TALLINN UNIVERSITY OF TECHNOLOGY
School of Information Technologies

Kristine Hovhannisyan 177337IVCM

Applying Confidence-Building Measures to Cyber Conflict: Computer Emergency Response Cooperation and Cyber Espionage

Master's thesis

Supervisor: Eneken Tikk, PhD
Olaf Maennel, PhD

Tallinn 2019

TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia teaduskond

Kristine Hovhannisyan 177337IVCM

Usaldusmeetmete rakendamine küberkonfliktis: arvutivõrkude intsidentide alane koostöö ja küberluure

Magistritöö

Juhendaja: Eneken Tikk, PhD
Olaf Maennel, PhD

Tallinn 2019

Author’s declaration of originality

I hereby certify that I am the sole author of this thesis. All the used materials, references to the literature and the work of others have been referred to. This thesis has not been presented for examination anywhere else.

Author: Kristine Hovhannisyan

13.05.2019

Abstract

Confidence-building measures (CBMs) originate from the period of Cold War and were introduced to help restrain unintended escalation based on misunderstanding that could derive from regular military activities. These measures are now being introduced to ICT field by the UN Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context on International Security.

This thesis is interested in the effect that the implementation of the proposed and recommended CBMs would have on prevention and mitigation of cyber conflict. To facilitate understanding of the anticipated impact of the proposed CBMs, the author develops a methodology for applying those confidence-building measures on cyber conflict. Firstly, the author studies the history and theory of the CBMs and identifies seven characteristics (Regular, Continuous, Precise, Specific, Present, Systemic, Measurable) of the CBMs that are used as a blueprint for the analysis of the proposed CBMs for the ICT field. The author identifies that the proposed CBMs does not correspond to the set requirements of the CBM characteristics and concludes that these proposed CBMs are currently propositions for CBMs, they do not constitute as complete measures. The author makes recommendations for improving the proposed CBMs. Secondly, by analysing six publicly known datasets of cyber incidents, the author concludes that cyber espionage constitutes over 82% of cyber incidents and creates an ontology of cyber espionage. Thirdly, the author uses the model, the ontology of the cyber espionage and one of the proposed CBMs to conduct a cross-examination and concludes that the elements of the proposed CBM don't provide the anticipated effect of reducing the risk of misunderstanding and helping in preventing unintended escalation. The author develops a methodology for evaluating the effectiveness of the CBM of CERT cooperation that can be used for further evaluation of other proposed CBMs.

Keywords: Confidence-building measures, cyber conflict, international peace and security, cyber espionage.

This thesis is written in English and is 66 pages long, including 4 chapters, 7 figures.

Acknowledgements

Thank you, Dr. Eneken Tikk, for guiding me throughout this journey. You challenged me, and I tried to keep up, you made me stronger. You allowed me think independently and make my own choices. I very much appreciate the time and effort you invested in helping me to arrive to the finish line.

Thank you, Dr. Olaf Maennel, for your continuous support and the positive approach and energy you spread around your students. Without your support and understanding this thesis would not have been completed.

Thank you, Dr. Mika Kerttunen, for helping to translate sometimes simple, sometimes difficult thoughts. Without your methodological guidance, I would not have been able to construct my own thoughts.

I would like to thank Klaid Magi, knowing his busy schedule, for dedicating time for our discussions about CERTs/CSIRTs. Thank you for taking your time to provide an overview of the differences that national, governmental, military and other CERTs have.

I, in particular would like to thank, Brady Maxwell. You supported me day and night, cheered me and encouraged me to not give up. Thank you, all friends and colleagues, who shared this long and tough journey together.

List of abbreviations and terms

UN	United Nations
UN GGE	United Nations Group of Governmental Experts
CBM	Confidence-building measures
MAD	Mutual Assured Destruction
MoU	Memorandum of Understanding
OSCE	Organization for Security and Co-operation in Europe
ICT	Information and Communications Technologies
CERT	Computer Emergency Response Team
CSIRT	Computer Security Incident Response Team
CIRT	Cybersecurity Incident Response Team
CFR	Council on Foreign Relations
CSIS	The Center for Strategic and International Studies
GReAT	The Global Research and Analysis Team
COT	Cyber Operations Tracker
ENISA	European Network and Information Security Agency
CIIP	Critical Information Infrastructure Protection
SOP	Standard Operating Procedure

Table of contents

List of figures

Introduction

Confidence-building measures (CBM) originate from the period of Cold War. These measures were introduced to help to restrain from unintended escalation based on misunderstanding that could derive from regular military activities. They aimed to eliminate the fear that could trigger a conflict between nuclear states, the United States and the Soviet Union. These measures have evolved over time internationally [1] [2] and regionally [3] [4] [5].

In 1950s the Soviet Union initiated a discussion on collective security in Europe by demanding new European Security Treaty. The efforts for establishing international peace and security resulted in Helsinki Conference on Security and Cooperation in Europe, and was concluded respectively by Helsinki Final Act in 1975 [6] that adopted certain CBMs. As a follow-up, another meeting took place in Madrid, then in Stockholm and was concluded with reshaped CBMs in the Document of Stockholm [4]. The negotiations on CBMs were then finalised in the Vienna Document in 1990 [5].

In parallel with the European initiatives, the CBMs have been developing on the United Nations level, and the comprehensive study conducted to define what these measures are and how these can be implemented was performed by the UN GGE on Confidence-building Measures back in 80s [7]. To showcase how CBMs have been evolving, a timeline with major events has been created and is presented in the Annex 1. To highlight, most of the CBMs suggested by states were "one or more of the non-military categories, but a high proportion of the proposals were related to military concerns, reflecting the high priority that many Governments accord to problems of security" (page 30 [7]). Almost two decades later, in 2010 these measures were introduced in information and communications technologies (ICT) environment [8], and expanded by the time [9][10].

This thesis is interested in the effect that the implementation of the proposed CBMs for the ICT environment would have on prevention and mitigation of cyber conflict. To facilitate understanding of the anticipated impact of the proposed CBMs, the author develops a methodology for applying those confidence-building measures on cyber conflict, using the examples of the CBM on computer emergency response teams and cyber espionage as the most prevalent type of cyber conflict.

To determine which cyber incidents are prevalent and which type can be regarded as examples of conflict, or precursors to it, and to where, accordingly, conflict prevention measures must be targeted, this thesis identifies a typology of cyber incidents, based on publicly known state sponsored cyber incidents and creates an ontology of the prevalent cyber incident. It then examines some of the measures that have been proposed by a UN Governmental Group of Experts to address situations where an activity in cyberspace could affect international peace, security or stability.

This thesis results in a methodical approach to evaluating the effectiveness of proposed confidence-building measures in mitigating various types of cyber conflict. It also provides recommendations on further developing the CBMs, by applying the theoretical considerations of CBMs, analyzed in Chapter 1, to the proposed measure in UN GGE 2015 report for CERT cooperation as additional confidence-building measures (see page 10 (d) [10]).

Motivation

There is a variety of documents, literature, papers available on CBMs and their conceptual analysis, their applicability for conflict prevention globally and regionally [11], but there is limited research conducted on practical implementation of the CBMs in the field of ICT and cyber conflict prevention. A partial relevance that "Confidence Building Measures for Cyberspace – Legal Implications" [12] paper has, is relatively limited analysis of nature and obstacles of CBM implementation on concrete cyber incidents. Despite that this paper concentrates on states legal and political obligations deriving from CBMs, Dr. Katharina Ziolkowski emphasizes that there is a need for research on this topic "It should be mentioned that, surprisingly, the current developments of CBMs for cyberspace, as far as respective documents are publicly available, do not contain any reference to a CBM of exchange of scientific research or of academic personnel (…), a measure, which could be considered as politically rather innocuous" (see page 24 [12]). Dr. Ziolkowski concludes that "an analysis of lessons identified with regard to CBMs, as collected by armed forces and peace research institutes during the last decades, would be beneficial in order to consider the findings during the current negotiations with regard to cyberspace" [12].

One relevant research that was identified is the "Confidence Building Measures for the Cyber Domain" by Erica D. Borghard and Shawn W. Lonergan. The authors emphasize that "while governments have taken initial efforts to establish cyber CBMs, current academic work on the topic is at a nascent stage" [13]. They use "(…) Cold War frameworks for evaluating CBMs as a benchmark for developing realistic CBMs for the cyber domain in light of the latter's distinct characteristics" [13]. Their approach in analyzing CBMs limits for cyber domain is based on Johan Holst's framework of the CBM analysis that was conducted in 1983. According to his framework there are four main categories of CBMs: (1) information, (2) notification, (3) observation, and (3) stabilization [13]. The authors have discussed the complexities of these categories, created a new framework for CBMs, and organized all the existing CBMs, the UN GGE and OSCE proposed, into one new framework [13]. This analysis is based rather on the categories of the CBMs, than the characteristics.

There are other relevant document(s) that address CBMs for cyberspace [14] [15] or suggest these measures for rival countries, for instance India and Pakistan [16]. But these materials unfortunately do not address or question the nature of the CBMs to understand if the applicability of those is realistic or not and what impact would potentially be obtained.

The research question that the author poses is: **How to apply the CBMs to cyber conflict, using the example of CERT cooperation and cyber espionage?**

Structure of the thesis

To be able to answer the research question, the author will conduct a study of the following sub-questions that also define the structure of the thesis:

- In the first chapter the author discusses what is a CBM and what are the characteristics of the CBMs. The author provides examples from the existing CBMs, by using the Vienna Document [5] and the Agreement Between the Government of The United States of America and the Government of The Union of Soviet Socialist Republics on the Prevention of Incidents On and Over the High Seas [17] to demonstrate and validate the characteristics of the CBMs.
- In the second chapter the author examines cyber conflict on the basis of datasets of cyber incidents and creates a typology of cyber conflict on the basis of these

sources. Concluding that cyber espionage constitutes over 82% of incidents qualified as cyber conflict, the author develops a conceptual model of cyber espionage as the object of application of the selected CBM.

- In the third chapter the author applies the theoretical framework of CBMs, discussed in Chapter 1, to the measure proposed in UN GGE 2015 report (page 10 (d) [10]). The purpose of this chapter is to determine if, and how, the proposed measure could be further developed. The author uses the following methods of examination: textual interpretation and teleological argumentation, and negation to examine the proposed measure and make recommendations on improving it.
- In the fourth chapter, the author develops an approach to examine the applicability of CBMs to cyber conflict, using the CERT cooperation as an example of CBM and cyber espionage as example of cyber conflict. This approach examines each element of the CBM to each factor of the type of conflict in question, keeping in mind the broader context of CBMs as verification mechanisms.

Methodology

The methodological approach used for this thesis is:

(1) Literature review: studying international documents, treaties, agreements and resolutions, non-profit international and regional organisations reports, papers and workshops outcomes, academic literature and articles, databases and other online resources. Several instances of online literature sources have been used. Predominantly, the United Nations resolutions and structural organisations reports are used to define CBMs.

(2) Research and identification of relevant publicly available dataset(s) of cyber incidents.

(3) Analysis of data: (a) a validation of data have been accomplished, by the method of comparison with other datasets, and (b) over fifteen years of data (from 2005 till 2018 October) was analysed to define the typology of cyber incidents, and (c) a statistical analysis was conducted to identify the prevalent type of cyber incident. Based on quantitative data analysis, charts are prepared and presented in the thesis.

(4) Interpretation: the author applied several methods of interpretation for the proposed CBM to define if the proposed CBMs meet the formal criteria of the

CBMs identified and discussed in Chapter 1. The author used (a) textual interpretation and negation to provide additional layer of justification and (b) teleological argumentation, that is based on identifying the goal(s) and the intent.

(5) Interviews: two interviews were conducted. One interview aimed at discussing the model of espionage, the ontology of cyber espionage. Another interview was to discuss the role of the national CERTs/CSIRTs and limitations in terms of operations and constituencies.

(6) Modelling of cyber espionage, as an object for applying the selected CBM for evaluating the effectiveness.

Limitations

The following limitations are recognized for the current thesis:

1. This thesis does not study or define *cyber conflict*.
2. The identified datasets of cyber incidents are limited in data points to the extent of where the incidents have been disclosed by states. It is presumed that there are cyber operations or incidents that have not been made public, taking into consideration that disclosing those might reveal or in any other way affect the national security aspects of the nation states.
3. The typology of cyber incidents and the methodology for including those into the dataset is limited to the specific selected dataset used for the analysis.
4. The identified datasets are in English, thus there is a likelihood that the data reflected in the datasets are western-centric.
5. The application of the CBMs on cyber incidents is limited to typology of cyber incidents defined in one dataset.
6. The author has not concentrated on specific UN GGE report, rather on one concrete CBM.

Contribution

There are various studies on confidence-building measures that address their nature, effectiveness for serving as verification mechanisms and contributing to the objective for arms control, but there are a few addresses these measures for the ICT field.

The author's contribution with this thesis constitutes the development of a methodology on how to apply the proposed CBMs by UN GGE for the ICT field. The methodology allows to analyse and define the characteristics of the confidence-building measures, to be able to examine their applicability. Knowing what CBMs are, it will help to understand what effect they can have on cyber incidents. The author bases the analysis not on the categories of the CBM, but rather understanding the characteristics of the CBMs. The author creates a blueprint of CBM characteristics to evaluate the nature of the proposed CBMs.

Additionally, to be able to picture the landscape, the typology of cyber incidents that are prevalent, the author has studied six datasets of cyber incidents. The author uses the most prevalent type of cyber incident to create a conceptual model that can be used as an object to apply the proposed CBMs for the ICT field for further analysis.

1 Confidence-Building Measures

1.1 Historical overview of Confidence-Building Measures

The confidence-building measures (CBMs) that have been used for minimizing misunderstanding and preventing unintentional escalation, tensions between states, and additionally served as measures for the arms control and disarmament. One successful example of CBM implementation back in 70s was the Memorandum of Understanding (MoU) signed between the United States of America and the Union-of Soviet Socialist Republics for establishing a direct communication link [7].
These measures are now being introduced to ICT field. And because "the vulnerabilities in ICTs as well as *de facto* exploitation of these vulnerabilities by state (…) has been acknowledged and problematized" [18] and threatening to international peace and security, the UN GGE [8] [9] [10] have been proposing to use CBMs for the ICT environment as well. This testifies that states view the information and communication technologies as a potential weapon and threat to national security as well.

