ICRC EXPERT MEETING
14–16 NOVEMBER 2018 – GENEVA

THE POTENTIAL HUMAN COST OF CYBER OPERATIONS

REPORT

ICRC

ICRC EXPERT MEETING
14–16 NOVEMBER 2018 – GENEVA

THE POTENTIAL HUMAN COST OF CYBER OPERATIONS

Report prepared and edited by Laurent Gisel, senior legal adviser, and Lukasz Olejnik, scientific adviser on cyber, ICRC

Table of Contents

Foreword

One of the main aims of international humanitarian law (IHL) is to protect the civilian population from the effects of military operations. Cyber warfare is the subject of growing concern, and there is no consensus around the question of how IHL will protect civilians against its effects.

But what *are* the effects of cyber warfare on civilians? Since most known operations have been conducted outside conflict settings, the potential human cost of cyber operations in armed conflict is a matter of risk analysis.

To move towards a realistic assessment of the potential human cost of cyber warfare, the International Committee of the Red Cross (ICRC) invited scientific and cyber security experts from all over the world to share their knowledge. In a three-day meeting, experts analysed some of the most sophisticated known cyber operations, regardless of whether they occurred during conflict or in peacetime, focusing on the risk that cyber operations may result in death, injury or physical damage, affect the delivery of essential services to the population, or affect core internet services.

The meeting included participants working for global IT companies, cyber threat intelligence companies, computer emergency response teams, a national cyber security agency, participants with expertise in cyber security (including that of hospitals, electricity grids and other services), participants with expertise in the development and use of military cyber operations, lawyers and academics.

The rich discussions provided a nuanced picture of the risks that cyber warfare can entail for the civilian population. One of the main fears of those working on cyber warfare and IHL is perhaps the idea that in cyber space, the principle of distinction will be difficult if not impossible to uphold. Yet, the expert meeting showed that the global digital infrastructure that can be targeted through cyber operations is in fact rather resilient to widespread effects. While a number of the cyber attacks analysed were indiscriminate, many others have been precisely targeted from a technical perspective. Nonetheless, while many systems are resilient, others are particularly vulnerable, and health-care systems are among those. Furthermore, the threats are evolving at a faster pace than anticipated, and the most sophisticated cyber capabilities may be largely unknown.

Another area of concern highlighted in the meeting is the risk of proliferation of cyber tools, because they may linger in digital systems and can potentially be accessed from anywhere in the world, modified and reused.

In the view of the ICRC, many of the operations described in the report would be contrary to IHL if carried out during armed conflict. However, there is insufficient consensus today as to the interpretation of IHL in cyber space to provide clear legal protection for the civilian population.

We are grateful to the experts for having shared their deep knowledge and expertise. With this report, we hope to help develop a realistic picture of the risks to civilians that can arise from cyber warfare and to highlight the need to address those risks on several levels: through cyber security measures, but also through clarity and agreement about IHL as the most important international legal framework for the protection of civilians in armed conflict.

Cordula Droege
Chief Legal Officer and Head of the Legal Division, ICRC

Acknowledgements

The present report is the outcome of an expert meeting held in November 2018 on the potential human cost of cyber operations.

The conceptualization, drafting and publication of the report would not have been possible without the commitment and contributions of many people.

Our gratitude goes, first of all, to the experts who participated in a personal capacity and whose expertise, knowledge and contributions were essential to the success of this meeting. They provided invaluable expertise on the various types of cyber operations, the threats they may raise for civilians and for the delivery of essential services for the population, and the feasibility of some avenues that could be explored in an effort to prevent or reduce the human cost of cyber operations.

We would also like to express our gratitude to Laurent Gisel, senior legal adviser, and Lukasz Olejnik, scientific adviser on cyber, who were in charge of organizing the meeting, drafting the background document and preparing this report. We would like to extend our thanks to Tilman Rodenhauser, legal adviser, and Sophie Huve and Guillem Adrià Puri Plana, associates, who assisted in the preparation of this report, as well as Christopher Scala, the editor who reviewed it.

Finally, we would like to sincerely thank all our other colleagues at the ICRC who commented on the draft report – in particular Lindsey Cameron, head of the Thematic Legal Advice Unit in the Legal Division – provided valuable support in organizing and following up on the expert meeting, or helped publish this report.

Cordula Droege
Chief Legal Officer and Head of the Legal Division, ICRC

Executive summary

Cyber operations during armed conflicts: assessing the challenges for international humanitarian law

The use of cyber operations during armed conflicts is a reality. While only a few States so far have publicly acknowledged that they use them, cyber operations are a known feature of present-day military operations and the use of them is likely to increase in the future.

This new reality has triggered a debate regarding the rules of international law that apply to such operations. In this debate, the ICRC has recalled that during armed conflict, cyber operations are subject to the rules of IHL.[1] It is nevertheless clear that cyberspace and these new military operations raise a number of questions as to precisely how certain rules of IHL – which were drafted primarily with the kinetic realm in mind – apply to cyber operations.

Assessing these questions requires an understanding of the expected use and military potential of cyber technology. What aims may belligerents want to achieve by using new tools at the strategic, operational or tactical levels during conflicts? How does this new technology compare to other, existing means of warfare?

Furthermore, to assess how IHL protects civilians in armed conflict, and whether further regulation is needed, lawyers and policy makers require an understanding of the actual or potential human cost of cyber technologies. Indeed, one of the main aims of IHL is to protect civilians from the effects of military operations.

Purpose and scope of the meeting

As part of its mandate to work for the clarification of IHL and, if necessary, prepare any development thereof, the ICRC monitors the development of new technologies that are, or could be, used as means and methods of warfare during armed conflicts. This approach is based on legal, technical, military and humanitarian considerations, which are interrelated.

To develop a realistic assessment of cyber capabilities and their potential humanitarian consequences in light of their technical characteristics, the ICRC brought together scientific and cyber security experts from all over the world to share their knowledge about the technical possibilities, expected use, and potential effects of cyber operations. The three-day meeting drew on the expertise of participants working for global IT companies, cyber threat intelligence companies, computer emergency response teams, a national cyber security agency, participants with expertise in cyber security (including that of hospitals, electrical grids and other services), participants with expertise in the development and use of military cyber operations, lawyers and academics.

States and militaries remain reluctant to disclose their cyber capabilities, including the details of cyber operations conducted in the context of armed conflicts, and little is known about the few acknowledged cases. Therefore, the experts discussed a number of the most sophisticated known cyber operations, regardless of whether they occurred in the context of an armed conflict or in peacetime. Examining the technical features of these attacks and the specific vulnerabilities of the respective targets provides a powerful evidence base for what is technically possible also during armed conflict.

The meeting focused in particular on the risk that cyber operations might cause death, injury or physical damage, affect the delivery of essential services to the population, or affect the reliability of internet services. It looked at the specific characteristics of cyber tools, how cyber threats have evolved, and the cyber security landscape.

Approaching the subject from a humanitarian law and humanitarian action perspective, the ICRC seeks a sober and – to the greatest extent possible – evidence-based understanding of the risks of cyber

[1] See in particular: ICRC, *International Humanitarian Law and the challenges of contemporary armed conflicts*, ICRC, Geneva, 2015, pp. 39–44 (hereinafter ICRC 2015 IHL Challenges report) (all web addresses accessed April 2019). The restrictions imposed by IHL do not legitimize the use of force in cyber space, which remains governed by the United Nations Charter.

attacks[2] for the civilian population. The meeting allowed the ICRC to confirm much of its own research (submitted in the background paper, included as Annex 3), and to supplement it with highly valuable additional expert knowledge. The meeting was extremely useful in that it contributed to a nuanced picture of cyber operations, demystifying some of the assumptions that often surround discussions on cyber warfare.

Areas of concern

Discussions helped to put the spotlight on four areas of particular concern in terms of the potential human cost of cyber operations:

a) the specific vulnerabilities of certain types of infrastructure
b) the risk of overreaction due to potential misunderstanding of the intended purpose of hostile cyber operations
c) the unique manner in which cyber tools may proliferate
d) the obstacles that the difficulty of attributing cyber attacks creates for ensuring compliance with international law.

a) Specific vulnerabilities of certain types of infrastructure: cyber attacks that may affect the delivery of health care, industrial control systems, or the reliability or availability of core internet services

Apart from causing substantial economic loss, cyber operations can harm infrastructure in at least two ways. First, they can affect the delivery of essential services to civilians, as has been shown with cyber attacks against electrical grids and the health-care sector. Second, they can cause physical damage, as was the case with the Stuxnet attack against a nuclear enrichment facility in Iran in 2010, and an attack on a German steel mill in 2014.

Cyber attacks that may affect the delivery of health care

The health-care sector is moving towards increased digitization and interconnectivity. For example, hospital medical devices are normally connected to the hospital's information technology (IT) system to enable automatic electronic filing. Connected biomedical devices, such as pacemakers and insulin pumps, make it possible to remotely monitor individual patients' health as well as the functioning of the medical devices themselves.

This increased digital dependency, combined with an increased 'attack surface', has not been matched by a corresponding improvement in cybersecurity. Consequently, this infrastructure is particularly vulnerable, with potentially serious consequences for health and life.

