namely the large-scale cyber-attacks against Estonia and Iran, and cyber-attack evaluation results are been presented. The correlation of both qualitative and quantitative methods of analysis allows to achieve an improved cyber-attack evaluation assessment and as a result a more accurate and complete cyber-attack classification. The proposed systematic methodology is applied in order to portray a better modelling evaluation of cyber-attacks. However, it was not the purpose of the authors to provide an absolute algorithm for producing the "right answer" given any input.

The threshold of a "use of force" must be balanced between on the one hand state's willingness to avoid any harmful consequences caused by the actions of others states and on the other hand its motivation to preserve their freedom of action. The evaluation criteria proposed by the IGEs in the Tallinn Manual seek to balance these conflicting objectives through consideration. As such, the usefulness of the methodology is perceived in areas where there is uncertainty or disagreement in a number of legal analyses, and for making available a means for addressing all issues having to do with "use of force". Finally, the methodology could act as a basis for the assessment and classification of cyber-attacks that are intended towards Software-Intensive (SI) systems, component of a state's CII.

References

Blank Stephen: "Web war I: is Europe's first information war a new kind of war?", Taylon & Francis Online Journal, 2008, Vol. 3 No. 2, pp. 227-247.

Bridgeman Percy: "Dimensional Analysis", New Haven, Yale University Press, 1922.

Collins Sean and McCombie Stephen: "Stuxnet: The emergence of a new cyber weapon and its implications", Journal of Policing, Intelligence and Counter Terrorism, 2012, Vol. 7, No. 1, pp. 80-91.

Council of Europe: Convention on Cybercrime. European Treaty Series 185, 2001.

Eisenberg Ted, Gries David, Hartmanis Juris, Holcomb Don, Lynn Stuart, Santoro Thomas: "The Cornell Commission: On Morris and the Worm", Communications of the ACM 32, no 6, 1989.

Elkus Adam: "Moonlight Maze in A Fierce Domain: Conflict in Cyberspace", Cyber Conflict Studies Association, 2013.

European Parliament and the Council: Directive 1148 concerning "measures for a high common level of security of network and information systems across the Union". Official Journal of the European Union, 2016.

Farwell James and Rohozinski Rafal: "Stuxnet and the Future of Cyber War" Survival, Global Politics and Strategy, 2011 Vol. 53, Issue 1, pp. 23-40.

Fishburn Peter: "Additive utilities with incomplete product set: Applications to priorities and assignments". Operations Research, 1967, Vol. 15, No. 3, pp. 537-542.

Haizler Omry: "The United States' Cyber Warfare History: Implications on Modern Cyber Operational Structures and Policymaking", Cyber, Intelligence and Security, vol. 1, no. 1, 2017.

Hwang Ching-Lai, Yoon Kwangsun: "Multiple Attribute Decision Making: Methods and Applications", Berlin/Heidelberg/New-York: Springer Verlag, 1981.

International Court of Justice Advisory Opinion on the Legality of the Threat or Use of Nuclear Weapons Advisory Opinion, 1996, par. 39.

Jason Jolley: “Article 2(4) and Cyber warfare: How do Old Rules Control the Brave New World?”, Canadian Center of Science and Education, Vol. 2, No 1, 2013.
Jougleux Philippe, Mitrou Lilian, Synodinou Tatiana-Eleni: “The Legal Regulation of Cyber Attacks”, Kluwer Law International, 2016.
Karnouskos Stamatis: “Stuxnet worm impact on industrial cyber-physical system security”, in Proc. of the 37th Annual Conference on IEEE Industrial Electronics Society, Australia, 2011.
Langer Ralph: “Stuxnet: Dissecting a Cyber warfare Weapon”, IEEE Security & Privacy, 2011 Vol. 9, Iss. 3.
Langner Ralph: “Stuxnet’s Secret Twin: The real program to sabotage Iran’s nuclear facilities was far more sophisticated than anyone realized”, Foreign Policy, 2013.
MacCrimmon Kenneth: “Multiple Criteria Decision Making: An Overview of Multiple Objective Decision Making”. University of South Carolina Press, 1973, pp. 18-44.
Morth Todd: “Considering our position: Viewing information warfare as a use of force prohibited by article 2(4) of the UN Charter”. Case Western Reserve Journal of International Law, 1998, Vol. 30, Is. 2, pp. 567-600.
Panetta Leon: “Defending the Nation from Cyber Attack”, Business executives for National Security, October 2012. Available at: URL: http://www.bens.org/document.doc?id=188
Permanent Court of International Justice: “The Case of the S.S. Lotus France vs Turkey, Judgment No. 9, 1927.
Pipyros Kosmas, Mitrou Lilian, Gritzalis Dimitris and Apostolopoulos Theodoros: “A review of obstacles in applying international law rules in cyber warfare”, Information & Computer Security, 2016, vol. 24, Iss. 1, pp. 38-52.
Pipyros Kosmas, Thraskias Christos, Mitrou Lilian, Gritzalis Dimitris and Apostolopoulos Theodoros: “Cyber-Attacks Evaluation Using Simple Additive Weighting Method on the Basis of Schmitt’s Analysis”. In Proc. of the 10th Mediterranean Conference on Information Systems, MCIS Proceedings, 41.
Rao Venkata: “Decision Making in the Manufacturing Environment Using Graph Theory and Fuzzy Multiple Decision Making (MADM) Methods”, Springer – Verlag London, 2013.
Robinson Michael, Jones Kevin and Janicke Helge: “Cyber warfare: Issues and Challenges”, Computers & Security, 2015, Vol. 49, pp. 70-94.
Roscini Marco: “Cyber Operations and the Use of Force in International Law”, Oxford University Press, 2014.
Sanger David: “Obama Order Sped up Wave of Cyber-attacks against Iran”, The New York Times, 2012, Available at: http://www.nytimes.com/2012/06/01/world/middleeast/obama-ordered-wave-of-cyberattacks-against-iran.html
Schmitt Michael: “Computer Network Attack and the Use of Force in International Law: Thoughts on a normative framework”. Columbia Journal of Transnational Law, 1999, Vol. 37, pp. 885-937.
Schmitt Michael: “Cyber operations and the Jus ad Bellum Revisited”, Villanova Law Review 56, 2011, pp. 569-606.
Schmitt Michael (Ed.): “Tallinn Manual on the International Law Applicable to Cyber Warfare”, Cambridge University Press, 2013.
Schmitt Michael (Ed.): “Tallinn Manual 2.0 on the International Law Applicable to Cyber Warfare, Cambridge University Press, 2017.
The Economist: “Cyber War in the fifth domain: Are the mouse and the keyboard the new weapons of conflict?” 2010. Available at: www.economist.com/node/16478792
Tikk Eneken, Kaska Kadri and Vihul Liis: “International Cyber Incidents: Legal Considerations”, Cooperative Cyber Defence Centre of Excellence (CCD COE), 2010.
Tsagourias Nicholas: “Cyber-attacks, self-defence and the problem of attribution”, Journal of Conflict & Security Law, 2012, Vol. 17, Is. 2, pp. 229-244

Virvilis Nikos, Gritzalis Dimitris: “The Big Four - What we did wrong in Advanced Persistent Threat detection?”, in Proc. of the 8th International Conference on Availability, Reliability and Security (ARES-2013), pp. 248-254, IEEE, Germany, September 2013

Yoon K. Paul and Hwang Ching-Lai: “Multiple Attribute Decision Making: An Introduction”, Sage University Paper Series on Quantitative Applications in the Social Sciences, 1995, pp. 7-14.