The UN GGE explains that CBMs "can increase interstate cooperation, transparency, predictability and stability" [10] . As to their task, however, the UN GGE makes reference to the Guidelines for Confidence-building Measures adopted by the Disarmament Commission in 1988: "A centrally important task of confidence-building measures is to reduce the dangers of misunderstanding or miscalculation of military activities, to help to prevent military confrontation as well as covert preparations for the commencement of a war, to reduce the risk of surprise attacks and the outbreak of war by accident; and thereby, finally, to give effect and concrete expression to the solemn pledge of all nations to refrain from the threat or use of force in all its forms and to enhance security and stability." [19].

Historically, a recommendation on consideration for confidence-building measures came out in the United Nations General Assembly resolution 33/91 B of 16 December in 1978, where the "General and Complete Disarmament" [1] was the discussion subject, aiming to facilitate the arms control, disarmament and with the help of policies strengthen peace

and security. All the United Nations member states were encouraged to consider arrangements for confidence-building measures, basing on the specificities of regional dynamics, and to later on share their "experiences regarding those (…) measures they consider appropriate and feasible" [1].

In 1982 the UN GGE on Confidence-building Measures [7] published a comprehensive study on confidence-building measures. For these experts the major reference to initiate the study was the General Assembly's resolution 34/87 B dated of 11 December 1979, that emphasized the "(…) the need and urgency of first steps to diminish the danger of armed conflicts resulting from misunderstandings" and "recognizing; that a minimum of trust among states in a region would facilitate the development of confidence-building measures"[2].

These initiatives related to confidence-building within the context of international relations emerged as mechanisms to help to release tensions between superpowers during the Cold War. The nuclear arms race between United States and Soviets was on-going, and the fear of being attacked, could have triggered an escalation and result in an armed conflict. The western countries had an understanding that the Soviets were aiming in winning the nuclear warfare, although according to military personal of the Soviet Union, there were "no winners in nuclear war" (page 156 [20]). In the fewest possible words, the Mutual Assured Destruction (MAD) served as a deterrence for launching the nuclear weapons the first, as the other side would retaliate with an equal force. In distinction to MAD doctrine, the CBMs aimed at creating the collaboration between states.

The study on CBMs presented by the UN GGE on Confidence-building Measures discussed various measures and formats of collaboration. It suggested that these measures have a potential of evolving and broadening the scope of the application, from neighbouring states to regional scope, and even broader to a continental confidence-building measures. Although, when designing the CBMs, there are factors to be taken into consideration such as "past experiences, present perceptions and future expectations between states" [7]. The experts do mention that building confidence is a continuous process, and CBMs should be regarded as long-term on-going undertaking.

There are three main sections in the study: (a) Objectives, (b) Characteristics and (c) Opportunities. The study begins with defining the objectives of the measures, rather

discussing the general notion of what confidence-building measures are. Briefly, in a broader context, the CBMs aim to "strengthen international peace and security and to contribute to the development of confidence, (…)" [7]. These measures aimed on strengthening cooperation between states by eliminating fear and possible speculation that can result into escalation of conflict [7]. For instance, regular military activities can be misleading for other states, and provoke an unintended response, due to "lack of reliable information" [7]. Therefore, the CBMs with special emphasis on military and security related aspects could mitigate the possible escalation. Between other examples, "*Prior Notification on Military Activities"* (page 31(c) [7]) is one of those.

The implementation of this measure can be conducted in one of the following ways: (i) Prior notification of major military manoeuvres under agreed criteria, (ii) Prior notification of other military manoeuvres on a voluntary basis, (iii) Prior notification of major military movements [7]. Any CBMs related to military activities can be implemented in conjunction with additional mechanisms/measures. For instance, verification measures including "hot lines", given the complexity of the situations, direct communication channels could have a "stabilizing effect" [7].

Despite the fact that the experts acknowledged the need for information sharing on military activities, the degree of openness is left as a subject for further discussion. How much details would state(s) reveal about the military activities yet to be agreed [7]. But it was agreed by all experts that personal contacts being established on both political and military decision-making would facilitate and ease co-operation between states [7]. It should be kept in view, though, if and when specific CBMs are not complied with, this perhaps could be interpreted as a "warning sign for launching a surprise attack" [7].

Another major objective is that these measures should facilitate the arms control and efforts for disarmament [7]. The opportunities on implementing CBMs are various, and those can be accompanied with a series of other actions. For instance, as a primary requirement to maintain communication and contacts with other state, could be complemented with (a) preventing and containing international conflicts, (b) introducing peace-keeping forces or (c) cessation of hostilities [7].

The CBMs have specific characteristics. Firstly, the basis for establishing confidence entail an on-going, continuous process of concrete actions [7]. Secondly, assented

concrete related actions should be, simply, executed, and not omitted. The term "related" is crucial within the current context, as unrelated and broadly elaborated CBMs will not ensure the impact. They will distrust the concept, will degrade the effect of the CBMs and will convert into a "dog and pony show". Thus, leaving space for speculation of the effectiveness for disarmament, and in general undermining the importance of the concept for confidence-building measures in overall (page 9 [7]). These measures should not be implemented as one-time project, but repeated, steady footsteps. The CBMs can be introduced on various levels such as regional, interregional, international and global (page 25 [7]). For instance, on a regional basis, the Organization for Security and Co-operation in Europe (OSCE) was assigned to support the arms control mechanisms, and establish confidence-building measures [21].

To summarize, the confidence-building measures are voluntary set of actions to minimize the risk of misunderstanding and prevent escalation into conflict. For these measures to become legally binding obligations, they need to transform from voluntary commitments to political statement(s) or announcement(s), more broadly into an expressed political will, and later turn into agreements or treaties [7].

1.2 Analysis of Characteristics of Confidence-Building Measures

To be able to apply these measures, one should study and reflect upon their characteristics. It is necessary, for the purpose of this analysis, to understand how these characteristics shape, foresee and define the practical implementation/application of these measures. The comprehensive study conducted on CBMs by UN GGE on Confidence-building Measures describes these measures as "dynamic process (…) taken step-by-step within the framework of appropriate policies and international commitments" (page 8 (39) [7]). These characteristics deserve individual, one by one focus and analysis.
The author have identified seven characteristics and in the interest of discussing each characteristic, relevant examples will be presented from the Vienna Document [5] and from the Agreement on the Prevention of Incidents On and Over the High Seas signed between the United States of America and the Union of Soviet Socialist Republics (The Agreement) [17] to explain how in practice the implementation happens.

1.2.1 Regular

These measures should be implemented on a regular basis "The seriousness, credibility and reliability of a State's commitment to confidence-building, without which the confidence-building process cannot be successful, can only be demonstrated by the continuous, regular and full implementation of confidence-building measures and policies" (page 9 (41) [7]). One example of regularity is the "*Annual Exchange of Military Information on Military Forces"* (page 3 [5]). This exchange of information is conducted with definite pattern, which in the current context represents an annual implementation. Another example can be given from the Agreement on the *"Annual meetings to review the implementation of the Agreement"* [17].

1.2.2 Continuous

Continuity is reflected in activities that are accomplished without interruption. For instance, under the Section IX. Communications of Vienna Document "Each participating State will designate a point of contact capable of transmitting and receiving such messages from other participating States on a 24-hour-a-day basis" (page 42 (144) [5]). In the Agreement, it is mentioned that "the International Code of Signals, or other mutually agreed signals, shall be adhered to for signaling operations and intentions" [17]. The usage of signals cannot stop or arbitrary change, it is there to serve for its purpose continuously. It may, on mutual agreement, change the form of the signal, but the existence and usage of the signals are continuous.

1.2.3 Precise

These measures are concrete enough to be implemented. When referring to the example given before, the "*Annual Exchange of Military Information on Military Forces"* should be accomplished "not later than 15 December of each year" (page 3 [5]). As highlighted in the UN GGE study " (…) the effectiveness of a concrete measure in creating confidence will increase the more it is adjusted to the specific perceptions of threat or the confidence requirements of a given situation" (page 9 (40) [7]). The Agreement mentions that the parties should provide information and warning to mariners if actions planned (both on the sea and in air) "represent a danger to navigation or to aircraft in flight (…), not less than 3 to 5 days in advance as a rule" [17]. Another example from the Agreement would be that in case of collision, when damages have occurred, the parties should exchange information about the incident "through the Soviet Naval Attache in Washington and the

Soviet Navy shall provide such information through the United States Naval Attache in Moscow" [17].

1.2.4 Specific

The exchange of information on military forces is as specific as providing information on command organisation, such as, for each formation and combat unit of land forces and air forces [9]. Narrowing down the information, "for each formation and combat unit of land forces down to and including brigade/regiment or equivalent level the information will indicate the major organic weapon and equipment systems, specifying the numbers of each type of: (a) battle tanks; (b) helicopters; (c) armoured combat vehicles; (d) anti-tank guided missile launchers permanently/integrally mounted on armoured vehicles; (e) self-propelled and towed artillery pieces, mortars and multiple rocket launchers (100mm calibre and above); (f) armoured vehicle launched bridges" (pages 3-4 (11.2) [5]). In the Agreement, example of specific action would be when "At night, or in conditions of reduced visibility, or under conditions of lighting (…) when signal flags are not distinct, flashing light should be used to inform ships of maneuvers which may hinder the movements of others or involve a risk of collision" [17].

1.2.5 Present

These measures, both in the Vienna Document and in the Agreement, are "neither declarations of intent or a repetition of generally recognized principles nor mere promises for a certain behaviour in the future" (page 9 (39) [7]) these are existing and now occurring, real life conducted and ongoing cases.

1.2.6 Systemic

The measures are structured and interconnected. Each previous measure forms basis for next, upcoming measures, as mentioned "(…) by nature a process in which each previous measure forms the basis for further measures which progressively and cumulatively consolidate and strengthen the building of confidence, (…)" (page 8 (39) [7]). These measures are interlinked both vertically and horizontally and can be implemented on both levels independently in parallel. For instance, the exchange of "*Annual Calendars*", *"Information on Military Budgets"* of military activities [5], can be implemented in parallel with exchanging organization of three services, referred to land, air and maritime. This applies to the measures in the Agreement as well. The "*Annual Review of the*

implementation of the agreement" can be implemented in parallel with the *"warning notification to mariners"* [17]. Thus, the operational and organisational measures are taking place in parallel.

1.2.7 Measurable

The measurability should be applicable both to actions and results. These measures should be examined and accessed in advance, and "(…) States must, at each stage, be able to measure and to assess the results achieved" (page 8 (39) [7]).

In terms of measuring the actions, under the section IV of the Vienna Document prior notification of certain military activities is based on predefined and measured parameters. For instance, "(…) military activity will be subject to notification whenever it involves at any time during the activity: (a) at least 13,000 troops, including support troops, or (b) at least 300 battle tanks if organized into a divisional structure or at least two brigades/regiments, not necessarily subordinate to the same division" (page 15 [5]). In this example, there are measurable thresholds defined.

With regards to measuring the implementation of measures from the Agreement, it can be identified and stated if the warning notification have been provided within the given time-frame or time range or not.

When applying the measurability to results, we then evaluate if the agreement to deliver certain measures, both from the Vienna Document and the Agreement, have been performed or not. In this context is to assure that the agreed confidence-building measures are also practiced and delivered.

1.3 Introduction to Confidence-Building Measures in the field of Information and Telecommunications in the Context of International Security

The CBMs remained as major element of UN GGE recommendations when addressing international security in the field of information and telecommunications. The three UN GGE recommendations reports dated on 2010[8], 2013 [9] and 2015 [10] address common points and matters, but the later elaborates in much details on confidence-building measures, guiding and assisting states on creating collaboration, hereinafter "to

increase transparency, confidence and trust" [9] and "increase interstate cooperation, transparency, predictability and stability" [10] between states.

Yet, what is the significance of introducing CBMs in ICT field? They are aimed to help "to reduce the risk of conflict" between states [10], which emphasizes that the ICT are being viewed as a weapon among states. An explicit citation from the 2015 UN GGE report pointing weaponization of the ICT is as follows: "A number of States are developing ICT capabilities for military purposes" [10].

On one hand, the Group of Governmental Experts agrees that international law, in particular UN Charter applies to ICT field, and that "states must meet their obligations regarding the internationally wrongful acts" [9] or activities that relate to use of ICT. On the other hand, the experts acknowledged that establishing attribution for the misuse of the ICT is a complex process, thus eases the profiteering from the use of ICT for harmful purposes. One reason why states should collaborate, is to restrain from conflict which can be escalated (page 7 [9]) due to misinterpretation of incident(s) and/or ICT related activity(ies) and "the danger of destabilizing misperceptions" (page 6[10]).

As an observation, the 2013 UN GGE report highlights the importance of building capacity for the ICT sector in developing countries, taking into account that these countries can be used as a potential proxy, and unknowingly allow malicious activities to pass through their information/cyber space (page 2 [9]). Since ICT is nowadays a dominant topic in the international affairs, therefore, as a starting point to address any matters related to ICT, it is advised to "elaborate common terms and definitions" (page 6 [9]). The process of developing commonly and jointly accepted glossary, will open doors for mutual understanding and facilitate the discussions on mitigation processes into a higher level. Conceivably, this is the way to lay foundations.

Essentially, the 2010 UN GGE report laid foundations for the CBMs in the ICT field to be elaborated further. It, firstly, with the objective of reducing the risk of misperception resulting from ICT disruption, suggested that states should hold a dialogue on how the ICT should be used, and how important the protection of critical national and international infrastructure is. Later on, a cooperation for the critical infrastructure, and particularly the ICT-enabled industrial systems protection was highlighted in the 2013 report (page 9 [9]), while the 2015 report (page 9 [10]) took the discussion even further, and as a measure to

reduce the risk of conflict, stressed the international cooperation for sharing information related to critical infrastructure vulnerabilities [10].

Secondly, the 2010 report suggested to exchange information on ICT national strategies, policies and practices as a confidence-building measure. Even though the 2013 report supported the initial idea, it kept a reservation on to what extend such information should be shared. That will be up to states (page 9 (26 (a)) [9]).

The 2013 report expanded the CBMs. It introduced the "points of contact" concept, the type of CBMs that were and still are being used within the military context, referring to "direct communications between capitals" (page 42 [5]). The 2015 report reflected on this point as well and highlighted that these "points of contact" should be on both policy and technical levels (page 9 (16 (a)) [10]).