Cyber attacks against industrial control systems, including those used in critical civilian infrastructure

Industrial control systems are protected by complex safety mechanisms and often have built-in redundancy to guarantee safety and reliability. For example, electrical networks are grids with multiple power sources to avoid widespread effects when one of their parts is affected. Nonetheless, attacks on specific nodes might still cause a significant impact, such as if a critical system (like a hospital) depends on a specific sub-system or node, or because they have cascading harmful consequences.

Carrying out a cyber attack against an industrial control system requires a certain expertise and sophistication, and, often, custom-made malware. Such attacks have been less frequent so far than other types of cyber operations. Nonetheless, their frequency is reportedly increasing, and the severity of the threat has evolved more rapidly than anticipated only a few years ago. There is a risk that tools developed by the best-resourced actors may be repurposed or purchased by other actors who lack the expertise required to develop them from scratch. Moreover, there is a possibility that a number of undetected actors are capable of attacking industrial control systems.

Cyber attacks that may affect the reliability or availability of internet services

Cyber attacks that disrupt core internet services, such as the domain name system (DNS), which supports communications on the internet, or disrupt the functioning of major cloud services, may impact all services that rely on them. However, the risk of seriously compromising these core internet services was assessed by the experts as unlikely at the present moment thanks to the high degree of redundancy in the DNS and because major cloud providers tend to offer high security standards. If,

[2] The terms "cyber attacks" and "cyber operations" are used throughout the report in a technical (mainstream or colloquial) sense and not as they may be understood under international humanitarian law (IHL), unless specifically stated (see the first paragraph of Session 1 below for more details).

however, such disruption were to occur, it could have widespread and potentially serious consequences, for example when life-saving services such as ambulances rely on the cloud.

Finally, "distributed denial of service" (DDoS) attacks have been used against services provided by governments for the population. Such attacks are carried out through increasingly large botnets. The arrival of the internet of things will further increase the number of connected devices that could be used in such attacks. Furthermore, DDoS attacks might have a wider impact than expected by their author, in particular when information about the targeted network is incomplete.

b) Risk of overreaction due to the potential misunderstanding of the intended purpose of hostile cyber operations

Cyber operations can be broadly divided into two categories, depending on their purpose:

- activity encompassing reconnaissance, surveillance and the exfiltration of data and information, for example for espionage, often referred to as computer network exploitation (CNE), or "access operations"
- activity to generate effects on a targeted system or device, such as tampering with data integrity (deletion, modification), affecting availability (disabling, including for prolonged periods of time), or causing physical effects, such as damaging the system, often referred to as a computer network attack (CNA), or "effects operations".

The distinction is primarily one of purpose. From a technical perspective, the initial steps of a CNE and a CNA to gain and maintain persistent access to the target may be identical. CNEs can then be turned into CNAs relatively simply, mostly through the use of specific payloads of a different nature. While the initial steps of the attacks may be tracked, it is often difficult to fully assess the attacker's purpose until the effect on the end target is actually achieved.

When the target does not know the actual purpose of the operation, its reaction may be to consider the potential worst-case impact that the attacker could achieve though a CNA and react in a stronger manner than it would have if it had known that the intended purpose of the attack was CNE. This escalation risk factor may give rise to a potentially harmful over-reaction.

c) Proliferation of cyber tools

A third concern is the proliferation of cyber tools – an issue that in some respects raises concerns similar to those that may exist with regard to weapons proliferation or the proliferation of dual-use technology, although the specific nature of cyber tools must be taken into account.

Cyber tools and methods can proliferate in a unique manner that is difficult to control. First, cyber space is a global domain: provided that the attacker can overcome the cyber security and defence measures in place, any network node and information residing on the network can be accessed from anywhere in the world. At the same time, cyber tools can be repurposed or reengineered. The combination of these two characteristics means that when cyber tools have been used, stolen, leaked or otherwise become available, actors other than those who developed them might be able to find them, reverse engineer them, and reuse them for their own purposes.

Finally, the fact that cyber tools and methods can be repurposed and reused is one of the factors making rapid and reliable technical attribution of cyber attacks a challenging process.

d) Attribution of attacks

While not a primary focus of the meeting, the discussions also touched upon the anonymity of attacks and the difficulty to attribute them to a specific actor, which is a fourth area of concern.

Cyber space is a complex domain where multiple actors operate: individual hackers; criminal groups, potentially motivated by financial gain; States; non-State armed groups; and other non-State actors. Actors may also cooperate: for example, States may buy cyber tools or have an operation performed on their behalf against a target they have identified.

Digital forensics and the capabilities of attribution of malicious cyber activity appear to be improving. Nonetheless, the ability of threat actors to obscure or effectively hide the origin of their operations on the internet, compounded by the ability to buy, repurpose or reengineer cyber tools developed or used by other actors continues to make it difficult to rapidly and reliably attribute cyber attacks to a specific actor. This hampers the possibility to identify actors who violate IHL in cyberspace and hold them responsible. This is a concern because to hold such actors responsible is one way to ensure compliance with IHL. It may also lower the threshold of using cyber attacks and of using them in violation of international law, because attackers can deny responsibility.

Cyber operations during armed conflicts: implications for international humanitarian law

It is well-established that international law applies to cyber operations. More specifically, IHL and its principles of distinction, proportionality, precaution, military necessity and humanity restrict the use of cyber means and methods during armed conflict. Further discussions may however be needed to clarify how IHL applies and whether it is adequate and sufficient or requires further development, building on existing law.

The meeting helped to clarify which areas of humanitarian concern should be the focus of attention. In brief, based on the detailed knowledge of cyber operations during peacetime, and somewhat lesser knowledge of cyber operations in times of armed conflict, the following picture emerges:

Distinction in cyber space

First, cyber attacks are not necessarily indiscriminate. As the report illustrates in more detail, cyber tools can be designed to self-propagate or not. Even if they self-propagate and cause cyber security concerns for all those infected, they can be designed to only cause damage to a specific target. While some self-propagating malware that caused indiscriminate harmful effects has made headlines, many cyber operations have in fact been rather discriminate from a technical perspective (which does not mean they were lawful).

Furthermore, certain types of cyber attacks require custom-made cyber tools, such as those that would aim to cause physical damage to industrial control systems. In many cases this would also effectively hamper the ability to carry them out in a large-scale indiscriminate manner.

This is important from an IHL perspective, because contrary to the assumption often heard that the principle of distinction might have become meaningless in cyber space because of the interconnectivity that characterizes it, not all offensive cyber tools are inherently indiscriminate. On the contrary, they may well be very precisely tailored and create effects on specific targets only.

Highlighting the potential human cost

Secondly, and of equal importance, it is nonetheless clear that cyber tools can cause substantial damage and can be – and have sometimes been – indiscriminate, and that certain systems are particularly at risk, first and foremost, perhaps, health-care systems. Moreover, the threats that can be observed have been evolving faster than anticipated, in particular regarding attacks against industrial systems. Finally, much is still unknown in terms of the rapid evolution of the technology, the capabilities and the tools developed by the most sophisticated actors, and the extent to which the increased use of cyber operations during armed conflicts might be different from the trends observed so far. In other words, while the risk of human cost based on current observations does not appear extremely high, especially considering the destruction and suffering that conflicts always cause, the evolution of cyber operations still merits close attention due to existing uncertainties and the rapid pace of change.

Legal protection through IHL

Many of the attacks described in the report targeted or indiscriminately affected civilian infrastructure. In the view of the ICRC, if carried out in times of armed conflict, such attacks would be prohibited. First of all, direct attacks against civilian infrastructure and indiscriminate attacks would be prohibited. Secondly, even if the infrastructure or some parts of it had become military objectives (such as a part of an electricity grid), IHL would require that only this part be attacked, and that there be no excessive damage to the remaining civilian parts. Thirdly, IHL would require parties to the conflict to take all feasible precautions to avoid or at least minimize incidental harm to civilians and civilian objects. Finally, even when they do not amount to attacks under IHL,[3] such operations might also be prohibited by the specific protection afforded by IHL to medical facilities or objects indispensable to the survival of the population. These are powerful protections that remain entirely relevant in view of the technical characteristics of cyber operations. For IHL to truly provide legal protection to civilians against the effects of cyber warfare, however, States must commit to its applicability and to an interpretation of its rules that is effective for the protection of civilians and civilian infrastructure. In particular, it would require a clear recognition that cyber operations that impair the functionality of civilian infrastructure are subject to the rules governing attacks under IHL.[4] This report will hopefully help illustrate the need for such an interpretation to ensure that civilian infrastructure is protected.

[3] Under IHL, "attack" has a specific meaning which would not encompass all cyber operations that are referred to as cyber attacks in a colloquial sense. See Chapter 2(c) below and Part 3(f) in the background document contained in Annex 3.

[4] See ICRC 2015 IHL Challenges report, p. 41 (see note 1 above).

Avenues that could be explored to reduce the potential human cost of cyber operations

Cyber security measures

Beyond the restraints imposed by IHL upon those carrying out cyber operation, it is critical to enhance the cyber security posture and resilience of the actors potentially affected. While cyber security and defence are constantly improving, older systems with outdated or even non-existing cyber security are particularly vulnerable to cyber attacks and will remain a concern in the years to come. Both the public and private sectors have a role to play through industry standards and legal regulation.

In the health-care sector, for instance, the regulatory environment should be adapted to the increased risk, such as through standardization requirements, with a view to ensuring resilience in the event of a cyber attack. Cyber security needs to be taken into account in the design and development of medical devices and updated throughout their lifetime, no matter how long they last. Similarly, for industrial control systems, industry standards, whether imposed or self-imposed, are critical. This includes reporting incidents and sharing information between trusted partners.

In terms of IHL, parties to armed conflicts must take all feasible precautions to protect civilians and civilian objects under their control against the effects of attack. This is one of the few IHL obligations that States must already implement in peacetime.

Disclosing vulnerabilities

The preferred option for enhancing the safety of cyber space should be disclosing vulnerabilities to the appropriate software developer so that the vulnerabilities can be fixed. Some States have therefore put in place equity processes to balance competing interests and risks and decide whether to disclose the vulnerabilities they identify.

Measures to prevent proliferation

Those who develop cyber weapons should consider creating obstacles in order to make repurposing difficult and expensive. While it is hardly possible from a technical standpoint to guarantee that malware cannot be repurposed, methods like encrypting its payload and including obstacles in different components of the code, for example, could raise the bar in terms of the expertise required to reengineer malicious tools. While there is currently no express obligation under IHL to create obstacles to the repurposing of cyber tools, this could prevent at least some actors from doing so and therefore reduce the risk of subsequent misuse that their proliferation entails. The unique way in which cyber tools proliferate also raises the question of whether existing law is adequate or sufficient to address this phenomenon.

Marking of certain civilian infrastructure

Another avenue, which builds on existing international law, could be to create a "digital watermark" to identify certain actors or infrastructure in cyber space that must be protected (such as objects that enjoy specific protection under IHL). The aim would be to help their identification and prevent them from being targeted during armed conflicts. The potentially positive effects in terms of protection against unintended harm by law-abiding actors would however need to be balanced against the risk of disclosing information on critical infrastructure to potential adversaries, including criminals. The prospects of positive effects might depend in part on attribution becoming easier.

Improving attribution and accountability

Finally, enhanced attribution capacities would help ensure that actors who violate international law in cyber space can be held accountable, which is a means to strengthen compliance with the law and more generally encourage responsible behaviour in cyber space.

Way forward

The use of cyber operations in armed conflict is likely to continue and might remain shrouded in secrecy. Analysing its consequences is a complex and long-term endeavour that requires multidisciplinary expertise and interaction with a wide variety of stakeholders.

Building upon the conclusions reached at the expert meeting, the ICRC would like to pursue the dialogue with governments, experts and the IT sector. It looks forward to the feedback to this report to continue to follow the evolution of cyber operations, in particular during armed conflicts, and their potential human cost, explore avenues that could reduce them, and work towards a consensus on the interpretation of existing IHL rules, and potentially the development of complementary rules that afford effective protection to civilians.

Introduction

Cyber attacks[5] occur on a regular basis and cause substantial economic costs. The attackers, targets, victims, level of sophistication, purposes and impacts of cyber operations vary widely.

Today, most known cyber operations have no apparent link to an armed conflict. However, a few States have publicly acknowledged that they have used cyber operations during armed conflicts; cyber operations have affected other countries involved in armed conflicts; and an increasing number of States are developing military cyber capabilities.

As part of its mandate to work for the clarification of IHL and, if necessary, prepare any development thereof, the ICRC monitors the development of new technologies that are, or could be, used as means and methods of warfare during armed conflicts. This approach is based on legal, technical, military and humanitarian considerations, which are interrelated.

To develop a realistic assessment of cyber capabilities and their potential humanitarian consequences in light of their technical characteristics, the ICRC brought together scientific and cyber security experts from all over the world to share their knowledge about the technical possibilities, expected use, and potential effects of cyber operations.

In particular, we sought to gain a better understanding of the risk that cyber operations may result in death, injury or physical damage, affect the delivery of essential services to the population, or cause systemic effects on the internet.

The structure of this report follows the agenda of the expert meeting. We began by discussing cyber operations in general and how they can be analysed (Chapter 1). We then delved into cyber attacks that may affect specific sectors, namely the health-care sector (Chapter 2), or various industries providing essential services such as energy or water (Chapter 3), and cyber attacks that may have a global or systemic impact (Chapter 4). We looked at the specific vulnerability of these sectors and the risk that cyber attacks may cause harmful consequences. We then turned more specifically to the use of cyber operations during conflicts (Chapter 5). We concluded by analysing the protection afforded by existing law and looking at avenues that could be explored to prevent or alleviate the human cost of cyber operations (Chapter 6).

While the various points made in the discussion summarized here are not attributed by name to the experts who made them, a list of participants is provided (Annex 1). The agenda also included a number of questions with a view to guiding the discussions (Annex 2). A background document was submitted to the experts in advance of the meeting (Annex 3).

Neither the background information nor this report necessarily represents the view of the ICRC.

Cyber operations raise many other issues and challenges such as cyber espionage, intellectual property theft, surveillance and privacy concerns, and the use of cyber means to further information operations (for example, by leaking hacked information or using social media or other cyber means for propaganda or disinformation purposes). While some of these operations may be related to or occur during armed conflict, they were outside the scope of this expert meeting.

[5] The terms 'cyber attacks' and 'cyber operations' are used throughout the report in a technical (mainstream or colloquial) sense and not as they may be understood under international humanitarian law (IHL), unless specifically stated (see the first paragraph of Session 1 below for more details).

Session 1: Cyber operations in practice

During the first session, the discussion focused on gaining a better shared understanding of cyber operations and how to analyse them. It should be noted that the terms "cyber attacks" and "cyber operations" are used throughout this report in a technical (or mainstream) sense and not as they may be understood under IHL, unless otherwise mentioned. In general, cyber attacks refer to any cyber operation carried out without the consent or knowledge of the owner of the targeted system, to obtain access, extract data and/or encrypt, degrade, delete, modify or disable data or services. This understanding is far broader than the meaning these terms would have under IHL, as IHL applies only during armed conflicts, and the notion of attack has specific meanings under this body of law.[6]

A. Understanding cyber operations with the cyber kill chain model

The experts agreed that the cyber kill chain model is a useful tool for describing cyber operations. The cyber kill chain model comprises seven phases, namely:
1. Reconnaissance
2. Weaponization
3. Delivery
4. Exploitation
5. Installation
6. Command and Control
7. Actions on Objectives[7]

The cyber kill chain needs to be understood as a non-linear model. In practice, kill chain steps are repeated in order to achieve the final aim. The timespan of a given operation will vary depending on factors such as the aim, the type of target and its environment, the circumstances, the urgency of achieving the aim and the risk that the attacker is prepared to accept for the operation – including the risk that the attack is subsequently attributed to it. One expert gave the example of the Olympic Destroyer campaign (2018), where the entire operation was estimated to have taken around two months.

The command and control phase enables the implanted malware to be controlled and situational awareness to be maintained. Command and control offers the controller the ability to decide which actions to take, and when, including with a view to reducing the risk entailed by the operation. Command and control is also necessary if the operators want to maintain the ability to perform additional actions (such as cleaning up after the operation). On the other hand, if the operators already know the environment, malware can be designed to omit some of the kill chain phases. For example, when access to the target facility and the right knowledge already exists (i.e. from past reconnaissance), there may be no need for the command and control phase: the malware can be set to operate on its own, in a "fire and forget" manner. One expert noted that Stuxnet had specific rules in deciding when the destructive payload had to actually be delivered; this measure was probably meant to reduce the risk of attracting unwanted attention or damaging the wrong target, and to ensure functionality in case the command and control channel was disrupted. It also had the command and control component, although it wasn't clear whether that was really needed in the final stage of the operation. In the more recent case of Olympic Destroyer, the malware did not need command and control either, since the operators knew exactly how to reach their goals, and the malware was set in motion through a timer and a self-propagation algorithm.

B. Operational purpose

The experts emphasized that the characteristics, operational approaches and impacts of cyber attacks could vary widely depending on the purpose of the operation and the tools and techniques employed.

[6] Art. 49 of Additional Protocol I of 8 June 1977 (AP I). See also Session 6 below.
[7] For more details on these phases, see Part 1(d) in the background document contained in Annex 3.

They noted that the most common operations were conducted for purposes of reconnaissance, surveillance and the exfiltration of data and information (for espionage or other purposes, often referred to as computer network exploitation (CNE)) and would usually involve gaining access to, and often maintaining a persistent presence on, the targeted system or device. These operations are generally designed to avoid detection and are not aimed at harming the targeted system or device, which could nevertheless be disrupted or destroyed unintentionally.

In contrast, the experts pointed out that cyber operations might be intended to generate effects on or against a targeted system or device. This is often referred to as a computer network attack (CNA). CNAs deploy a harmful effect on the targeted system – such as deleting or tampering with data, or disabling or physically damaging the system – and can even have physical effects on human beings, as discussed in the next session. The distinction is primarily one of purpose, and from a technical perspective CNAs may be achieved by the execution of a specific payload at phase seven of the kill chain, although simple command-line access may suffice. It was underscored, however, that not all cyber tools and operations are the same. While some may cause substantial impacts on the targeted system or device by design, others may employ precisely targeted technical means that barely affect how the targeted system works but that support or achieve greater operational objectives.