Additionally, the 2013 report offered to conduct workshops, seminars and exercises, envisioning that these will help to identify, map and develop incident management techniques and mechanisms. It also pointed out that the cooperation between law enforcement agencies would facilitate the investigation of incidents "that could otherwise be misinterpreted as hostile State actions" [9].

To be able to map the developments of the CBMs within the ICT fields, Annex 2 has the colour coded illustration. The themes within the above mentioned three reports that relate one to another are highlighted by the same colour. This allows to visualise the common themes and showcase how these topics have been expanded within the time-frame between the year of 2010 and 2015.

1.3.1 Computer Emergency Response Teams (CERTs) as suggested Confidence-Building Measures

The UN GGE 2013 report is referencing the UN 2010 64/211 resolution, where three directions were stated for "incident management and recovery" [22]. Firstly, it was suggested to states to "identify the Government agency that serves as the coordinator for incident management". The purpose for this agency would be to develop and maintain "capability for watch, warning, response and recovery functions" (page 4, (19)[22]). Secondly, "identify national-level computer incident response team" that could be given responsibilities, tools and mechanisms to protect Government computer networks, and

"procedures for the dissemination of incident-management information" (page 4, (11) [22]). Thirdly, international cooperation, bilateral and multilateral, should be established that could potentially improve "incident response and contingency planning" (see page 4, (12) [22]).

The 2013 report then, as confidence-building measure, suggested that national CERTs should established bilateral cooperation for the purpose of exchanging information, and that CERTs should also contribute knowledge and expertise for policy and political decision making [9]. Later, the 2015 report incorporated two relevant CBMs. The initial suggestion of exchange of information between the national CERTs found its description in a more expanded version in 2015 report under the (d) point. In the above-mentioned report under point (c), the experts also suggested to establish national CERTs. There was also a suggestion that states could assign to CERTs a status equal to critical infrastructure. This status would allow CERTs to receive protection on much different level than any other organisation. The CERT would then be positioned under the aegis of the state, as a high priority asset. This will also allow states to manoeuvre, when and if necessary, their claim(s) under the international law.

The abovementioned report provided additional safeguards when addressing the norms, rules and principles for the responsible behaviour of states. The experts called for omitting/restraining to conduct any activity or support one that could harm the CERTs, as stated in the report "States should not conduct or knowingly support activity to harm the information systems of the authorized emergency response teams (…) of another State" (see page 8 [10]). And these norms and responsible state behaviour also refers to not only harming other states CERTs, but restraining to engage their own CERTs "A State should not use authorized emergency response teams to engage in malicious international activity" (page 8 [10]).
For the purpose of this thesis, I have selected the CBM on CERT cooperation (page 10 (d) [10]), to understand if it meets the set requirements, the characteristics defined for the CBMs, as well as use it further for the cross-examination in Chapter 4.

2 Cyber Conflict and Cyber Espionage

Before we start the discussion for establishing cyber conflict resolutions measures, it is required to have a comprehensive understanding of cyber incidents that constitute or might serve as precursor to cyber conflict itself. And, before we make conclusion about if there is an ongoing cyber conflict, violent or nonviolent, a more detailed and nuanced understanding of conflict in cyber space is necessary. It is particularly, within the scope of this thesis, important to understand, what type of cyber incidents are frequently occurring. For this purpose, the section below studies (a) the methodology how these publicly available datasets have been elaborated, and (b) analysing the typology of the incidents to define what is the most prevalent cyber incident.

2.1 Datasets of Publicly-known Cyber Incidents

To be able to detect characteristics of cyber conflict, datasets of incidents have been selected to inform the study. There are existing datasets of cyber incidents, but unfortunately those have been put together for various purposes and with various methodologies, in some cases with no consistent methodology at place.

This section of the study aims to examine six identified dataset(s), as those address cyber conflict and collect data relevant to this study. Some datasets do not necessarily address cyber conflict but analyze incidents more broadly.

By looking into already existing, publicly available datasets, analyzing the methodology based on which the datasets have been created, it provides an understanding on what the author(s) aimed to showcase. Why is this important? Because this process allows to follow the logic of the author(s) for the purpose of reflecting what aspect(s) of the methodology could potentially be relevant to current study, and what is not relevant.

For example, the Cyber Operations Tracker (COT) of the Council on Foreign Relations (CFR) [23] has a publicly available dataset of incidents, which has been put together "exclusively to track incidents such as denial of service attacks, espionage, defacement, destruction of data, sabotage, and doxing" [24]. The CFR mentions that their dataset encompasses incidents where the threat actor is known to be affiliated with a nation-state.

The Center for Strategic and International Studies (CSIS), as well, has made publicly available list of significant cyber incidents since 2006 that "focuses on cyber attacks on government agencies, defense and high-tech companies, or economic crimes with losses of more than a million dollars" [25]. The inclusion of economic crimes might seem irrelevant to international security, peace and stability, but when the incident triggers a public statement by the state authorities, thus escalating the incident to state level, this might result in potential conflict or cyber conflict.

Another publicly available source for cyber incidents is the Kaspersky Lab's Targeted Cyberattack Logbook [26]. It has been compiled by experts from the field, by the Global Research and Analysis Team (GReAT) and has the cyber campaigns that have been investigated by the abovementioned team.

These datasets are there, and the methodology for putting those together is different from one source to another. During the study, another dataset has been identified which was developed by a private person [27], but this dataset also includes cyber crime related data.

As Brandon Valeriano and Ryan Maness mention in their article "Coding Cyber Security Incident Data", that by collecting the dataset or making a list "is not enough to produce social science inferences or data analyses" [28]. They further discuss the common parameters [28] that datasets should have in place, to be usable for a bigger community. The authors, in a different publication discuss the importance of applying evidence-based approach to cyber incidents analysis: "we offer facts and evidence to help evaluate how cyber tactics have been used, will continue to be used, and will be used in the future"[29]. Additionally, authors mention that "to understand cyber conflict in the international relations realm, we must understand who uses the tactic, where, how, and for what ends" [28].

To provide a clear understanding of terms used within the current study, the definition of terms such as "data element" and "data point" are given below.
"Data element", for the purpose of this study, is regarded as a categorization of information that is being extracted from the facts about cyber incidents. For instance, the *"Type"* data element in the dataset would be the information on type of cyber operation. The "Type" is one of the data elements in the dataset.

"Data point", following the example given above, will be an independent piece of information, describing the "Type", which will constitute for instance an "espionage".

2.1.1 The Cyber Operations Tracker (The COT) dataset

The methodology behind the COT is based on multi-sourcing. They collect data from three other sources such as "(…) from existing repositories of state-sponsored incidents, such as Florian Roth's APT Groups and Operations spreadsheet, the Center for Strategic and International Studies' list of significant cyber events, and Kaspersky Lab's Targeted Cyberattacks Logbook" [30], the last two being analysed later in the thesis.
The COT dataset is compiled of over fifteen years of data, since 2005 and it is on-going. It has nine data elements in the dataset, which are the following: (1) date of report; (2) description of the incident; (3) suspected victims; (4) suspected state sponsor; (5) type of incident (e.g. espionage); (6) target category (e.g. government, private sector); (7) victim government reaction (yes or no); (8) policy response (e.g. criminal charges); (9) suspected state sponsor response (e.g. denial) and some additional references.

As an observation, the COT uses the phrase "suspected state sponsor" instead of "attacker" or "offender", in order to keep the language neutral. Since the CFR, within the scope of its mission, aims to inform about foreign policies, they have included the "policy response" as a data element into the dataset, to be able to detect policy development and/or policy course changes deriving from cyber incidents. Even though this element was included in the dataset, there were not many cases when this element was covered.

One recent example of data point inclusion is the case of the indictment of officials from the Mabna Institute [31][32]. What was impressive in the announcing remarks, is that it included "Accompanying Mitigation Efforts" what described the mitigation efforts, including revealing the vulnerabilities to private sector for further mitigation purposes [32]. This is an interesting development from the mitigation point of view and broadening the potential for public private partnership (PPP) efforts.

2.1.2 The Center for Strategic and International Studies (CSIS) dataset

The Center for Strategic and International Studies has the dataset in the format of a list made in chronological order since 2006, which provides a summary of incidents. Further, the CSIS, in collaboration with the Hackmageddon (a project/dataset to be described later in this section), has published a graph with a statistical data of number of incidents per

country, and 14 countries in total. The graph represents combined number of data points such as “cyber espionage and cyber warfare”[33], excluding cybercrime. The list of incidents has two types of data elements, such as “offender” (about 392 offences registered) and the “victim” (about 531 attacks, therefore victim(s) registered).
These data elements are concentrated around the states, providing statistical data, for instance, how many “offences” did Iran conduct and how many times have Iran been a victim of a cyber attack. The methodology behind the categorization of the data points is broad and does not provide details on “coding frames or dimensions”[34].

2.1.3 The Kaspersky Lab’s Targeted Cyberattack Logbook (GReAT) dataset

The Logbook has a different approach of providing data elements with regards to incidents. In general, this dataset’s main emphasis is on APTs[35][36], that “experts have chosen to share information about few samples belonging to the rarest and most menacing”[26]. This dataset is interactive and more sophisticated in its representation. It allows to filter data by type of APTs and number of targets (e.g. 100-1000, 1000-3000 etc.). It has the following data elements in the dataset: (1) name of the attack; (2) first known sample; (3) discovery; (4) current status; (5) type; (6) targeted platforms; (7) targets; (8) top targeted countries; (9) connected attacks; (10) the way of propagation; (11) purpose/function (e.g. remote control) and (12) special features which guides to an additional reading.

It is noteworthy to mention, as an observation, that this dataset has (1) technical data points and came from the (2) private sector experts. The fact that “targeted platforms” as a data element is included in the dataset contributes to information and knowledge sharing, for further purpose of initiating discussion on possible technical remediation, introducing standards and other forms of long-term solutions. For instance, The Moonlight Maze [26] APT, for the “targeted platforms” element mentions “Linux and Windows” as a data point, and the “special features” data element refers to a separate report [37] with technical details and evidence from a digital forensics perspective, as well as the details on the process of examination[38].
Further, a data element such as “connected attacks” brings additional value to information sharing, but unfortunately, there is no direct indication, based on which aspects of APTs the connections are made, presumably based on the artifacts found during the investigation.

2.1.4 The IMEMO dataset

The IMEMO dataset (**The** Institute of World Economy and International Relations), referred to within this study, has been identified in Information Security Threats during Crisis and Conflicts of the XXI Century [39] publication, related to threats analysis in informational sector of international relationships. The authors address cyber and ICT as "influence methods and new concepts of opposition with usage of informational and tele-communicational tools" [39]. Some examples from the publication are of the "Georgian-Ossetic conflict in 2008 and Ukrainian crisis in 2013–2015" (page 15 [39]), "Stages of Color Revolution, 2013-2014" (pages 68-71 [39]).

This dataset of incidents is presented in the form of a table and is being titled as "Examples of cyber-attacks of the XXI century" (page 106 [39]). The dataset elements are the following: (1) name of attack; (2) date of attack; (3) description; (4) damage and consequences and (5) supposed creator. The very first data point is "I love you" cyber attack from back in 2000, and with a last entry of "Wild Neutron" APT. The data points in this dataset are descriptive, unlike other datasets mentioned-above, and do not have a representation of categorization. It is a written description of the event or data, instead of specific data point entry, for instance "intrusion", "data destruction" etc.

2.1.5 The Hackmageddon dataset

The Hackmageddon dataset or project is being framed as "Information Security Timeline and Statistics"[40]. This project has an extensive of data, but unfortunately with little methodology in place. The cyber attack timeline is chronological and are from 2011 till 2018. The data elements in this dataset are the following: (1) ID number; (2) date; (3) author; (4) target; (5) description; (6) attack; (7) target class; (8) attack class; (9) country. The distinct data elements worth examining would be "target class" and "attack class". The "target class" data element has various inputs, as data points, such as "human health and social work activities", "information and communication", "education", "public administration and defense", "multiple targets", "individual" and others. Based on the observation of these inputs, and the information provided by the author, the categorization of these data points is based on International Standard Industrial Classification [41].

With regards to "attack class", the dataset includes following data points such as "cyber crime", "cyber espionage", "cyber warfare" and "hacktivism". Unfortunately, a

clarification, on how these data points have been developed, has not been provided. When looking into a dataset back in 2011, the data elements were different, for instance, the "attack class" didn't exist. This data element has been added later.

This dataset is dynamic, in terms of continuous changes and adjustments implemented by the author. The Hackmageddon project is an example of cyber incidents data collection as a private effort. It is an extensive dataset, but unfortunately this extensiveness leads to broad of a context to be useful for this study.

2.1.6 The Dyadic Cyber Incident and Dispute dataset

One relevant study that was made for the purpose of analyzing cyber conflict through a dataset [42], is elaborated in the Cyber war versus cyber realities: cyber conflict in the international system [29] publication. The authors, in Chapter 4, provide description of the methodology they have put together to analyze the cyber incidents, for the purpose to identify a "shift in the international cyber security landscape" [29] .

The authors have been collecting publicly available "cyber incidents and cyber disputes" (see page 84 [29]). This study has a comprehensive methodology and provides criteria for inclusion and exclusion of cyber incidents over a ten-year time-frame (between years of 2001-2011), as well as the methodology of coding the data. For instance, if the attribution of dispute was in a "serious doubt", the authors have not coded this incident as a state-on-state cyber operation. And this goes the other way around, the incidents are included in the dataset if a) the "state have admitted" or b) "cyber security companies have confirmed the involvement" (see page 84 [29]).

This dataset has a distinct categorization, it has the "cyber disputes" as a main category, that encompasses multiple "cyber incidents". Further, the data elements are the following: (1) cyber disputes; (2) cyber incidents; (3) rivalA; (4) rivalB; (5) name; (6) start; (7) end; (8) type; (9) method; (10) APT; (11) target type; (12) Initiator; (13) objective; (14) interaction type and (15) severity.

Each of these above-mentioned data elements has its own coding type, its own data categorization. For instance, the methods for cyber incidents are the following: (1) vandalism (e.g. website defacement); (2) denial of service; (3) intrusion; (4) infiltration; (5) APTs; (6) vandalism and denial; (7) intrusion and infiltration.

This dataset parameters have been developed, as authors mention, “based on history of relations, intent of the tactic, the likelihood of government complacency, (…)” [29]. Their analysis is “confined to rivals because they are the most disputatious members of the international system” (see page 86 [29]).