When the defender knows little or nothing about the actual purpose of the operation, its reaction may be to consider what the enemy could achieve after compromising the system (in a CNA) and to identify the potential worst-case impact of such an operation. The fact that CNEs can be upgraded to CNAs relatively simply adds to the defender's uncertainty about how best to react. This escalation risk factor may give rise to a potentially harmful over-reaction.

The experts noted that some actors worked to obtain and maintain a persistent presence on devices and in systems for various purposes. For example, VPNFilter is a malware that has been found to have targeted and infected large numbers of enterprise, small office and home routers and network-attached storage devices, and it remains on the devices even after they are rebooted. This and other malware can be used in various ways, such as for reconnaissance or anonymization (when using the compromised hosts as proxy servers to connect to other systems in potential future attacks), or to turn compromised nodes into a botnet. Finally, the experts added that it might also be possible to use them to cause a wide-spread impact, such as knocking large groups of people off the internet at the same time. More generally, when persistent access has been achieved, it could be upgraded to obtain effects if and when the actor's objective changes. For example, at the start of an armed conflict or during an ongoing one, States or other belligerents might decide to leverage their access capabilities in order to deliver effects. These actors will have already developed and completed stages one to six of the kill chain when building the capabilities during peacetime, which reduces the time required to achieve an effect. However, this might not necessarily always be the case, given the dynamic nature of cyber space and the ephemeral nature of accesses and capabilities.

Various experts noted that most operations currently studied occurred in peacetime and not during armed conflicts. This influences how the operations are carried out and how they are perceived. Cyber operations during armed conflict might look different. For example, actors may accept a higher risk of detection or attribution in view of the ongoing hostilities and the time available to achieve operational objectives. One expert illustrated this with Stuxnet, noting that the timeline might have been dictated by the diplomatic context. The time constraints that belligerents can face during armed conflicts are, of course, of a totally different magnitude. It may therefore be easier to attribute attacks during armed conflict. More operations to create effects might be seen, such as temporarily disabling cyber systems and/or the facilities that rely on them. However, the targeted facilities may not necessarily need to be destroyed or rendered inactive for a long period. Shutting them down for a certain amount of time could be sufficient to obtain the objectives sought by the belligerent.

One expert explained that experience in tracking threat actors showed that their reconnaissance activities were sometimes visible and could make it possible to identify probable end targets. Less frequently, it may even be possible to identify the technologies being researched and developed by the attacker before they are used, which again may enable conclusions to be drawn about the type of potential targets. But in general, this kind of visibility is rare. It might be difficult to fully assess the attacker's purpose until the actions on objective are actually implemented (last phase of the kill chain model).

C. Trusted systems and software supply chain attacks

The discussion turned to trust in the operating systems on devices or computers and the software running on them, and the notion that trust is often an implied assumption. In particular, connected devices trust other devices to supply components such as software (or updates) or to input data. This means that the device receiving the data from the trusted source will assume that the command, update or other data received are correct and will implement them accordingly.

The experts agreed about the risks posed by attacks that compromise trusted systems or the supply chain, to the extent that the compromised system or device offered a means of accessing any connected devices or networks. For example, the 2017 NotPetya campaign is reported to have started with attackers gaining access to the infrastructure of an accounting program called M.E. Doc, which was widely used by companies in Ukraine. This access was then used to deliver the malware.

In addition, some systems are trusted to a much higher degree than others. Compromising the trusted systems and maintaining persistent access could provide further access to many other systems.

In view of this, a State actor developing its readiness to act in the event of an armed conflict might want to achieve persistent access to one or more of these important trusted systems. One expert estimated that the majority of the computer devices in the world were only one or two steps away from a trusted system that a determined attacker could compromise.

One expert provided an example,[8] observed in 2018, of a slightly different method of reaching the same goal when the attacker had been unable to penetrate the target. Looking for an alternative method, the threat actor decided to build an actual product. It is likely that legitimate developers were hired to create real software. The product was promoted and ended up being installed at companies that did not realize the true nature of the threat. At some point the software received a malicious update, compromising the system. In this case, even though the malicious update may appear to be a supply chain attack on a legitimate product, in reality the entire operation was a supply chain attack cover-up from the start.

D. Cyber capabilities and exploits

One expert explained the notion of exploit as the combination of two elements: knowledge of a programming mistake in a target operating system or software, and a sequence of steps to be performed to exploit that mistake and cause an undesired and unexpected effect in the targeted program or device. In most cases, exploits offer options for gaining unauthorized access to a targeted computer system (including privilege escalation) and, in some cases, for delivering a follow-on effect. One expert considered that Stuxnet was still one of the most sophisticated "cyber weapons" ever used.

One expert noted that one of the reasons exploits existed was that software vendors may want to hurriedly publish new features and programs in response to market demand. By failing to focus enough on security during the engineering process, they create risks for customers and users. This expert pointed out that it was possible to write software that was not vulnerable, citing OpenSSH[9] and DJBDNS[10] as examples of longstanding serious vulnerability-free software, although the underlying economics do not align with the needs of most commercial businesses. Another expert noted that the enormous complexity of many software products (Windows, for example, has more than 60 million lines of code) could exceed the capacity of humans to prevent every unintended interaction. For this expert, current engineering practices are insufficient to ensure large, completely vulnerability-free software products for the foreseeable future, and actors actively looking for vulnerabilities with a view to developing exploits for malicious use contribute to the concern.

Experts noted that a market existed for selling and buying exploits. Prices skyrocketed over the last decade, with a hundredfold increase since 2000. This explosion in prices is driven by an increased scarcity of exploitable vulnerabilities, to some extent due to the increasing attention to security and good engineering practices within large software development firms, and price inelastic demand for exploits for important platforms. One expert offered the view that States' interest in developing

[8] Kaspersky Lab, *Operation AppleJeus: Lazarus hits cryptocurrency exchange with fake installer and macOS malware*, APT Reports, 23 August 2018.
[9] For more on OpenSSH see https://www.openssh.com or https://en.m.wikipedia.org/wiki/OpenSSH.
[10] D.J. Bernstein, https://cr.yp.to/djbdns.html.

intelligence and/or military cyber capabilities might be one of the factors fuelling the zero-day exploits market.

Experts noted, however, that while some systems had a very good security posture, most did not fare so well, and that this was reflected in the shelf price of exploits available for purchase. In 2018, the cost of exploits for some systems or products were hundreds of thousands – or even potentially one million – dollars, while exploits for less secure systems cost much less (e.g. $10,000 for a bug in an Internet of Things device). The possible need to rely on systems or products with weaker security may make it more challenging to achieve or maintain a strong cyber security posture. It also shows that systems and products differ widely when it comes to their security level and the difficulty in exploiting them, and it is expected that these discrepancies will remain. Furthermore, the price of acquiring exploits could go up if they prove to be more cost-effective than other means for espionage purposes.

It was noted that many exploits were developed and then put on the shelf, available but not actually used. Some of the exploits may be retained for two or three years without being used, while others are eventually used when needed. For example, software engineering practices at the targeted system or facility may lead to some functionality being removed, requiring actors to use a different exploit, which they had been keeping on the shelf. In this sense, stocking exploits could be a measure meant to diversify the portfolio of access capabilities. On the other hand, security updates, vulnerability patches or other action at the targeted organization could render such exploits useless.

This presents a parallel with some traditional armed conflict activities, where actions need to be taken in a timely manner. There may be a specific window of opportunity in which to strike. In such cases, the belligerent needs to have both the required capabilities at hand and situational awareness. There is also reason to believe that during conflicts there would be more cyber attacks relying on tools that had already been installed in the targeted facility and were waiting to be activated.

Furthermore, the expert noted that although some malware cost millions of dollars to develop, the tools lost value (in some cases because they had become publicly known) and were not being used. There is indeed a move towards using tools that are available for free. Such tools remain effective as long as the target does not have appropriate cyber security or detection systems. In fact, it is often the case that the targets' security posture is too weak to detect or deter cyber operations. Furthermore, using such tools, which are less specific and not custom-made, might help protect the user's identity; the tools' technical characteristics could help identify the developers, not the users – and there could be numerous unrelated users.

E. Evolving nature of the threat actors and the growing attack surface

The experts noted the wide range of actors carrying out cyber operations: individual hackers; criminal groups, potentially motivated by financial gain; States; non-State armed groups; and other non-State actors. Furthermore, various actors may cooperate, whether it be State alliances, States supporting groups, or criminal groups selling cyber capabilities to other actors. Some of the active sophisticated actors are known under the term advanced persistent threats (APTs), namely threat actors that establish a persistent, long-term access to the targeted system(s).

One expert noted that the private sector was increasingly interested in "cyber capability development", as is already the case with the development of traditional means and methods of warfare. This is an ongoing process, with some States being more open than others to engaging in such collaborations. There are also certain groups that already advertise malicious payloads for sale. These groups could cooperate with other vendors, and they could provide services to different countries at the same time. While some States may still prefer to keep all the operational aspects in-house, others – whose intelligence services or military have not developed cyber capabilities – can still be active by buying the appropriate toolkit, or even have an operation performed on their behalf against a target they have identified. Outsourcing may often be cheaper than developing and maintaining in-house capabilities. This creates a viable business for those who can supply services and find customers. As a result, attributing attacks may become more difficult: the technical data may point to one single actor, the developer, although there could be multiple users.

Experts also emphasized that the exponential growth of cyber space through the Internet of Things (IoT) increased the attack surface. Any connected device can become a target or part of an offensive

cyber operation (e.g. a bot in a botnet). It is no longer merely about exploiting operating system vulnerabilities. Now, devices like pacemakers[11] and automobiles are becoming connected to some degree. In autonomous cars, software will be doing what humans now do. But this software will need communication features, and this will create new opportunities for attacks.

One expert recalled that the human element would always be one component of the attack surface, because systems were actually operated and used by humans. Social engineering will remain part of the attackers' tools.

F. Cyber vs kinetic attacks[12]

The experts offered some considerations with regard to the strategic and operational nature of cyber operations, including how they compared with kinetic weapons.

Some experts noted that cyber operations might enable one State to attack another State in the absence of the kinetic capability to do so. Also, most advanced weapon systems rely on connected computing systems, which could present vulnerabilities (even if such systems are probably well-protected). Cyber capabilities could therefore be used in an asymmetrical manner by less sophisticated or powerful belligerents.

One expert said that, in general, it was easier to conduct espionage through cyber means than through traditional means such as spies, yet it was easier to achieve destructive effects through kinetic attacks than cyber attacks, at least for now. Looking more specifically at the cost factor, experts noted that it could not be said in general that cyber operations were, or would become, necessarily cheaper than kinetic operations. This might however be the case depending on the circumstances. For example, one expert speculated that it might have been cheaper to launch missiles to physically damage the Natanz plant than to do so with Stuxnet. But a kinetic operation might have raised different political and legal consequences. Similarly, it may be cheaper to obtain specific information on an adversary with a human spy in the territory of another State than with a CNE; however, the spy runs the risk of being captured. So even assuming that a specific cyber operation would cost more than its equivalent human or kinetic operation, other reasons might make it more advantageous to opt for the cyber operation. Among other factors, actors operating in cyberspace, including States, might deem it easier to deny responsibility for operations conducted in cyber space than for activities carried out in the physical world.

Another expert suggested that there may be a point where resorting to cyber rather than kinetic means might become the preferred choice based on moral, ethical and legal considerations. The expert provided the example of the alleged Israeli operation to shut down the Syrian air defence network by cyber means in 2007,[13] without causing any casualties or long-term physical destruction. The expert wondered if, in the future, the prioritization of cyber means may become an obligation for parties to armed conflicts.

G. Attack and defence

The experts discussed attack and defence in detail. They agreed that the defensive side was improving – some of them considered it to be improving fast. But the problem of the security posture in general is complex. The experts provided various views.

One expert said that while malware was becoming more advanced, the ability to defend against it was also getting better. The expert acknowledged that the question of access was one issue where eliminating the risk was difficult: since the role of networks is to facilitate communication, cutting network access is often practically impossible. However, looking only at successful attacks puts too much focus on organizations at the lower end of the security scale. In the view of this expert, despite the contrary stance often mentioned publicly, actual defensive capabilities had greatly improved, in particular with regard to detecting threats within systems and removing them. The expert said that, as a result, the persistence of APT operations was decreasing, although this was disputed by another expert. Even if progress has been made in developing software with fewer vulnerabilities, the ability to

[11] See Chapter 2(c) below.
[12] For more details on the military use of cyber operations, see Session 5 below.
[13] See Part 3(b) in the background document contained in Annex 3.

create invulnerable systems will remain elusive. One potential solution is to build layers, with particular emphasis on making the critical operating system highly secure. But, the expert emphasized, the real objective was to reach a point where the efforts that needed to be devoted to offensive and defensive actions became balanced so that attackers found their task increasingly difficult; this was already occurring.

Another expert noted that only a few really novel offensive techniques had been observed over the last decade. Also, not all vulnerabilities are necessarily problematic. Although the IoT carries risks related to the growing number of connected devices and their decentralized operation, services such as those offered by cloud providers bring benefits of centralization, along with the strong security expertise and resources of the cloud operator. In essence, this is an example of the democratization of defence, where solid defensive tools are provided to everyone at a very low – or even no – cost. The expert predicted that in 10 to 20 years we would see a shift in the focus of attacks, including those by States, to the supply chain, where the system and software providers might become the targets.

Several experts shared less optimistic predictions, highlighting among other things the difficulty of anticipating what would happen in 20 years. Part of the unknown is the extent of serious cybersecurity problems caused by legacy systems (old, outdated or unmaintained systems still in use, instead of newly available or upgraded systems); this will remain a relevant issue in the years to come. One expert warned against thinking solely of the "Hollywood military hacker" model; what was important was to focus on the actual consequences of the very basic techniques used in cases such as in Estonia (2007). Today's attackers do not need to have "super powers" to cause significant economic harm; they merely need to be smart. This will still be the case in 20 years. Another expert used the Mirai botnet as an example, in which IoT devices with simple vulnerabilities were massively compromised. The expert expressed doubts about the security posture of new devices such as smart TVs and smart cameras.

Furthermore, a couple of experts also noted that while it was possible to build better defence systems, the economic incentive to do so tended to be misaligned. Exploitation is the act of taking advantage of programming errors, and many of these errors are the result of business practices prioritizing the release of products quickly at the risk of jeopardizing user security. The experts did not believe that actors would stop exploiting errors in the next few years – every defensive measure can be overcome – but some of the experts expected the entry barrier for exploit development to continue rising.

In the longer term, hardware changes that make previous exploit techniques obsolete could mark a clear improvement. One expert drew attention to some future processor design changes (in particular ARM v8.5A Memory Tagging[14]), which would significantly reduce the exploitability[15] of most software vulnerabilities, potentially even undercutting the exploit development market and making it less viable. This could conceivably change the nature of cyber attacks as we know them today.

A couple of experts referred to possible risks created by the hypothetical construction of quantum computing devices. Another one noted the risk that, because of nano-technology, small chips with major capabilities could be surreptitiously embedded in manufacturing processes. Finally, one expert highlighted that the importance of IT was growing in certain countries that did not necessarily have a strong defensive infrastructure or heightened user awareness. Such countries could end up as big "bot centres". This expert stressed, however, that the situation was not the same around the world, which made it difficult to identify general trends.

One expert cautioned against being overly confident that some services would be as secure as was widely believed. It may cost a few dozen dollars to obtain credentials to email accounts, perhaps even at some big email providers. Flaws that are found are sold to cyber criminals who can later weaponize them. Just because there is not widespread knowledge of the flaws does not mean they do not exist. This expert also speculated about whether governments were forceful enough in promoting good defensive practices, considering their desire to retain the ability to conduct various types of operations. On the positive side, this expert considered that the level of operating system and software security was often very high. However, attackers are now moving to target firmware, where there is a lack of resources, expertise and defensive measures. If defences improve there as well, attackers may move on

[14] "Arm A-Profile Architecture Developments 2018: Armv8.5-A", Arm Community, October 2018.
[15] K. Serebryani *et al.* "Memory Tagging and how it improves C/C++ memory safety", ArXiv, February 2018.

to other targets, such as to the supply chain, updating processes, signing keys, encryption and digital certificates. Various experts acknowledged the problem, noting that actors carrying out offensive operations would focus on those areas and systems with the weakest defences. Cyber offence and defence will continue their game of cat and mouse. Experts concluded that it was important to imagine what might happen, and act accordingly to prevent it. One expert summed up his opinion by noting that, in cyber space, the best defence was defence.

H. Importance and challenges of attribution

One expert warned of a world in which attribution became impossible; in such a world, attackers would not feel constrained by the potential consequences. Another expert illustrated the challenges raised by attribution with the 2018 Olympic Destroyer attack. Olympic Destroyer was a worm that propagated over the network with a component designed to delete and destroy data, rendering systems unusable. However, this malware did not self-delete (a common technique) after executing destructive payloads. The expert assessment was that the attackers wanted to be discovered. Curiously, the malware appeared to be created in such a way as to point towards other threat actors than the ones actually responsible for this attack. The attack was unusual because of this, and it is among the rare examples where the attacker's primary goal was apparently to cause misattribution.

Session 2: Cyber attacks that could affect the delivery of health care

The experts discussed various types of cyber attacks[16] that could affect the health-care sector: attacks affecting hospitals or other medical facilities, those affecting medical devices in hospitals, and those affecting connected biomedical devices. The experts then dwelled on the specific challenges to patching vulnerabilities in medical or biomedical devices. The final section in this session summarizes the discussion about the resilience of the health-care sector when faced with a cyber attack, including some avenues for improvement.

Experts agreed that the health-care sector was particularly vulnerable to cyber attacks because of, *inter alia*, the sector's relatively weak cyber security posture, and that the risks were serious and potentially fatal. They highlighted that the growing use of connected devices was increasing the attack surface and, thus, the cyber security challenge. Throughout the session, experts underscored the importance of adopting regulations designed to enhance cyber security in the health sector.