After an extensive analysis of the methodology of these six datasets, the author concludes that some of them are not detailed or meticulous enough to help to learn more about cyber conflict. The CSIS and the IMEMO do not offer filtering or categorization mechanisms to help to conduct analysis of typologies. The rest of the datasets are detailed enough to analyse the cyber incidents.

For the purpose of analysing the typology of the incidents, the author has chosen the COT dataset. The motivation to use particularly this dataset, is the following:

- The CFR is an independent think-tank providing an overview of foreign policies and state strategies;
- The COT has a clear methodology defined for data/incident collection;
- The COT has a comprehensive data collection since 2005 and has been continuously updated;
- The dataset is based on reports from various sectors, that include private sector analysis, technical details, tools, techniques and some repetitive methods used in the cyber incidents.

2.2 Typologies of Cyber Incidents

The purpose of this section is to identify and categorise major types of cyber incidents that are publicly known. For this purpose, out of six datasets, the Cyber Operations Tracker of the Council on Foreign Relations has been analysed. This dataset, as mentioned previously, incorporates data points from three other datasets, and applies concurrent filtering mechanism on the inclusion of cyber incidents typologies. The CSIS dataset has about 417 datapoints, and 114 datapoints are included in COT dataset, which constitutes 27.3% of CSIS total.

The typology of cyber incidents as per COT as follows:

Sabotage

The traditional definition of *sabotage* as follows "deliberately destroy, damage, or obstruct (something), especially for political or military advantage" [43]. In the current context *cyber sabotage* "refers to any sabotage activity facilitated by or using cyber space" [44]. This type of cyber activity was tagged only in COT dataset, but some similarities were identified in the Dyadic Cyber Incident and Dispute Data(set). The later, for instance, uses "Coercive objectives for initiators" [45] to tag Stuxnet as "degrade" type, but purely for the purpose/objective not for the type of activity. It defines *degrade* as an "attempt physical degradation of a targets' capabilities" [45].

Doxing

The doxing appears to be only in the COT dataset, and is defined as "Searching and publishing private or identifying information about an individual or group on the internet, typically with malicious intent" [46] which also matches the definition given by the Oxford dictionary [47].

Denial of service (DoS) | Distributed denial of service (DDoS)

The DoS/DDoS [48] as a type of cyber incident is included in the Hackmageddon, the Dyadic Cyber Incident and Dispute datasets, and the COT itself. The later defines in the Glossary as follows "Intentionally paralyzing a computer network by flooding it with data sent simultaneously from many individual computers" [46].

Defacement

The CFR doesn't provide definition for this type, but commonly used definition can be found in various sources [49]. The primary definition for *defacement* would be "Spoil the surface or appearance of (something), for example by drawing or writing on it" [50]. Defacement as a type is present in the COT, the Hackmageddon datasets. The Dyadic Cyber Incident and Dispute dataset included as well, but has been coded as an isolated incident, defined as methods of cyber-incidents and marked as "vandalism" [45].

Data Destruction

There is no definition in the CFR Glossary for *data destruction*, but the common characterization is "Data destruction is the process of destroying data stored on tapes, hard disks and other forms of electronic media so that it is completely unreadable and cannot be accessed or used for unauthorized purposes" [51] or in the case of a cyber incident, used for authorized purposes. This type of cyber activity is included in the COT and in the The GReAT Logbook as a "data destroyer".

Espionage

Typically, *espionage* is defined as "the practice of spying or of using spies, typically by governments to obtain political and military information" [52]. According to ENISA's Overview of Cybersecurity and Related Terminology, *cyber espionage* is understood as an activity that can have "two types of espionage vectors: (a) state espionage (intelligence, when state actors are involved) or (b) industrial espionage (when commercial actors are involved)" [44]. The *espionage* as a type of cyber incident can be found in the CSIS, the Hackmageddon, the Targeted Cyberattacks Logbook (GReAT) and the Dyadic Cyber Incident and Dispute datasets. The GReAT Logbook as a type of cyber incident has *cyberespionage toolkit*, and the Dyadic Cyber Incident and Dispute dataset has *long-term* and *short-term espionage* as an objective. Additionally, the authors of the Dyadic Cyber Incident and Dispute dataset as "methods for disputes" categorize *intrusion* and *infiltration,* and *intrusion* is mentioned as a method for *espionage* [45] [42]. Despite that the CFR's Glossary separates two types of espionage: (a) cyber espionage and (b) industrial espionage, when categorizing the cyber incidents, it marks *espionage* solely. It defines *cyber espionage* as "The use of computer networks to collect information on the activities, movements, and plans of a target" [46], and *industrial espionage* as "Spying directed toward discovering commercial secrets from a rival manufacturer, other company, or held by a government" [46].

2.3 The Prevalent Cyber Incident

The statistical analysis of the COT dataset shows that between 2005 and October of 2008 in total 288 cyber incidents are registered in the dataset. The figure below illustrates the dynamics of the incidents annually.

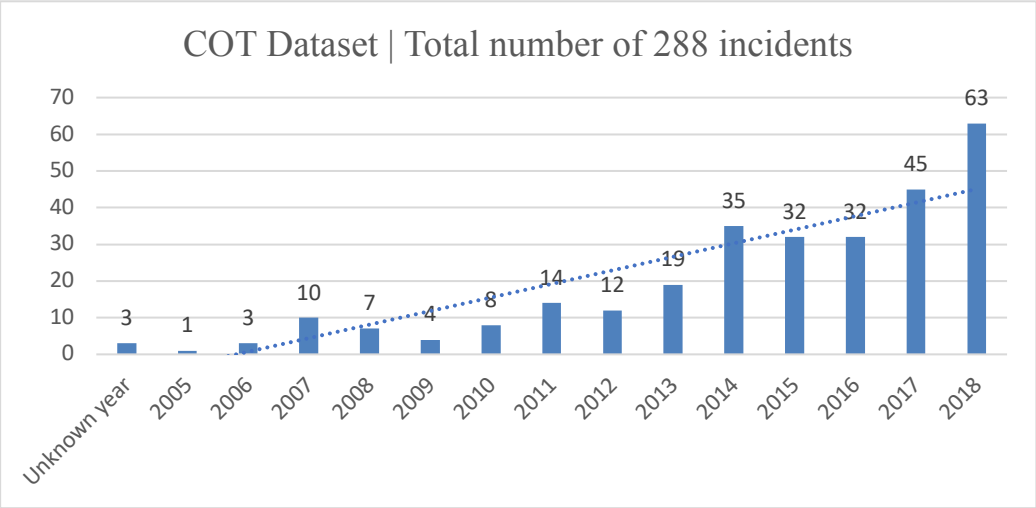


Figure 1. Chronological Development of Cyber Incidents based on COT dataset (between 2005-2018 October)

Based on the typology analysis, below are presented the numbers. It is immediately noticeable that espionage overweighs all the other types of cyber activities.

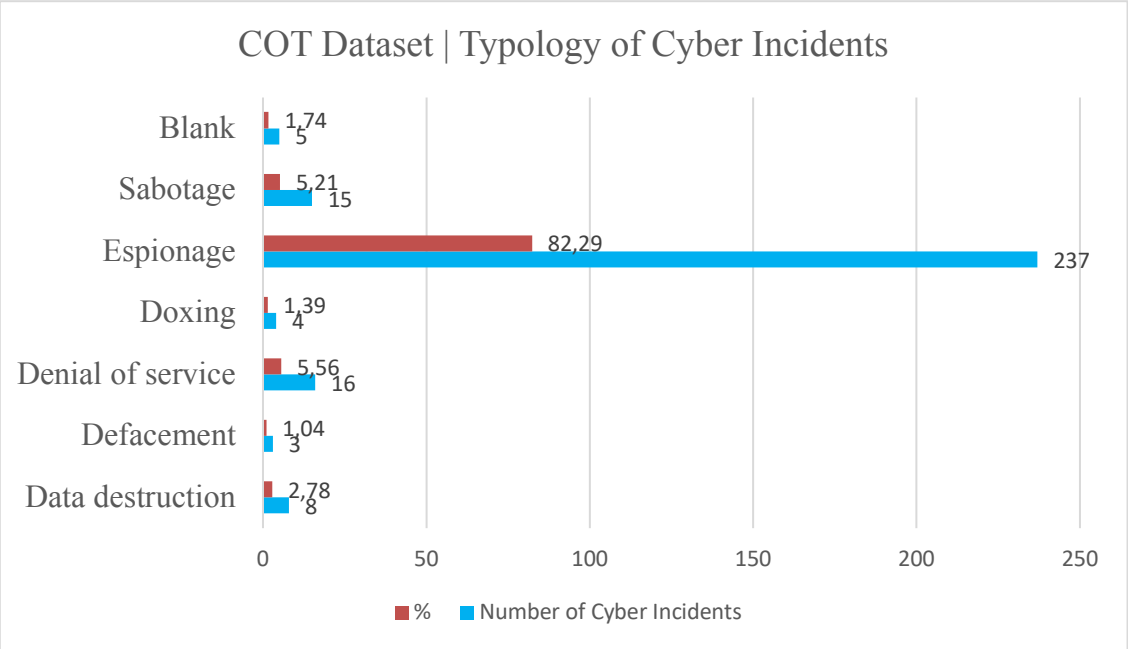


Figure 2. Typology of Cyber Incidents based on COT dataset

To validate this finding, I have excluded the CSIS datapoints from the COT dataset, and results showcase that still the prevalent type of cyber incident in the COT dataset is cyber espionage.

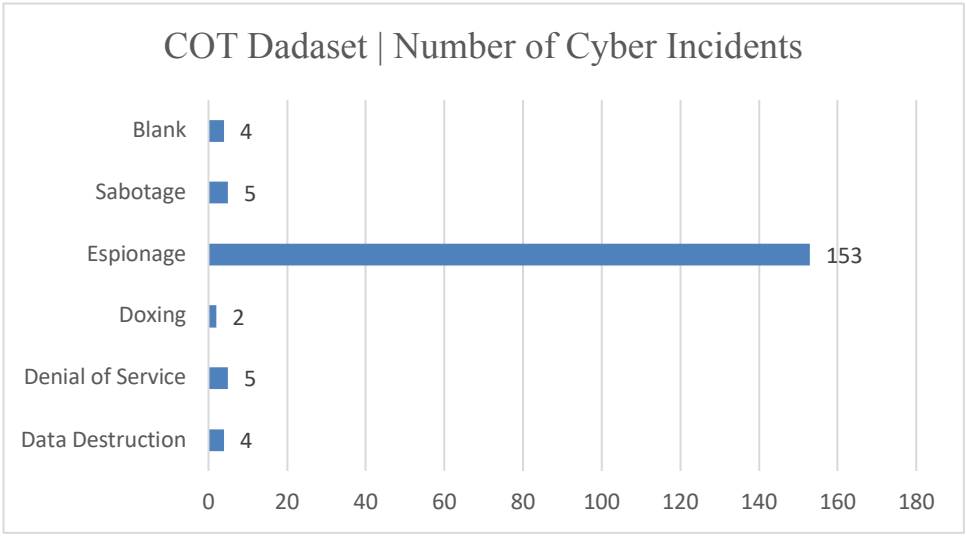


Figure 3. COT dataset excluding CSIS datapoints

It is important to mention that the Council on Foreign Relations highlights the limitations of the Cyber Operations Tracker (the CFR dataset). Firstly, there is a limitation with regards to techniques and methods applied for attribution, and the CFR defines the attribution for "(…) suspected threat actors and their state sponsors based on what the reporting suggests and whether the tools, techniques, and procedures used by the threat actor conform to what is known about a state sponsor's preferred methods of intrusion" [30]. Secondly, there are no claims made that the dataset is complete. It includes incidents that were made (a) public and (b) communicated widely in English [30], thus making the dataset western-centric. And thirdly, the scope of the evidence can potentially change by the time, thus affect the initial attribution [30].

By analyzing the typologies if cyber incidents, it is evident that the vast majority of the cyber incidents that are state sponsored, according to the CFR methodology, cyber espionage is prevalent. The author conducts the further analysis upon this finding and creates a model of cyber espionage and uses the ontology of cyber espionage as a test case to discuss how the CBM of CERT cooperation can be applied.

2.4 Ontology (conceptual model) of Cyber Espionage

To be able to apply the proposed CBM in the context of cyber espionage, the author firstly defines what cyber espionage is and what elements does it have. The author uses the term "intelligence gathering" or "intelligence process" throughout this subsection to discuss and analyse what constitutes "espionage" and "cyber espionage".

When reflecting about the "intelligence process", it does not have a value on its own, it is for the purpose of making decisions. And so, it is a process that is taking place in parallel to decisions making, where the latest can relate to high political level, economic level, military or any other levels of decision making. For instance, recent industrial espionage cases, the compromises the networks of SingHealth, the largest health-care provider in Singapore for the purpose of obtaining the pharmaceutical prescriptions of 160,000 patients [53] [54]. This example testifies that some threat actors are motivated to learn about the medical prescriptions to patients. This information can be useful to a range of different corporations, one of them could be insurance companies. Potentially other pharmaceuticals could be interested to learn what type of medicine is being commonly prescribed and for the decision makers to approve the development of similar solutions for a much cheaper price, to be able to overtake certain market shares. Another espionage case is recorded primarily in the oil, gas, and electricity production in the Gulf region [55][56].

Others find motivation or need in learning and/or stealing defence related classified information, research and development project can come in handy for duplicating military technologies and/or weapons. For instance, private sector cyber espionage case of the networks of a contractor for the U.S. Navy. These operations resulted in exfiltrating 614 gigabytes of sensitive information on the new U.S. anti-ship missile and submarine communications systems [57] [58].

And so this "intelligence process" for gathering data is for the purpose to make economic decisions, decisions that relate to defense, and also political decisions, when targeting states institutions. For instance the compromise of the Finnish Ministry of Foreign Affairs back in 2013 [59], and particularly the communications between the EU and Finland [60] [61]. Another example of the compromise of the German Foreign Office [62].

It is important to highlight, where there is a decision making, there exists the "intelligence process". Before the "intelligence process" is initiated, the "tasking" or "intelligence order" or "request" takes place.