A. Cyber attacks that could affect hospitals (or other medical facilities)

Examples of cyber attacks that have affected hospitals

The experts recalled various recent cyber attacks that affected hospitals: the WannaCry ransomware in 2017,[17] the 2016 ransomware campaign against a hospital in Hollywood (which seriously impeded patient care for several days),[18] and the 2016–17 attack in Singapore.[19] The investigation into the Singapore attack revealed that the infection lasted for more than ten months, and that the data of some 1.5 million users (including 16,000 medical prescriptions) were exfiltrated. The attackers performed database queries for specific patients, including the prime minister. So far, no attempts to tamper with prescriptions have been reported, but this is a potential risk. While details of the operation have not been made public, one expert noted that the attack appeared to have been targeted, and the tools used were comparable in sophistication to those employed by State or State-sponsored actors. The attack did not appear to have affected the delivery of health care, and the motive or ultimate purpose of the attack remains unclear. In the cases of WannaCry and the Hollywood hospital, the attacks prevented the medical facilities from operating normally by hampering system and data availability.

Growing digitization and interconnectivity

The experts noted that hospitals around the world were strongly moving towards increased digitization and interconnectivity, both for their own operations (internally and when relying on cloud-based services) and for communications with other actors in the health-care sector (other hospitals, laboratories, patients, suppliers, insurers, etc.). While greater connectivity increases the potential attack surface, necessary improvements in cyber security have not taken place at the same pace. One expert mentioned a case where a machine used to sterilize medical equipment provided a channel to reach into a hospital network. The experts warned that the situation could get worse in the near future. For example, policies that allow staff to use personal IT devices mean that devices with a potentially lower security posture can connect to hospital networks.

It was noted that the more digital dependencies were ingrained in the system, the more difficult it might become to operate when and if these dependencies stop functioning. Depending on the type and

[16] On the notion of "cyber attack" as used in this report, see note 2 above.

[17] See the background document contained in Annex 3 and the text in relation to notes 72 to 74.

[18] See S. Ragan, "Ransomware takes Hollywood hospital offline, $3.6M demanded by attackers", CSO Online, 14 February 2016; and C. Sienko, "Ransomware Case Studies: Hollywood Presbyterian and the Ottawa Hospital", INFOSEC Institute.

[19] M. Field, "Cyber attack on Singapore health database steals details of 1.5m including prime minister", *The Telegraph*, 20 July 2018; J. Au Yong, "Info on 1.5 SingHealth patients stolen in worst cyber attack", *Straitstimes*, 21 July 2018; Ministry of Communications and Information, *Public Report of the Committee of Inquiry into the Cyber Attack on Singapore Health Services Private Limited's Patient Database*, Ministry of Communications and Information, Government of Singapore, 10 January 2019.

number of facilities affected and the severity of future cyber attacks, it could be made impossible to treat patients. For the experts, the public might not fully appreciate these risks, because no significant crisis had occurred so far.

Likelihood of attacks against hospitals: trends and factors
While some attacks might affect hospitals and other types of targets indiscriminately or incidentally, there is a growing trend of attacks that target hospitals directly – for example with the Orangeworm group.[20] However, the experts noted that most of the attacks appeared to be opportunistic, and even the more advanced ones were not meant to interfere with the delivery of health care. Indeed, cybercriminals may see hospitals as soft targets and attacking them as an economically viable option. For hospitals, the priority is to provide health-care services to patients whose lives are at stake. They are therefore more likely than other types of potential targets to pay a ransom in case of a cyber attack that limits their ability to use their systems and data (e.g. through a ransomware attack). Paying a ransom to recover the data or regain access to the system may be faster than using backup systems (assuming such backups exist in the first place). This will remain the case until suitable defensive measures are developed, although one expert believed that many technologies already existed to recover from ransomware wiper attacks.

Cyber attacks on other actors in the health care sector
It was also noted that hospitals were not the only potential victims in the health-care sector. Other health-care facilities can be attacked and infected as well. For example, pharmaceutical companies may be targeted for intellectual property theft. Even though these operations may not be destructive in nature, they may still cause unintentional damage beyond the property theft. Even more concerning would be attacks on research facilities that store dangerous materials, such as viruses.

B. Cyber attacks affecting medical devices

Cyber security challenges raised by medical devices
Hospital computers are typically divided into two groups. First, there are the regular computers for the hospital's management; these include the hospital's patient files, typically controlled by the hospital's IT department. Second, there are the computers embedded in medical devices such as magnetic resonance imaging scanners, which are not usually controlled by the hospital's IT department. IT-controlled systems run on general-purpose operating systems and are more likely to be incidentally affected than medical devices. Medical devices are normally connected to the hospital IT system, as this enables immediate electronic filing and reduces the risk of error. However, such connections represent a potential entry point for a malware that could affect medical devices.

Experts noted that, in their view, the manufacturers of medical devices did not take cyber security fully into account in designing and developing these products. Certification requirements for connected medical devices do not usually focus on cyber security. The certification requirements may actually compound the problem, since responsibility for managing and repairing the devices and their software components is held by the manufacturer. As a result, if the hospital's IT services were to install security updates on such devices, the hospital could be held liable if the devices then malfunction. This problem is aggravated by the typically long life span of medical devices; they may very well be running on outdated software (see Chapter 2(d) below).

Likelihood of cyber attacks against medical devices
One expert expressed doubts about the likelihood of attacks incidentally affecting medical devices, because medical devices were typically very specific in the way they operate. For this expert, attack tools would have to be purpose-built in order to affect medical devices. Other experts noted that while some medical devices were run by purpose-built software with limited specialized functionality, others were more vulnerable because they ran on often obsolete general purpose operating systems. This exposes them to the risk of infection by malware (including self-propagating malware) that takes advantage of the general purpose operating system used to control the device.

[20] "New Orangeworm attack group targets the healthcare sector in the US, Europe, and Asia", Symantec Blogs, 23 April 2018.

Furthermore, many experts underscored the fact that offensive tools could usually be repurposed. So malware originally designed for one type of target might be repurposed for use in specific attacks on the health-care sector, including those targeting medical devices.

One expert emphasized that significant numbers of cybersecurity incidents at medical facilities were not, or not sufficiently, investigated. This expert highlighted the need to rely on people with cyber security expertise, especially in cases where a patient's death might be linked to the malfunctioning of a biomedical device (see below). Cyber forensic inspections could help establish if the malfunction was purely accidental or if it may have been caused by a cyber intrusion, targeted or not.

C. Cyber attacks affecting biomedical devices

The experts turned to the specific risks posed by connected biomedical devices such as pacemakers and insulin pumps. They noted the advantages offered by such connectivity, such as to remotely monitor a patient's health or operate the medical device. If the device allows changes to be made remotely, doctors can react immediately to a patient's evolving health situation by changing the device's settings themselves. Depending on the system, this may require a Bluetooth connection between the biomedical device and another internet-connected device (e.g. a phone or server), and it may involve data transiting over the internet and cloud storage for both data and applications. Each of these layers presents vulnerabilities that could be exploited. One expert illustrated some of the potential risks: battery depletion through a type of DDoS attack; reprogramming the device to malfunction; death threats or extortion; and even murder attempts.

The expert pointed out that, in many cases, there might be no need for specialized malicious code to exploit the vulnerabilities in the biomedical devices themselves (such as in the firmware) in order to conduct an attack. Instead, it may be sufficient to produce an update designed to cause the device to crash. To protect against such risks, implantable medical devices should be equipped with safety mechanisms. Some devices already have this type of feature. This expert emphasized that patient survival might depend on such safety mechanisms, especially in the case of pacemakers. Like medical devices, connected biomedical devices have a long life span (a decade or more). The question of cyber security is unlikely to have been taken fully into account in their design. Furthermore, subsequent cyber security updates may still be insufficient, given the need to proceed cautiously and consider the potentially fatal risk of uploading a security update that causes the device to malfunction.

The experts noted that a large-scale cyber attack on connected biomedical devices would probably be quickly noticed (and hopefully stopped) due to its visible effects. One expert, however, added that attackers might have ways of decreasing the chances of advance detection and therefore increase the potential number of casualties.

D. The challenge of fixing vulnerabilities in medical devices

One expert suggested that one systemic reason behind the cyber security weaknesses of medical products might be that most programmers involved come from engineering fields, with no specific background in software coding. This may result in the source code not meeting the appropriate standards and being difficult to maintain in the long run, and in systems not being properly documented.

The experts highlighted the fact that hospitals might not be in a position to know the vulnerabilities in the medical devices they use because the device details, including the list of the system's software and hardware dependencies, were not usually available. The long life span of medical devices is another challenge, as products still in use may no longer receive support or updates. One expert called for more product transparency, including the use of standardized protocols, and for external security researchers to scrutinize the devices. The creation of regulations to compel device makers (i.e. suppliers) to be more open about these details could also help. Another expert called for a responsible end-of-life procedure for systems and products in widespread use, which could include keeping a source code register; such a register could help in the development of security updates for products that were no longer supported by their original vendors. That expert also suggested requiring suppliers to have someone available who can access the source code and update the product for a specific period of time.

However, when source code grows organically for 20 years, patches may become impossible to develop in practice, because nobody really understands the code and how it functions anymore. One expert noted that, even if the source code was made accessible through a register, creating patches outside of the original development process would be difficult. Moreover, such patches could even cause the systems to malfunction.

This software engineering problem may be systemic in the industry. One expert argued that the U.S. Food & Drug Administration (FDA) did not regulate cyber security on purpose, in order to promote innovation. Certifying medical devices for safety is already a lengthy process; including cyber security requirements would make it even longer.

E. Resilience of the health-care sector to cyber attacks

The discussion in Session 2 eventually turned to the resilience of the health-care sector to cyber attacks.

The experts discussed concepts such as redundancy and network segregation. Hospitals often have redundant systems so that they can quickly replace compromised systems with functioning ones. This can work especially well for centrally managed devices and computers. However, this type of redundancy is impossible in many cases, such as with magnetic resonance imaging (MRI) devices. Network segregation can also ensure resilience but, as noted above, hospitals' systems and networks often cannot be perfectly isolated. One expert emphasized the point that critical systems should not depend on an external internet connection, so that they would continue to function in case of an internet outage.

Another expert brought up how difficult it was, in practice, to recover from a cyber attack that had affected medical devices. Because the recovery process has to be done by the device suppliers, a simultaneous attack on many hospitals could, depending on the circumstances, overwhelm the suppliers' capacity to repair the systems in a timely manner. Urgent health-care needs at hospitals would remain manageable, but the situation would become unsustainable in the longer term.

The experts generally agreed that cyber security negligence was a challenge. However, one expert added that hospitals were working towards ensuring cyber resilience, although the measures they took might not always be fully effective. This is especially apparent in the wake of WannaCry and other recent high-profile cyber attacks against hospitals, which were a cyber security wake-up call for hospital managers. Each time a new IT process is introduced, medical facilities should consider the consequences. Another expert noted that standard measures in many hospitals included contingency planning and exercises for emergency situations. These contingency plans and exercises should also take into account the risk of cyber attacks.

The experts emphasized the need for established procedures and backups. Several of them felt that if the sector had specific standards, the recovery time following a cyber attack could be reduced. One expert noted that, from experience, recovering from an incidental disruption was easier and quicker than recovering from a targeted attack. In a targeted operation, the attacker may also target the recovery plan, such as by removing backups while disrupting the main systems.

The experts agreed on the importance of adequate regulatory action and standardization, saying that they were critical to ensuring cyber resilience. Several experts mentioned the NIS Directive recently adopted by the European Union.[21] They hoped that these kinds of regulations would lead to the development of mature cybersecurity standards. The NIS Directive also requires essential service operators, including health-care facilities, hospitals and private clinics, to take specific measures to avoid or minimize the impact of incidents affecting system security, and to notify the designated authorities of incidents. These requirements could help reduce risks and recovery times in hospitals and in the medical sector more broadly, improving the resilience of that sector. Another expert drew attention to the Hippocratic Oath for Connected Medical Devices, a commitment by individuals aimed

[21] *The Directive on security of network and information systems (NIS Directive)*, European Commission Policy, adopted by the European Parliament on 6 July 2016.

at improving the cyber safety of connected medical devices through design, third-party collaboration, evidence capture, resilience and containment, and updates.[22]

A few experts concluded that, while the distinguishing factor of the health-care sector was that human life was at stake, most cyber security problems observed in that sector were similar to those in other sectors. Consequently, similar approaches to enhancing the security posture, such as by establishing dedicated standards and educating people, would largely improve the situation. Furthermore, although deploying cyber security measures can be costly and complicated, the technological problems can be solved. One expert noted that cyber security was not rocket science, and that it should be possible to have medical devices with a reasonable level of security. That said, the health-care sector will always remain vulnerable to cyber attacks to some degree.

One expert noted finally that cyber attacks appeared able to cause human death and wondered why there had been no indications of this already occurring. One possible answer could be that no actors that have attacked the health-care sector with lethal intent have been identified. However, the discussion also highlighted the fact that incidents tended not to be thoroughly investigated and, consequently, it would be difficult to even establish whether the fatalities caused by a medical device malfunctioning were the result of a cyber attack.

[22] "Hippocratic Oath for Connected Medical Devices", I am the Cavalry.

Session 3: Cyber attacks that target critical civilian infrastructure or that may otherwise affect the delivery of essential services to the civilian population

A. Specific features of cyber attacks against industrial control systems

The experts described the three stages that cyber attacks[23] on industrial control systems generally involve:

1. accessing the IT network in the industrial facility
2. accessing the industrial control system itself
3. creating a (harmful) effect in the industrial process.

There is a significant difference between attacking an asset and attacking a process to create a physical effect. Taking advantage of the access established in the first two stages, the attacker needs to gather information on how the industrial processes and industrial control systems function, often by observing how the industrial control system works from within the internal network. Some experts had observed situations in which all the information necessary to attack an industrial control system, sometimes including the source code, was actually available online. However, they also noted that the available information was generally not fully accurate, including the documentation located in the IT networks of the targeted industrial facility. The experts noted that even if there might be many vulnerabilities, only a few might actually make it possible to create an effect on physical processes. This means that the attacker must really understand the industrial process and how all the monitoring, control and safety systems in place actually work. The safety mechanisms used in industrial processes depend to a large degree on sensors, signals, actuators and control loops. For an attack to be successful and create an effect, the attacker must systematically disable all these safety systems during the third stage of the attack. The harmful effect can be caused directly, by manipulating the industrial process (i.e. taking over control), or indirectly by deceiving the control systems about the state of the process (reducing their ability to monitor processes or interfering with safety systems). When the objective is to disable the system, facility or service for a long time, the attack may need to cause physical damage. Creating a harmful effect on the industrial process therefore requires both cyber and engineering expertise.

With regard to the second and third stages, the attackers are racing to exploit and disrupt systems that are not defended or are insufficiently defended. One expert explained that the 2015 attack against Ukraine's power grid targeted the human-machine interface level. This expert felt that the Windows-based human-machine interface was easy to harden and equip with appropriate security or defensive measures. Provided these security features are used, there is little sense in attackers looking for exploits at this level. Another expert qualified this view, stating that Windows-based malware remained a major threat. The 2016 attack took place at the level of the industrial protocols and was facilitated by the scarcity of means available to monitor the control networks. In terms of what this means for future attacks, the first expert felt that the industrial protocols' cyber security would improve and reach a satisfactory level in the near future. Finally, some attacks in 2017 directly exploited the embedded system:[24] Triton/Trisis malware targeted the safety instrument systems, which are the last line of defence before an incident.

Some groups have begun targeting vendors. Vendors build persistent access into products and maintain that access when the product is placed in the industrial control system environment, such as to enable remote monitoring. Attackers may attempt to compromise this form of legitimate access, and use it as a channel for malicious access.

[23] On the notion of "cyber attack" as used in this report, see note 2 above.
[24] Embedded systems are software controls that are usually tightly coupled with the hardware, including mechanical components.

Furthermore, in cases of supply chain attacks carried out through firmware, the many computing devices that an industrial control system comprises can be compromised. This type of attack may make it possible to gain large-scale access (subject to the uniformity of the devices within the market segment concerned and of their update schedule). However, the attackers' success in creating large-scale effects would depend on the target's characteristics.

In terms of the scalability of risks, the experts set out three scenarios: the risk of deliberate attempts to cause large-scale harm to industrial control systems; the risk that malware used on industrial control systems could cause unintended collateral damage to other industrial control systems; and the risk that existing self-propagating malware available in the wild could cause unintended collateral damage to industrial control systems.

In this regard, one expert pointed out that some diversity in the operational technology used in an industrial facility or network (e.g. in its operational processes or configuration) might prevent, or at least hamper, attack automation and scalability. Regarding the scalability of payloads in particular, another expert noted that some payloads might be reusable among different facilities – even irrespective of the equipment vendor – when targeting equipment that was subject to the same modes of failure (such as overstress). This makes the engineering part of the attack slightly easier. However, overcoming the devices' security boundaries remains a challenge and, in principle, the payloads must generally be adapted to the targeted facility.

B. Threat actors: number, purposes, resources, capabilities, and evolution

Experts underscored the limited amount of information available with regard to threat actors and risks. This is partly due to the fortunately limited number of successful attacks on industrial control systems seen so far.

Purposes

With regard to the objective of attacking industrial control systems, some experts noted that the attackers did not seem to be interested in the infrastructure in particular, but rather in secondary effects, for example undermining the confidence of the population. In such cases, an attack on a facility is a means to an end, and some level of penetration may be sufficient to cause grave concerns. Another expert suggested business competition as a possible driver of cyber attacks against industrial control systems.

Threat actors and resources

With regard to threat actors, one expert noted that there were 13 known and tracked groups that had the capability to prepare cyber attacks against industrial facilities. These groups appear to specialize in operations against particular sectors. However, there may be many more actors that have not been identified, a fact that undermines the validity of general assessments and predictions.