The "intelligence process" is comprised of other general stages, or as C.Falk mentions in the Ontology for "Threat Intelligence paper", the "Intelligence is a cyclical process" [63]. According to Falk, the cycle, the process is comprised of the following presented in Figure 4 below:

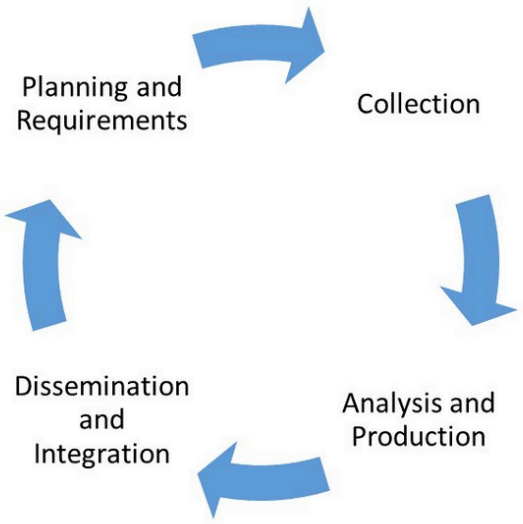


Figure 4. Source C. Falk (2016) - Demonstrating the nature of the intelligence cycle

To summarize, the "intelligence process" has four stages:

a) Planning: takes place for the purpose of defining or identifying and acquiring devices and /or tools or activities.

b) Collection or gathering: takes place for a long period of time, it may even take years. At this stage the gathered data is raw.

c) Analysis: at this stage, the raw data is being analyzed. The analysis can be conducted in various formats, such as statistical or linguistic analysis, deductive or cross analysis.

d) Dissemination or delivery: this stage happens when the analysis is concluded, and so the "intelligence" can be delivered. There might be cases when an order given by a decision maker to deliver raw data urgently, without analysis. For instance,

if specific criteria are met or specific target were identified (e.g. weapons of mass destructions have been identified).

To visualize conceptually how this takes place, and to combine the decision making with the general stages of the "intelligence process", the figure below presents the high-level abstraction process, without details on what the general processes, the stages are constructed of.

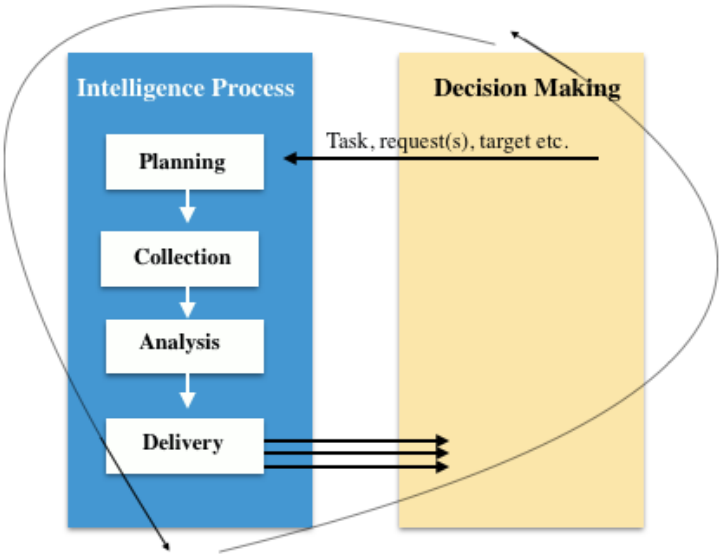


Figure 5. Intelligence Process Complete Cycle developed further based on the suggested model from the following source: C. Falk (2016)

The "Decision Making" process and the "Intelligence Process", these two together make an independent cycle on its own. This cycle restarts when a new order, request or task is being issued. The task, request or the target may vary every time, but the cycle is the same in terms of its stages and it is continuous. Despite that the "Decision Making" takes in parallel to "intelligence process", in terms of positions, the "Decisions Making" is on top of above the "intelligence process", it is a hierarchical relationship. It is the ultimate decision maker that may decide the specific activities or provide restrictions or provide a scope for the "intelligence process".

It is important to bear in mind that this is a conceptual theoretical model of the relationship between the "Intelligence Process" and the "Decision Making". In reality some stages of within the process might take place differently.

After the discussion on general conceptual model of the espionage or "Intelligence Process", a model illustrated below can be used to refer to an ontology, a conceptual model of the cyber espionage (Source: Interview with Dr. Kerttunen (2019)).

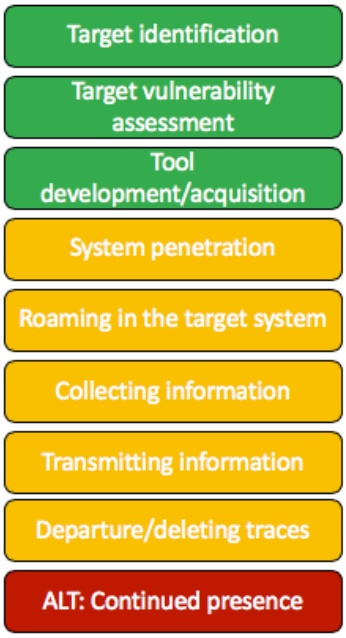


Figure 6. Source Kerttunen (2019)

This model presents the stages of the "Intelligence Process" without "Decision Making", it is only for the operational part of the cyber "intelligence process", of the cyber espionage.

The color coding in this model indicates different stages. The first three such as 1) target identification, 2) target vulnerability assessment and 3) tool developments and acquisition make part of planning stage. The following five actions make part of implementation stage, in a broader context, which then can be split into three separate phases. The "system penetration" and "roaming in the target system" could be considered as exploitation phase, the "Collection of information" could be collection phase, "transmitting information" would be the combination of analysis of data and delivery of data.

The last action is an alternative, it may happen that there is no departure planned. The last action suggests longer presence if necessary.

By conceptualizing cyber espionage, it becomes visible what are the elements that can be used for further analysis. The proposed CBM of CERT cooperation can be now applied on this conceptual model, using the ontology of cyber espionage to understand how it will minimize the misunderstanding between states.

This methodological approach can be used for applying to any other type of cyber incidents to evaluate the effectiveness of the proposed CBMs.

3 Analysis of the proposed Confidence-Building Measures

In the first chapter, the author has discussed and analysed what constitute a confidence-building measure deriving from the comprehensive study conducted on CBMs [7]. In the second chapter, the author has identified publicly available datasets [30] [33] [26] [39] [40] [42] and analysed typologies of publicly known cyber incidents to identify which type of cyber incident is the most prevalent. This third chapter is dedicated to the core analysis of the proposed CBMs for the ICT environment, basing the analysis on the outcomes derived from Chapter 1. The author has briefly introduced the proposed CBMs for the ICT field in Chapter 1, but in this chapter the selected CBM of CERT cooperation will be analysed. To highlight, the author is not concentrating on one specific UN GGE report for the analysis, rather choosing one theme specific suggested confidence-building measure.

The author has selected the CBM of CERT cooperation because most countries [64] [65] [66] have established national CERTs. Some developing countries are in the process of following the trends for creating CERTs. The author will be examining the CBM proposed by the UN GGE as to whether it corresponds to the formal criteria and requirements that define the effectiveness of confidence-building efforts mentioned in Chapter 1.

The selected CBM from the 2015 UN GGE report included a proposition for "(d) Expand and support practices in computer emergency response team and cybersecurity incident response team cooperation, as appropriate, such as information exchange about vulnerabilities, attack patterns and best practices for mitigating attacks, including coordinating responses, organizing exercises, supporting the handling of ICT-related incidents and enhancing regional and sector-based cooperation" [10].

3.1 Does the proposed CBM for CERT cooperation meet the formal criteria and requirements defined for effectiveness of the CBMs?

In this section the author analyses how the selected measure meets the criteria of a CBM using (a) textual interpretation and negation, if necessary, and by (b) teleological argumentation to conclude to what extent the proposed measure meets the criteria of CBMs as discussed in Chapter 1.

3.1.1 Addressing Regularity

(a) Textual interpretation

When looking at the proposed CBM text "expand and support practices for exchange of information", "*organising exercises*", "*enhance regional and sector-based cooperation*" these activities are not being explicitly requested with certain regularity and pattern. The experts didn't suggest for instance, conduct annual or quarterly exercises to share knowledge. The timeline for the proposed activities is undefined. Based on textual analysis, the requirement of regularity is not met.

However, the absence of explicit reference to regularity the proposed measure refers to the mentioned activities and events in plural, which allows suggesting that regularity is not ruled out in the wording. Negation is used to further investigate the element of regularity.

According to dictionary [67], the opposite to regular is occasional. Accordingly, the author asks whether it follows from the language of the measure that the proposed CBM is occasional? There is nothing in the language of the measure that would rule out regularity. The author concludes that based on negation, the requirement of regularity is met.

(b) teleological argumentation

The aim of the proposed measure is to develop an "international" (see page 2 [10]), "regional" (see page 10 [10]) and "interstate" (see page 9 [15]) cooperation with regular implementable and practical activities at place, such as (1) exchange of information on vulnerabilities and attack patterns. Exchanging knowledge and (2) "best practices for mitigating attacks" should imply that this is realised on a regular basis, as the patterns of attacks changes, so as the mitigation mechanisms. In terms of (3) "organising exercises" one should bear in mind that it is not only the same people participating in the exercises, but also newcomers. The regularity in these exercises is a key to consistently convey the knowledge and practice skills. Moreover, the *"expansion and support practices for exchanges of information and communication*" cannot appear sporadically, as well as *"organising exercises"* cannot be arranged randomly. In this analysis, it becomes evident that at least some coherence and consistency of ordered activities is presumed, for

instance “as requested”, “as agreed”. As to *“enhancing regional and sector-based cooperation”* cannot occur irregularly, it should be implemented regularly and consistently. This is an on-going, multi-sector commitment that without regularity will simply perish. Taking into consideration all the above-mentioned arguments, the requirement of the regularity can be marked as met.

3.1.2 Addressing Continuity

(a) Textual interpretation

Based on the textual analysis, there is no explicit reference to continuity in “*information exchange*”, “*coordinating responses and organizing exercises*”, “*supporting the handling of ICT-related incidents*” or “*enhancing regional and sector-based cooperation*” [10]. There is no specified sequence or stability proposed by the experts for the implementation indicated measures, or a certain chain of activities that could form the continuity. Based on the textual interpretation the requirement of continuity is not met.

The author uses “intermittence” [68] to discuss if the proposed measures could be implemented with interruption or intermittently. There is no implication that, for instance, the “*supporting the handling of ICT-related incidents*” shall happen intermittently, or no rejection or dismissal of the continuity in general either. Therefore, based on negation, the requirement of continuity is not ruled out.

(b) teleological argumentation

The continuity should be a goal and should be at the core of coordinating ICT incident responses as postponing or pausing the response may result in crucial consequences. In reality it may require more than a 24-hours coordination at CERT or at CIRT to tackle an emergency or an incident. It may also require weeks or months continuously to analyse the incident, come up with remediation mechanism(s) and to actually implement the remediation. If the attack has targeted more than one state, then the same argumentation should be applied to “supporting the handling of ICT-related incidents”. The author points that the aim of the proposed measures to create a continuous coordination and support mechanism(s) to achieve effective remediation in whatever form it would be. Basing my analysis on teleological argumentation, the author concludes that the requirement of continuity can be implied as met.

3.1.3 Addressing Precision

(a) Textual interpretation

The example given for the annual exchange of military information to be conducted by December 15 of each year [5] is as precise as possible. The current proposed measures don't address the implementation with precise schedule or deadline. These are suggestions for actions such as "expand and support practices for exchange of information" or "organise exercises" without precise and fixed time indication(s). Based on textual analysis, the requirement of precision is not met.

The author uses negation to verify if the "information exchange about vulnerabilities" can be conducted in a vague time-frame. There is no implication that the exchange of information should happen in a vaguely or undefined time-frame. Based on negation, there is nothing in the proposed CBM that would rule out the precision.

(b) teleological argumentation

The aim of the suggested CBM is to "strengthen cooperation" by sharing information on threats. To be able mitigate incidents efficiently, the information sharing between CERTs should take in a promptly manner. Considering the importance of the matter and what impact could be obtained by sharing information or intelligence, the proposed CBM should have a precise operational procedure or, as known standard operating procedures (SOP) for the efficient response and coordination. As this proposed measure is a suggestion for a CBM, it has a potential to be elaborated further and to include precise time-frame or deadline for informing partnering states on malicious activities taking place. This precise time-frame will be defined by the partnering states, states should agree what would be the optimal time-frame for sharing information for efficient mitigation. Basing my analysis on teleological argumentation, the author concludes that the requirement of precision can be implied as met.

3.1.4 Addressing Specificity

(a) Textual interpretation

This CBM proposes an expansion and support of CERT/CIRT practices and cooperation, and suggests certain areas, which are specific. It does not specify the methods of implementation but discusses the areas. Bearing in mind, that the discussed measure is a

proposition for an actual CBM, the experts succeeded to propose specific areas and topics. Their proposition included support for CERT practices and CIRT cooperation, which was narrowed down to:

(1) information exchange about vulnerabilities, attack patterns;

(2) best practices for mitigating attacks, including coordinating responses;

(3) organizing exercises;

(4) support the handling of ICT-related incidents;

(5) enhance regional and sector-based cooperation.

This CBM is specific by its areas of the implementation and based on the textual interpretation the requirement of specificity is met.

(b) teleological argumentation

This CBM aims at narrowing down specific areas of collaboration, it does not mention *support* or *cooperation* or *enhancement* in general, but it points out specific directions where collaboration can help to mitigate cyber attacks. The "information exchange about vulnerabilities" requires strong partnership and trust built between two or more parties. This type of partnership does not happen randomly but based on specific arrangements between states. Informing other states CERTs about vulnerabilities identified or the type of cyber attacks are being experienced requires strong cooperation at place. The experts highlighted specific areas, for instance, sharing "information on attack patterns" for then to be developed further by states. States may initiate discussions on which methods or which platforms will be used for sharing information, which type of information should be prioritized and similar questions that require further discussions. Taking into consideration that this CBM is a proposition for actual measures, based on the argumentation above, it can be concluded that the requirement of the specificity is met.

3.1.5 Addressing Presence

(a) Textual interpretation

The text of the CBM is guiding states on what should be done, considered or promised in terms of "behaviour in the future" as the proposed measure is a proposition for a CBM.

This CBM is not a currently existing measure, but an offer what to consider for implementation. Therefore, the requirement of the presence is not met.

To apply negation, the author asks if this proposed CBM is non-existing. As this CBM is a proposal/recommendation for a CBM, the author concludes that the actual measure is currently non-existent, therefore the requirement of presence is not met.

(b) teleological argumentation

The goal of the proposed CBM is to be practical, functional and exiting as they suggest "information sharing on vulnerabilities and attack patterns", "organizing exercises" and "supporting the handling of ICT-related incidents". The aim of the measures should be current, existing and on-going, else they will cease their functioning goal. Basing the analysis on the goal of the proposed measures, the author concludes that the requirement of presence is met.

3.1.6 Addressing Systemic approach

(a) Textual interpretation

When looking at the proposed measures, indication of horizontal and vertical implementation of measures is present. The experts suggested measures as "information exchange about vulnerabilities and attack patterns" and "organizing exercises". These two measures are independent one from another and can be implemented in parallel, by making it an independent system on its own. A parallel can be drawn between *information exchange of military activities* along with *exchange of annual calendars* or *military budgets*, as given in the Vienna Document [5]. Based on the textual interpretation the requirement of systemic approach is met.

3.1.7 Addressing Measurability

Looking at the proposed CBM, it does not offer to organize, for instance, "four" exercises, or that the support of the handling the ICT incidents should be accompanied by at least "two" cyber security engineers from each national CERTs. This CBM does not include any measurable output or milestone. Although it does not host a measurable activity, it is still possible to measure the outcome of the proposed CBM, basis on the fact if the proposed CBM has been implemented or not. For instance, the result or outcome can be marked as positive or as achieved and makes the result measurable. Basing the analysis

on the textual interpretation, the author concludes that the requirement of the measurability is partly met.

The analysis of the proposed CBMs is complex, therefore the author used different methods to uncover its effects. The textual interpretation didn't allow to accommodate fully correspondence to the formal criteria and requirements defined for the effectiveness of the CBMs. When applied negation to the proposed CBM, it did not rule out the requirement either. More over some requirements out of the seven (Regular, Continuous, Precise, Specific, Present, Systemic, Measurable) were meet. The teleological argumentation revealed more of the potential effect that the proposed CBM may have.

The recommendations for the improvements are the following: firstly, for the purpose of ensuring the effectiveness and impact of the CBMs, it is important to include the decision makers in the process. A recommendation would be to that the proposed CBM of the ICT environment address activities targeting decision makers as well. Additionally, suggested measures could address psychological aspects of cyber incidents. These aspects can be included, for instance in organising the exercises, potentially for both the technical teams and the decision makers.

Moreover, the recommendations to improve the CBMs can include:

- For the purpose of developing these measures further, an aspect or characteristic of regularity can be incorporated within this CBM in different ways. For instance, "enhancing sector-based cooperation" through "organising exercises" can be conducted on different levels. Firstly, on the national level national CERT and CIRTs (cybersecurity incident response teams) of the private sector organisations can organise exercises annually. On regional level for the purpose of "enhancing regional cooperation" national CERTs can organise exercises twice a year.

- Aspects of continuity can be incorporated in the proposed CBM by introducing a 24-hour-a-day basis point of contact for interstate or multi-state cyber incident investigations for the purpose of "handling the ICT-related incidents". Additionally, continuously "supporting the handling of ICT-related incidents" by the states can be conducted by the revising, approving or financing the operations of the national CERTs.

- For instance, exchange of vulnerabilities should be conducted not later than two days after it was discovered. The exchange of vulnerabilities should be conducted in accordance with CVE scale (Common Vulnerabilities and Exposures) and the time-frame agreed between the states.

4 Applying CBMs to Cyber Conflict: A Methodology

Although the UN GGE has not specified any particular type of cyber incidents that the CBMs should be applied to, the statistical analysis of the datapoints in the COT (the CFR dataset) indicates that 82% of the cyber incidents is cyber espionage (see Figure 2). In this chapter the author uses the ontology of cyber espionage, as an example of conflict, or precursor to cyber conflict, to apply the proposed CBM and evaluate if the CBM of CERT cooperation will minimize the misunderstanding that might lead to unintentional escalation between states.

This analysis is conducted according to conceptual characterization "that confidence-building measures are concerned with the security perceptions which States have in relation to each other" (page 25 [7]). But simply defining perception of security is not enough in this context. Defining the perception of a threat completes the situational awareness of what is security and what is threat for states. And the both perceptions (security and threat) are dependent on various factors. And so "the causes of mistrust vary from region to region or even within the same region" (page 4 [7]). The mistrust can relate to "complex of historical experiences, as well as geographical, strategic, political, economic, social and other elements" (page 4 [7]).

In the first section of this chapter, the author uses the stages of cyber espionage ontology to discuss which actions could potentially create misunderstanding and discusses how specific actions can create misperceptions, mistrust, lead to loss of security that can potentially result in tensions.

4.1 Misperceptions | Misunderstandings

As defined and modelled in Chapter 2, cyber espionage has stages. Each stage is comprised on actions. The generic stages, excluding the "Decision Making", would be *Planning* and *Implementation*. The *Planning* would be comprised of the three actions: 1) target identification, 2) target vulnerability assessment and 3) tool developments and acquisition. And so, when looking at the "planning" as a whole, how sensitive this stage can be? Can this stage trigger reactions or assumptions by the target?

4.1.1 Planning stage

The *Planning* stage could be regarded as an internalized process of preparation. It is the preparation based on the given task or order. So, the task or the order, that is being issued by the "Decision Makers", can be precise (as given target) and/or generic (the selection of tools). This preparation, the planning may have tentative interaction(s) with target (target as a broader concept, not one person). For instance, the "*target identification*" and *"target vulnerability assessment"* will require actions such as: "harvesting email addresses, identifying employees on social media networks, discover internet-facing servers" [69]. These actions, such as discovering internet-facing servers may create noise in the network. If the target's systems are correctly preconfigured for alerting scanning activities, thus the scanning of the network will trigger alerts. What perception can then a target acquires from these actions?

Behind the perception or misperception on what is happening in the network can be various factors. One should not forget that the perception or the misperception is an interpretation made by a human, and "we are quick to reach interpretations, to tell others (and ourselves) stories about what is happening, and to explain puzzles as soon as we can" [70]. And so, there can be a perception that an adversary is preparing for an attack, which is a misperception is this case.

The "attack" in his context constitutes actions for the purpose of destroying, altering or in any other damaging data and/or systems. The intruder's end goal is to collect, gather data, and damaging or altering the data is not the intention of the intruder. The actual penetration of the network can be regarded as an attack, but within this context of the analysis, the attack means actual alteration or destruction of data, devices or infrastructure.

For this current example, the intention of this specific activity is scanning, is a mapping activity that can be regarded as listing unattended, unprotected devices. But the intention behind this activity is not to attack or in any way harm the system, but to learn from the system.
One possible factor behind the perception or misperception can be that that the target believes it can be a potential target of a cyber attack. The believing of being a potential target can be set by the high-level decision makers, that have created and implanted that

idea. It can be concluded that not all actions of the planning stage can create misunderstanding or misperception of the reality. Some actions, that have direct contact with target may cause misunderstanding or misinterpretation of the actions.

4.1.2 Implementation stage

The *Implementation* is broader and can be regarded as exploitation, collection, analysis of the data and the delivery. This stage is much hands on than the previous stage, it engages with the target directly. And because it has direct contact with the target, the sensitivity can be high, and the reactions can be prompt.

When the exploitation phase, referring to "system penetration" and "roaming in the target system" actions for the purpose of data collection, is discovered, it may create misunderstanding and misperception for the target. This misperception can relate to the "intruder's capabilities and intentions" [71]. So, what may the target infer and what reaction this can trigger by the target?

One example of possible inference drawn from cyber espionage, discussed by Martin C. Libicki is the following "because a malware implant designed for cyber espionage is often identical to one designed for cyber attack, discovering and attributing one in a critical system could easily be viewed as a direct precursor to attack" [71]. Consecutively, this can lead to that the target raises "its alert levels, which, in and of itself may exacerbate tensions" [71]. By creating tensions, it will complicate the relations between two parties, and may lead to overreaction and unintended escalation. The factors that are behind this reaction can relate "to perceptions of threat which form an additional psychological component" [7]. One example of overreaction is given, of the American cruiser Vincennes shot down an Iranian airliner over the Persian Gulf in 1988 [70]. The reason why this has happened related to the staff being intensely "trained to expect an attack and to be hypervigilant" [70].

Based on the inferences made, the victim may "react" or a decision can be issued to react. The reaction is based on psychological aspects – because for not acting, some would imagine the repercussions of passive behavior, despite the fact that the intention of the intruder is not confirmed yet. As R.Jervis mentions in his book "Leaders may also pay a price domestically for hesitating because they may be seen as weak and indecisive" [70]. There can also be a political pressure to act quickly without reconsidering or reevaluating

the situation and “in most political and organizational contests it is particularly difficult to do this because the leader has to act confidently in order to inspire confidence in others” [70].

What if the action of “departure” doesn’t take place? And what if the victim cannot find any proof that the “departure” has taken place? In Chapter 2 when modelling cyber espionage, the ontology included the alternative phase the “continued presence”. How this phase can mislead the victim and threaten the sense of security?

Discovering technical evidences depends on the maturity of the technical teams. If there has been no evidence found that the “departure” has taken place, this situation can threaten the sense of security for the victim. The victim might obtain a perception of being potentially constantly monitored and may consider reacting, instead of delaying reaction to anlayse the situation and proceed further and deeper into evidence hunting. This behavior can be connected to, not explicit, but to consideration of “the mental costs of delaying (…) image” [70]. As R. Jervis discusses, the “sense of being confused or even confronting a puzzle is usually uncomfortable, in part because the faster we can make up our minds the sooner we can turn our intellectual energies to other pressing matters” [70]. For that reason, for the victim it might be harder, psychologically pressing to wait longer.

4.2 Cross-Examination: Application of CERT CBM to minimize the misunderstanding between states

This sub-section presents the methodological approach to evaluate the effectiveness of the proposed CBM. The cross-examination implies that the both components or parties of the examination are defined, and elements are identified. The ontology of cyber espionage allowed to define the general stages, phases and the actions. In the previous section, the author has discussed potential misunderstanding and misperceptions that can arise within different phases or from actions in the cyber espionage model.

To proceed with the cross-examination, the author describes the elements of the both parties of the examination. The author has identified five elements within the proposed CBM that are the following: 1) Information exchange about vulnerabilities, attack patterns and best practices for mitigating attacks; 2) Coordinating responses; 3)

Organizing exercises; 4) Supporting the handling of ICT-related incidents; 5) Enhancing regional and sector-based cooperation.

The specific elements, but not the generic stages, from the defined ontology of the cyber espionage in the chapter 2, are the following: 1) Target Identification; 2) Target Vulnerability Assessment; 3) Tool Development/Acquisition; 4) System Penetration; 5) Roaming in the Target System; 6) Collecting Information; 7) Transmitting Information; 8) Departure/Deleting Traces; 9) Alternative: Continued Presence.

The schematic below explains the methodology of the cross-examination. Based on this schematic the application of the proposed CBM is conducted to discuss the effectiveness of the proposed CBM.

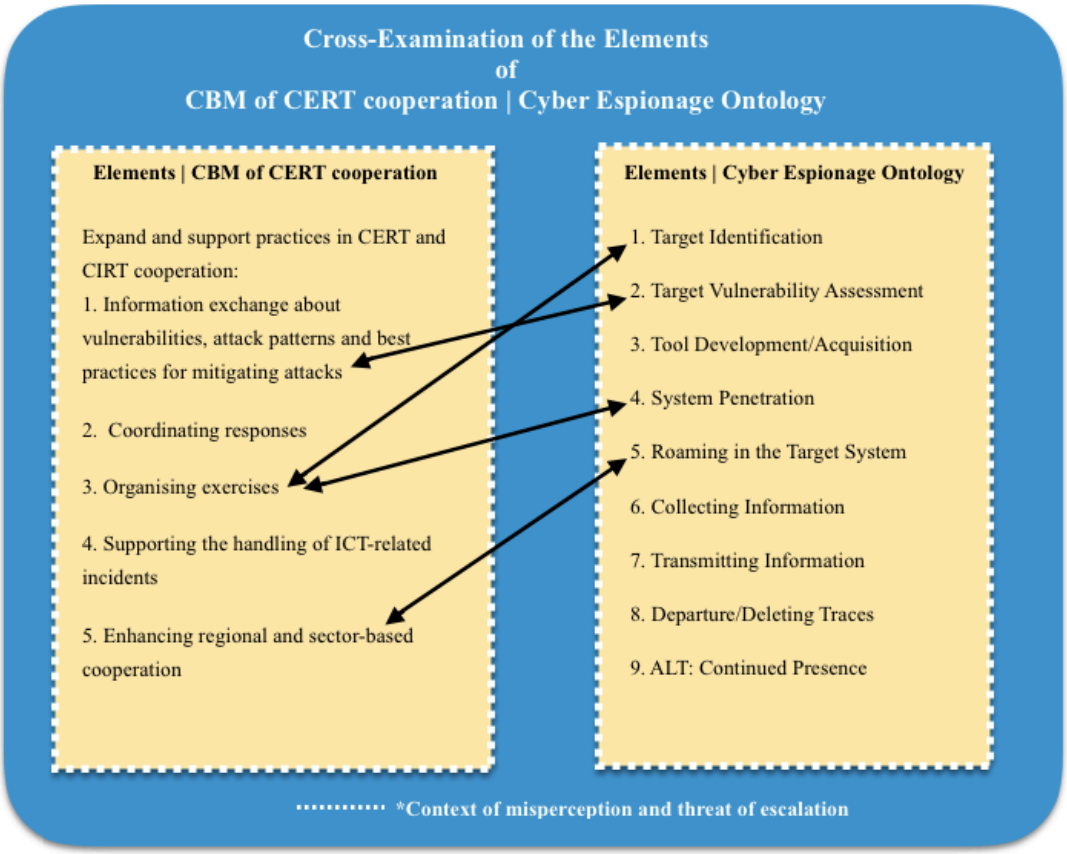


Figure 7. The Schematics of the Cross-Examination. Using the proposed CBM (UN GGE 2015 report) and Ontology of Cyber Espionage (Kerttunen, 2019)

A question to ask - how can information exchange about vulnerabilities, attack patterns between CERT and CIRT help minimize the misunderstanding, eliminate the

misperception that an intruder is preparing for an attack during the planning/preparation stage?

Firstly, it is important to identify to which type of CERT/CIRT does the proposed CBM refers to. As an observation, the 2013 UN GGE report explicitly mentions "national CERT", but unfortunately, the 2015 report doesn't specify which particular CERT should be engaged in cooperation and information exchange. For the purpose of this analysis, the author will refer to national CERTs.

By talking to experts in the field, it is understood that CERTs/CIRTs or CSIRTs vary by their constituencies and the mandate for operations. The European Union Agency for Network and Information Security (ENISA) has recorded over 387 CSIRTs (as of May 11, 2019), and the interactive map allows filtering of the CERTs/CSIRTs based on the type of constituencies [66] (statistical analysis of the constituencies based on the ENISA dataset in Annex 3). For instance, CERT of Austria identifies themselves as "as information hub which knows where to send the right incident reports to in order to help and facilitate the clean-up of IT security incidents" [72]. Their definition of the constituency is broad "The constituency of CERT.at is basically the whole country of Austria" [72].

In opposite to the Austrian CERT, the CERT of Luxemburg defines their constituency "CIRCL (Computer Incident Response Center Luxembourg) is the CERT for the private sector, communes and non-governmental entities for the Grand Duchy of Luxembourg". Despite that their mission is to provide services at a national level, they do not serve the governmental agencies. The more the types of CERT are being studied, is it being concluded that there is no unified approach when establishing CERTs and assigning the scope of the mission. Some nation states have only national CERTs, some have accelerated in establishing national, governmental and military CERTs in parallel.

If, for instance, the military or the governmental networks have been the target of the cyber espionage, and their networks are being scanned (as part of the *Planning* stage), then the exchange of information about vulnerabilities, attack patterns and best practices between national CERTs/CIRTs is not relevant to this case and particularly reducing the misunderstanding between states.

In the example given above, where, during the planning stage the perception or misperception is occurred for a potential attack to happen, the national CERT has no mandate of resolving this misunderstanding, firstly because the scope of the cyber incident is outside of its constituency. The information or knowledge still can be forwarded to military or governmental CERTs, to support the analysis or investigation. The information exchange about vulnerabilities and attack patterns between CERT and CIRT, assuming both internally at a national level, and externally between other national CERTs may bring additional knowledge and value about the re-occurring similar incidents somewhere else. This type of information sharing may create confidence for the target in terms of confirming similar exploitation techniques and mechanisms.

Secondly, national CERTs do not hold political mandate of tackling misperception and misunderstanding on the level of decision makers. Assuming that the national CERT is willing to resolve the misunderstanding, would this mean contacting the decision makers of the suspected state directly to clarify the actions?

How organizing exercises between CERTs will help to clear misperception of being under cyber attack when the network is being penetrated? Exercises are beneficial in training and improving technical skills of the staff working in CERTs, assuming that the relevant CERTs/CIRTs (as government and military) would also be involved. If the exercise concentrates on technical aspects of cyber incidents, it provides knowledge and improves technical skills, that can be used for identifying act of penetration. It may help technical teams to create the chronology of the incident. This can be empowering the teams in their technical and analytical skills, encouraging digital forensics practices. These exercises may help the teams to look for artifacts, analyse logs, identify what external software have been installed or any user accounts created on the systems. This “exercises” for looking for evidence maybe be helpful, but also biased. It can be biased in a way we look for evidence, and how we construct our case. The evidence may lead or point some behavior, but the evidence of actions will not verify or confirm the intentions.

Conclusion

The context from where the confidence-building measure originate shapes their characteristics and empowers their effectiveness. Their goal is to "contribute to, reduce or, in some instances, even eliminate the causes for mistrust, fear, tensions and hostilities, (…) regions and, ultimately, also on a world-wide scale" [7]. The fear and the mistrust can be built upon previous experiences between states, it can have factors as "geographical, strategic, political, economic, social and other elements" (page 4 [7]). And in some cases regional or geographical aspects don't play a significant role as "there may also be a lack of confidence among States which are not neighbours" [7].

During the past decade there has been discussions and concrete propositions by the UN GGE on Developments in the Field of Information and Telecommunications in the Context on International Security to use confidence-building measures for the ICT-related incidents to reduce the risk of conflict.

To discover the effect that the proposed CBMs is expected to offer, selected CBM by the author (page 10 (d) [10]), the author has analysed the characteristics of the CBMs defined by the theory of CBMs, and identified seven characteristics that the author used as a blueprint. The textual interpretation of the proposed CBMs didn't allow to accommodate fully correspondence to the formal criteria and requirements defined for the effectiveness of the CBMs. Two characteristics out of seven, the *Specificity* and the *Systemic approach*, were met based the textual interpretation. Further, negation was applied to the proposed CBM, that did not rule out the set requirements. Additionally, the teleological argumentation revealed more of the potential effect that the proposed CBMs.

The first part of analysis based on the characteristics of the CBMs, revealed that the ecosystem defined from the theory of the CBMs is not established. The specific proposed CBM by the UN GGE (page 10 (d) [10]) is not a measure itself but a proposition for a measure. The Second part of the analysis revealed that cyber espionage is a prevalent cyber incident affecting states. This conclusion was possible based on the analysis of six publicly available datasets of cyber incidents. Based on the outcome of the dataset analysis, the author created the ontology of cyber espionage, offering a methodological approach how to address further the application of the proposed CBM.

The methodology that the author suggested for the evaluation of the effectiveness of the proposed CBMs, helps to identify pieces and elements, their possible relationship and interaction. By using the offered methodology, the author was able to identify possible perceptions and misperceptions that the victim may obtain from the actions of the stages of the cyber espionage. Further, the methodology of cross-examination of the nine elements of the cyber espionage (from the ontology of the cyber espionage) and the five elements of the proposed CBM allowed to discuss, unveil what effectiveness the proposed CBM may have.

This suggested methodology for evaluating the effectiveness of the proposed CBM can be used to conduct further analysis of other CBMs that are being proposed for the ICT environment.

Additionally, the author highlights the importance of the context that the CBMs originally come from. The effectiveness of these measures is very much related to the context where these measures are introduced to. The study of confidence-building measures offered "the dissemination and exchange of pertinent information, regular personal contacts at all levels of political and military decision-making should be encouraged (…) to foster co-operation in the field of security-related communication" [7]. This suggestion speaks for itself and points towards the "Decision Makers", as a "target" and "source" for elimination of mistrust and fear. Essentially, the national CERT cooperation might add confidence in creating basis for evidence, but this will not help in clearing the perception or misperception or verifying the motives. Secondly, the exchange of information is done on the level of CERTs, which implies that there is no direct contact with decision makers or reaching out to suspected state sponsor of the cyber espionage for clarification.

To conclude, the elimination of mistrust aims not only in creating political healthy climate, but also "psychological climate (…) in which the importance of the military element will be gradually diminished and finally eliminated" [7]. The way we transfer our fears and how we perceive the sense of insecurity defines our action, more precisely our reactions.

One example back from 1983, that relates to "reacting" based on the situational awareness is about the Soviet Union's early-warning system for detecting incoming missile strikes. The duty officer Stanislav Petrov mentioned during his interview "there was no rule about how long we were allowed to think before we reported a strike" [73], but he breached his instructions and saved the world.

References

[1] The United Nations, “A/RES/33/91. General and Complete Disarmament,” in *United Nations, General Assembly*, 1978.

[2] The United Nations General Assembly, “A/RES/34/87. General and complete disarmament,” 1979.

[3] “Helsinki Final Act | OSCE,” 1975. [Online]. Available: https://www.osce.org/helsinki-final-act. [Accessed: 08-Mar-2019].

[4] C. on Security and and C. in E. (CSCE), “Document of the Stockholm Conference,” 1986.

[5] Organization for Security and Co-operation in Europe, “Vienna Document 1990 | OSCE,” 1990. [Online]. Available: https://www.osce.org/fsc/41245. [Accessed: 09-Mar-2019].

[6] U. D. of State, “Confidence- and Security-Building Measures.” [Online]. Available: https://www.state.gov/t/isn/4725.htm. [Accessed: 08-Mar-2019].

[7] “Disarmament Study Series: No. 7 – UNODA,” New York, 1982.

[8] G. A. United Nations, “Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security,” 2010.

[9] G. A. United Nations, “Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security,” 2013.

[10] G. A. United Nations, “Report of the Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security,” 2015.

[11] M. C. J. R. Michael Krepon, Dominique M.McCoy, Ed., *A Handbook on Confidence-Building Measures for Regional Security*, Handbook N. Washington, DC: The Henry L.Stimson Center.

[12] D. K. Ziolkowski, “Confidence Building Measures for Cyberspace – Legal Implications,” 2013. [Online]. Available: https://ccdcoe.org/uploads/2018/10/CBMs.pdf. [Accessed: 08-Mar-2019].

[13] Erica D. Borghard and Shawn W. Lonergan, “Confidence Building Measures for the Cyber Domain,” *Strateg. Stud. Q.*, no. Fall 2018, 2018.

[14] B. Baseley-Walker and J. A. Lewis, “Confronting cyberconflict,” 2011.

[15] J. Healey, J. C. Mallery, K. T. Jordan, and N. V. Youd, “CONFIDENCE-BUILDING MEASURES IN CYBERSPACE: A MULTISTAKEHOLDER APPROACH FOR STABILITY AND SECURITY,” Stockholm, 2014.

[16] T. Yamin, “Developing Information-Space Confidence Building Measures (CBMs) between India and Pakistan,” 2014.

[17] “Agreement Between the Government of The United States of America and the Government of The Union of Soviet Socialist Republics on the Prevention of Incidents On and Over the High Seas.” [Online]. Available: https://www.state.gov/t/isn/4791.htm. [Accessed: 05-May-2019].

[18] M. Kerttunen, “Tikk_Kerttunen Parabasis Cyber-diplomacy in stalemate.pdf,” *Parabasis: Cyber-diplomacy in stalemate*.

[19] T. U. D. Commission, “The Guidelines for Confidence-building Measures,” 1988.

[20] M. H. D. S. Dr. John A. Battilega, Dr. Mark T. Clark, Mr. Charles H. Fairbanks, Jr., Mr. Tod Lindberg, Dr. Richard R. Muller, Dr. James Mulvenon, Lietenant

General William E. Odom, Mr. Michael Quinlan, Mr. Henry S. Rowen, Dr. Harvey M. Sapolsky, *GETTING MAD: NUCLEAR MUTUAL ASSURED DESTRUCTION, ITS ORIGINS AND PRACTICE*. Nonproliferation Policy Education Center (NPEC), Army War College's Strategic Studies Institute (SSI), 2004.

[21] Organization for Security and Co-operation in Europe, "Arms control | OSCE." [Online]. Available: https://www.osce.org/arms-control. [Accessed: 09-Mar-2019].

[22] G. A. United Nations, "RES/64/211 - Creation of a global culture of cybersecurity and taking stock of national efforts to protect critical information infrastructures," 2009. [Online]. Available: https://undocs.org/A/RES/64/211. [Accessed: 17-Mar-2019].

[23] "About CFR | Council on Foreign Relations." [Online]. Available: https://www.cfr.org/about.

[24] "Our Methodology | Council on Foreign Relations," 2015. .

[25] A. J. Lewis and D. E. Zheng, "Significant Cyber Incidents | Center for Strategic and International Studies." [Online]. Available: https://www.csis.org/programs/cybersecurity-and-governance/technology-policy-program/other-projects-cybersecurity.

[26] Kaspersky Lab, "Targeted Cyberattacks Logbook," 2017. [Online]. Available: https://apt.securelist.com/#!/threats/.

[27] P. Passeri, "Cyber Attacks Statistics – Hackmageddon," 2017. [Online]. Available: https://www.hackmageddon.com/category/security/cyber-attacks-statistics/.

[28] R. Valeriano, Brandon and Maness, "Coding Cyber Security Incident Data." [Online]. Available: http://relationsinternational.com/coding-cyber-security-incident-data/.

[29] B. Valeriano and R. C. Maness, *Cyber war versus cyber realities : cyber conflict in the international system*. Oxford University Press, 2015.

[30] Council on Foreign Relations, "Cyber Operations Tracker." [Online]. Available: https://www.cfr.org/interactive/cyber-operations#OurMethodology. [Accessed: 14-Mar-2019].

[31] US departement of Justice, "Indictment of officials from the Mabna Institute | Council on Foreign Relations Interactives," 2018. [Online]. Available: https://www.cfr.org/interactive/cyber-operations/indictment-officials-mabna-institute.

[32] US departement of Justice, "Deputy Attorney General Rosenstein Delivers Remarks Announcing Nine Iranians Charged with Conducting Massive Cyber Theft Campaign on Behalf of the Islamic Revolutionary Guard Corps | OPA | Department of Justice," *Justice News*, 2018. [Online]. Available: https://www.justice.gov/opa/pr/nine-iranians-charged-conducting-massive-cyber-theft-campaign-behalf-islamic-revolutionary.

[33] "CSIS Technology Policy | Significant Cyber Incidents." [Online]. Available: https://csis-ilab.github.io/js-viz/tech-policy/cyber-incidents-bar/index.html.

[34] M. Schreier, *Qualitative Content Analysis in Practice.* SAGE Publications, 2012.

[35] FireEye, "Anatomy of an APT (Advanced Persistent Threat) Attack | FireEye," 2017. [Online]. Available: https://www.fireeye.com/current-threats/anatomy-of-a-cyber-attack.html.

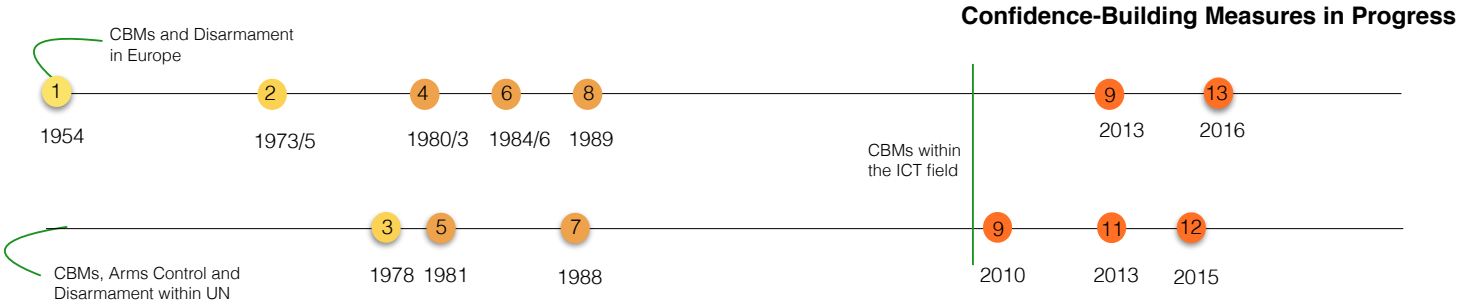
[36] N. Lord, "What is an Advanced Persistent Threat? APT Definition | Digital Guardian," *Digital Guardian*, 2016. [Online]. Available:

https://www.cisco.com/c/en/us/products/security/advanced-persistent-threat.html.
[37] J. A. Guerrero-Saade, R. (GReAT) Costin, D. Moore, and T. (King's C. L. Rid, "Penquin's Moonlit Maze | Securelist." [Online]. Available: https://securelist.com/penquins-moonlit-maze/77883/.
[38] The GReAT, "The art of finding Cyber-Dinosaur skeletons," 2014. [Online]. Available: https://securelist.com/the-art-of-finding-cyber-dinosaur-skeletons/67928/.
[39] A. V. Zagorskii and N. P. Romashkina, *Information Security Threats during Crisis and Conflicts of the Conflicts of the XXI Century*. Primakov Institute of World Economy and International Relations, Russian Academy of Sciences (IMEMO), 23, Profsoyuznaya Str., Moscow, 117997, Russian Federation, 2016.
[40] P. Passeri, "Hackmegeddon Information Security Timelines and Statistics," 2017. [Online]. Available: https://www.hackmageddon.com/.
[41] UNIDO, "International Standard Industrial Classification of All Economic Activities (ISIC)," 2018. [Online]. Available: https://stat.unido.org/content/learning-center/international-standard-industrial-classification-of-all-economic-activities-%2528isic%2529. [Accessed: 09-Mar-2019].
[42] B. J. Ryan C. Maness, Brandon Valeriano, "CYBER CONFLICT DATASET." [Online]. Available: https://drryanmaness.wixsite.com/cyberconflcit/cyber-conflict-dataset. [Accessed: 14-Mar-2019].
[43] "sabotage | Definition of sabotage in English by Oxford Dictionaries." [Online]. Available: https://en.oxforddictionaries.com/definition/sabotage. [Accessed: 23-Mar-2019].
[44] European Union Agency For Network and Information Security, "ENISA overview of cybersecurity and related terminology," 2017. [Online]. Available: https://www.enisa.europa.eu/publications/enisa-position-papers-and-opinions/enisa-overview-of-cybersecurity-and-related-terminology. [Accessed: 23-Mar-2019].
[45] B. J. Ryan C. Maness, Brandon Valeriano, "Codebook for the Dyadic Cyber Incidentand Dispute Dataset Version 1.1," 2017.
[46] Council on Foreign Relations, "Cyber Operations Tracker." [Online]. Available: https://www.cfr.org/interactive/cyber-operations#Glossary. [Accessed: 24-Mar-2019].
[47] Oxford Dictionary | Oxford University Press, "dox | Definition of dox in English by Oxford Dictionaries." [Online]. Available: https://en.oxforddictionaries.com/definition/dox. [Accessed: 24-Mar-2019].
[48] US-CERT | Department of Homeland Security, "Understanding Denial-of-Service Attacks | US-CERT." [Online]. Available: https://www.us-cert.gov/ncas/tips/ST04-015. [Accessed: 24-Mar-2019].
[49] "Website Security | US-CERT." [Online]. Available: https://www.us-cert.gov/ncas/tips/ST18-006. [Accessed: 25-Mar-2019].
[50] "deface | Definition of deface in English by Oxford Dictionaries." [Online]. Available: https://en.oxforddictionaries.com/definition/deface. [Accessed: 25-Mar-2019].
[51] "What is data destruction? - Definition from WhatIs.com." [Online]. Available: https://searchstorage.techtarget.com/definition/data-destruction. [Accessed: 25-Mar-2019].
[52] Oxford Dictionary | Oxford University Press, "espionage | Definition of espionage in English by Oxford Dictionaries." [Online]. Available:

https://en.oxforddictionaries.com/definition/espionage. [Accessed: 23-Mar-2019].
[53] “Compromise of SingHealth, a large health-care provider in Singapore | Council on Foreign Relations Interactives.” [Online]. Available: https://www.cfr.org/interactive/cyber-operations/compromise-singhealth-large-health-care-provider-singapore. [Accessed: 19-Apr-2019].
[54] “Singapore disconnects healthcare computers from the Internet after cyber attack - Reuters.” [Online]. Available: https://www.reuters.com/article/us-singapore-cyberattack/singapore-disconnects-healthcare-computers-from-the-internet-after-cyber-attack-idUSKBN1KE15X. [Accessed: 19-Apr-2019].
[55] “Chrysene | Council on Foreign Relations Interactives.” [Online]. Available: https://www.cfr.org/interactive/cyber-operations/chrysene. [Accessed: 19-Apr-2019].
[56] T. I. I. Dragos, “Chrysene | Dragos,” 2018. [Online]. Available: https://dragos.com/resource/chrysene/. [Accessed: 19-Apr-2019].
[57] “Compromise of a U.S. Navy contractor | Council on Foreign Relations Interactives.” [Online]. Available: https://www.cfr.org/interactive/cyber-operations/compromise-us-navy-contractor. [Accessed: 19-Apr-2019].
[58] “China hacked a Navy contractor and secured a trove of highly sensitive data on submarine warfare - The Washington Post.” [Online]. Available: https://www.washingtonpost.com/world/national-security/china-hacked-a-navy-contractor-and-secured-a-trove-of-highly-sensitive-data-on-submarine-warfare/2018/06/08/6cc396fa-68e6-11e8-bea7-c8eb28bc52b1_story.html?utm_term=.6cb651c1f812. [Accessed: 19-Apr-2019].
[59] “Compromise of the Finnish Ministry of Foreign Affairs | Council on Foreign Relations Interactives.” [Online]. Available: https://www.cfr.org/interactive/cyber-operations/compromise-finnish-ministry-foreign-affairs. [Accessed: 18-Apr-2019].
[60] “MTV3: Suomen ulkoministeriö laajan verkkovakoilun kohteena vuosia - MTVuutiset.fi.” [Online]. Available: https://www.mtvuutiset.fi/artikkeli/mtv3-suomen-ulkoministerio-laajan-verkkovakoilun-kohteena-vuosia/2369718. [Accessed: 18-Apr-2019].
[61] “Finland’s Foreign Ministry gets pwned by worse-than-Red October malware | Ars Technica.” [Online]. Available: https://arstechnica.com/tech-policy/2013/10/finlands-foreign-ministry-gets-pwned-by-red-october-malware/. [Accessed: 18-Apr-2019].
[62] “Compromise of the German Foreign Office | Council on Foreign Relations Interactives.” [Online]. Available: https://www.cfr.org/interactive/cyber-operations/compromise-german-foreign-office. [Accessed: 18-Apr-2019].
[63] C. Falk, “An Ontology for Threat Intelligence,” 2016.
[64] “FIRST Teams.” [Online]. Available: https://www.first.org/members/teams/. [Accessed: 31-Mar-2019].
[65] “Updated Map (v2.5) of ’Digital Fire-brigades’- CERTs — ENISA.” [Online]. Available: https://www.enisa.europa.eu/news/enisa-news/updated-map-of-digital-firebrigade-certs. [Accessed: 28-Mar-2019].
[66] European Union Agnecy for Network and Information Security, “CSIRTs by Country - Interactive Map — ENISA.” [Online]. Available: https://www.enisa.europa.eu/topics/csirts-in-europe/csirt-inventory/certs-by-country-interactive-map#. [Accessed: 28-Mar-2019].
[67] “Regular Synonyms, Regular Antonyms | Merriam-Webster Thesaurus.” [Online]. Available: https://www.merriam-webster.com/thesaurus/regular.

[Accessed: 11-Apr-2019].
[68] “Intermittence | Definition of Intermittence by Merriam-Webster.” [Online]. Available: https://www.merriam-webster.com/dictionary/intermittence. [Accessed: 12-Apr-2019].
[69] L. Martin, “GAINING THE ADVANTAGE | Applying Cyber Kill Chain® Methodology to Network Defense,” 2015. .
[70] R. Jervis, *Perception and misperception in international politics*. .
[71] M. C. Libicki, “Drawing Inferences from Cyber Espionage,” in *10th International Conference on Cyber Conflict*, 2018.
[72] “RFC 2350 - CERT.at.” [Online]. Available: https://www.cert.at/about/rfc2350/rfc2350.html. [Accessed: 12-May-2019].
[73] “Stanislav Petrov: The man who may have saved the world - BBC News.” [Online]. Available: https://www.bbc.com/news/world-europe-24280831. [Accessed: 12-May-2019].

Annex 1 Timeline for Confidence-Building Measures



1 Discussions on all-European Treaty on collective security with the goal of ensuring the status quo in Europe and establishing a collective European security agreement, with the United States having "observer" status. [1]

2 Conference on Security and Cooperation in Europe (CSCE) took place in Helsinki in July 1973, followed by the conference in Geneva in 1975. The Helsinki Final Act signed (1975, August), that had three main subjects ("Basket"). The Basket One concerns questions of security in Europe, including principles guiding relations among participating states and confidence-building measures. [1]

4 Follow-up meetings took place in Madrid, between November 1980 to September 1983. [1]

6 Stockholm Document, with principle measures signed (1986). The Stockholm Conference on Confidence- and Security-Building Measures and Disarmament in Europe. [1]

8 The Vienna Concluding Document, on resuming the negotiations on CSBMs. [1]

9 2013 12 OSCE Decision No 1106 Initial Set of OSCE Confidence-Building Measures to Reduce the Risks of Conflict Stemming from the Use of ICT.

13 2016 03 OSCE Decision No. 1202 OSCE Confidence-building Measures to Reduce the Risks of Conflict Stemming from the Use of ICT

3 United Nations General Assembly resolution 33/91 B of 16 December in 1978.

5 1981 08 UN Department of Political and Security Council Affairs United Nations Centre for Disarmament Comprehensive Study on Confidence-Building Measures (A-36-474)

7 Review of the implementation of the recommendations and decisions adopted by the General Assembly at its tenth special session (A-RES-43-78)

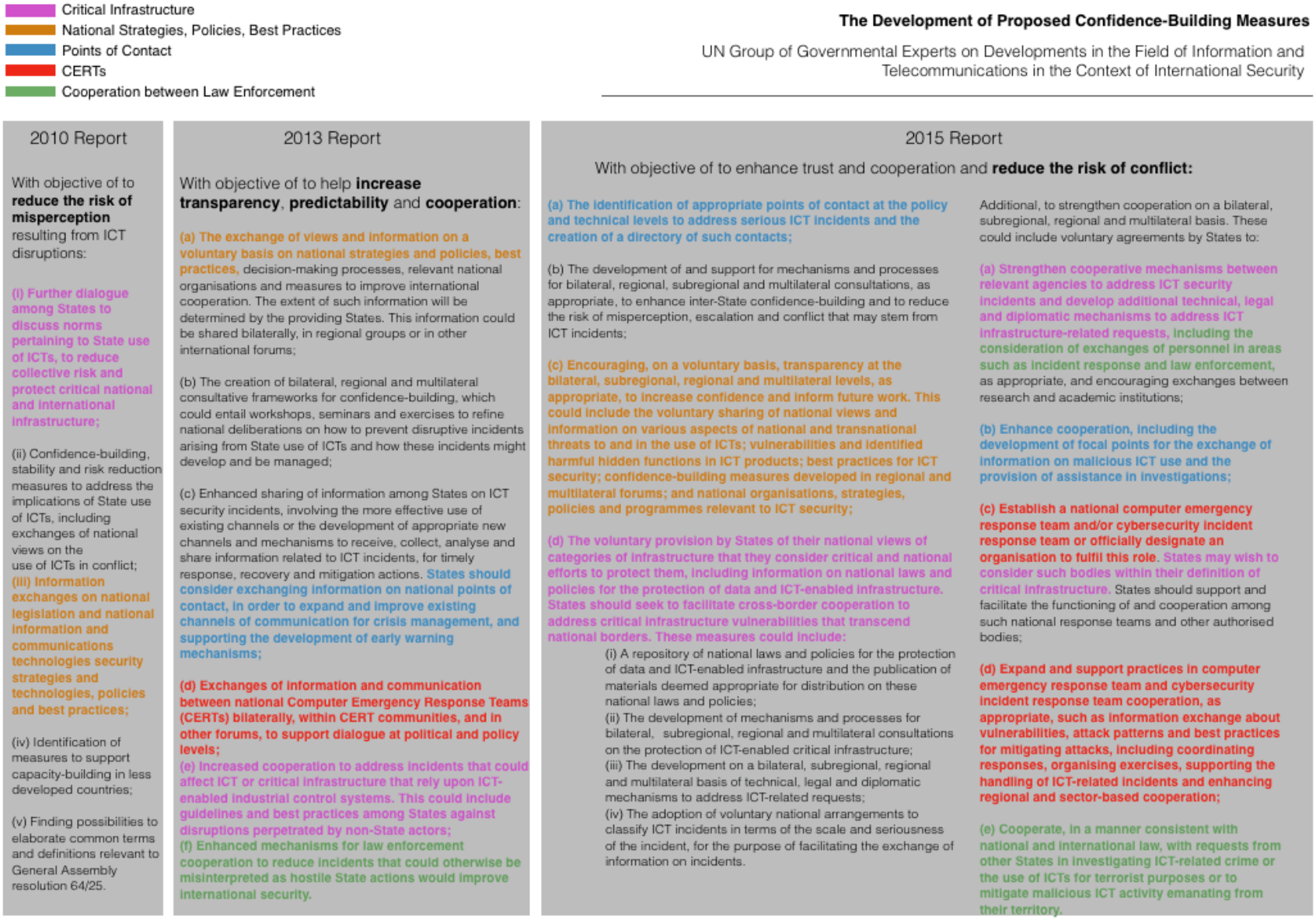
9 2010 07 UN Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security (A-65-201).

11 2013 06 UN Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security (A-68-98*).

12 2015 09 UN Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security (A-70-174)

[1] The US Department of State, Diplomacy in Action. Confidence-Building Measures. Online Source: https://www.state.gov/t/isn/4725.htm

Annex 2 Thematic representation of the content | UN GGE 2010, 2013 and 2015 Reports



Annex 3 ENISA CERT/CSIRT Dataset | Constituency Analysis