One expert recalled the 2015 cyber operation in Ukraine, where attackers sought to target five power distribution sites. However, due to resource constraints, the attackers could only focus on three facilities at the same time. When they went after the remaining two sites, they were discovered and the operators managed to disconnect the infrastructure. According to that expert's estimate, the access operation required around four or five cyber operators working across the sites, while the effects operation was highly labour intensive and probably required 20-25 cyber operators. This case highlights the direct relationship between the resources at hand and the effects cyber operators are capable of achieving in practice. In the case of Triton/Trisis, the same expert estimated that six people with cyber expertise were probably involved, consisting of two access operators, two developers, and one or two researchers developing the code (and possibly unaware of what they were working on). There is also evidence that the actor developing the attack tools had access to the actual hardware being targeted.[25]

More generally, this expert continued, for both stage 1 (access to IT infrastructure) and stage 2 (access to industrial control system infrastructure) operations, one or two people were typically needed. The challenge in cyber operations against industrial control systems comes in stage 3 (effects operations), which at the moment may require four to ten people with advanced expertise, depending on how much

[25] For a discussion on testing attacks, see Chapter 3(c) below.

time is needed and the availability of researchers. In general, it could take anywhere between a few people and a hundred people, depending on a wide range of factors, although most tasks will require tens rather than hundreds of people. It must be noted, however, that, beyond the number of people, such operations demand significant expertise, experience, tools and infrastructure. The following factors, among others, will affect the required capabilities: how strong the target's cyber security posture is; how wide and long-lasting of an impact the attacker seeks; the speed at which it needs to be done; the resources committed to the operation; whether the available resources need to be simultaneously spread among different tasks or targets or can focus on a single target; and whether the operation requires human intelligence and/or human involvement on the spot, since industrial control systems operations may be blended operations (i.e. not carried out solely through digital means).

Evolution of the capabilities and threats

Some experts noted with concern an acceleration in the evolution of attacks against industrial control systems, even if these attacks were not evolving as fast as traditional IT threats yet. Stuxnet was deemed a turning point in 2010, and over the years 2010 to 2015 a significant level of espionage, reconnaissance and weaponization of operational technologies (e.g. towards automated reconnaissance) has been observed. In 2015, however, experts would not have imagined that a threat as severe as Triton/Trisis could appear so soon.

One expert noted that all the groups that were being tracked had worked for at least 18 months before reaching stage 3 capability. In reality, so far only a small number of the actors that have the capability to create an effect have been detected. Another expert noted that, in the case of power grids, merely demonstrating the capability to affect the industrial control system was not necessarily proof of the capability to attack the whole process. However, groups lacking the required expertise could learn from and copy the smartest and best-resourced groups, including by repurposing what the latter may have developed and used. This could render attacks against industrial control systems cheaper, easier and less resource intensive – and thus more frequent. Experts also noted that this capability could be outsourced, and that expertise could be bought. Another expert, however, expressed doubt that attacks against industrial control system would become easier, given the rapid and constant improvement in cyber security.

Several experts stressed the difficulty of assessing how many operations remained undetected, how much reach the attackers really had into the infrastructure, or whether backdoors had been established for future use, for example as kill switches. Even bugs found in codes or products might be backdoors disguised as simple programming errors.

Finally, experts noted that States were probing industrial control systems as potential targets. Circumstances, such as an armed conflict, could trigger States' decision to go after these targets, and they may have the expertise, resources and access to information required to significantly reduce the timespan currently observed. In armed-conflict situations, operational requirements may dictate that the operations take place rapidly at the expense of stealth.

For these reasons, the experts found it hard to predict the pace and evolution of the threat with precision.

C. Attack testing

One expert noted that, in the same way that companies use penetration testing as a way of assessing their cyber security readiness, an increasing number of governments had well-established laboratory environments to test and anticipate the effects of potentially harmful operations carried out by criminals or adversaries. Conversely, the lack of readily available testing infrastructure for potential attackers may limit their ability to successfully attack industrial control systems. Yet if manufacturers and engineers choose to test their systems in a virtual environment, there may be an increase in leaks reaching malware developers – leaks related to both the system being tested and the virtual testing environment.

This expert also referred to Metasploit, a framework that helped in security testing but that could also be used by less skilled hackers to develop attacks, as it simplified the creation of exploits for IT systems. The expert wondered whether there would be an increase in effects operations against industrial control

systems if a framework similar to Metasploit also existed for operational technology (OT) equipment to help their security testing, as such a framework could also be misused to help develop malicious attacks against OT systems. The expert explained that effects operations were limited to single events because they were expensive and need time and human expertise.

Finally, one expert recalled that the malware used in the 2016 Ukraine attack included modular components, many of which were not used or even useful for attacking electricity grids in Ukraine. This observation raises the question of whether the attack was also a way to test the tools developed, with a view to potentially using them against other targets at a later stage.

D. Risk and quantification

Risk: severity, likelihood and uncertainty
The level of risk is a function of the likelihood that a harmful event will take place and of the severity of its impact if it does. The impact can be estimated through modelling and simulation. The likelihood of the event depends on the potential attackers' intent and resources and the objective difficulty of causing the event.

One expert noted that it is much more difficult to estimate the likelihood that a harmful cyber event will take place than it is to estimate the reliability of physical components. Physical components can be tested and their statistical failure rates determined, and on that basis safety measures can be taken (such as redundancy); however, the same is not possible for cyber. It is extremely hard to fully test sophisticated cyber systems. Furthermore, identical components have common vulnerabilities, so if one is discovered by a potential adversary, all the components can be exploited in the same way, possibly all at once. On the network, all critical infrastructure is nearby; the intent factor remains largely unknown; and there are too few comparable successful cyber attacks to have statistical models (no prediction can be made based on single events like Stuxnet).

Uncertainty also exists with regard to the effects of actions taken by the side that is targeted: introducing a new system results in changes in the environment that might interfere with the attacker's plans. If malware is already implanted in a system, occasional updates may effectively disable it by chance.

Another expert suggested a probabilistic risk assessment approach, in particular when looking at systems such as power grids or at hospitals in one region (taking into account the ability to refer patients from one hospital to another, rather than assessing the likelihood that a single specific hospital will be affected, which is much more difficult).

One expert stressed the importance of the people leading the risk assessments. Specifically, industrial control systems engineers and cyber security engineers may look at things differently, and the need to combine both views in the assessment teams was emphasized. To assess the risk, another expert underlined the importance of high-quality simulators for both the IT/OT systems and the physical systems, and to use the results from experimental tests together and include them in the simulations.

Likelihood of cyber attacks causing harmful effects in electricity grids
Some experts underlined the fact that it could be difficult for attackers to cause large-scale harmful effects in electricity grids because such grids had built-in redundancy and were composed of multiple entry points and nodes. One expert illustrated such difficulty with a discussion of the Lloyd's 2015 Business Blackout report. The report assesses the impact of a sudden loss of 18,000 MW of power after 50 to 70 generators have been damaged.[26] The expert noted that 50 generators chosen at random were unlikely to cause a loss of 18,000 MW; the attacker would need to target 50 specifically selected generators. It might therefore not be necessary to aim for 100% protection of every single device or node in the grid. The likelihood of the hypothetical event in the Lloyd's report actually taking place could be reduced by prioritizing the cyber security of the most important or powerful generators.

Another expert noted that the trend towards renewable energy sources had decentralized and diversified the power sources that make up the electricity network. The management of all these

[26] See Part 2(g) in the background document contained in Annex 3.

sources requires increasing digitalization of the system, which might increase the number of points potentially vulnerable to attacks. However, the number and diversity of power sources would complicate the ability of attackers to create a large-scale impact.

One expert noted, however, that while the grid as a whole might be resilient, if a critical system depended on a specific system or node, for example, and that node was vulnerable, attackers might still be able to achieve the desired effect. Furthermore, even small-scale attacks might have harmful cascading consequences.

Likelihood of cyber attacks causing harmful effects on nuclear power plants
A few experts considered that the likelihood of a catastrophic event being caused at a nuclear facility through a cyber attack was low because of the difficulty of causing such an event. Nuclear power plants are incredibly complex systems, including when it comes to the dozens of industrial control systems found in each plant. There is no typical design or standard for nuclear power plant IT infrastructure or for operating it; instead, each is a unique custom-made engineering project. This complicates the task of potential attackers and limits the scalability of attacks. At the same time, a lack of standardization can also hamper cyber security.

More than a dozen infections are known to have taken place at nuclear power plants, but most were confined to the periphery IT systems. The infection at Gundremmingen in April 2016, however, made its way to the monitoring system for the plant's fuel rods.[27]

E. Risk reduction and resilience

The experts then turned to potential avenues to reduce risk and improve the resilience of the infrastructures in question.

Risk reduction occurs within the limits imposed by three factors: technology, law and policy, and economics. First, the risk-reduction measure needs to be technically feasible. Second, laws and policies could set out positive cyber-security-related obligations or incentives, although they might also have negative side effects (see Chapter 2(b) above on how certifying medical devices could affect cyber security). Finally, it needs to be affordable: the cost of the security measure will be known, while the cost of the potential harmful impact that the security measure could prevent is highly uncertain. Its seriousness, cost and likelihood are estimates, and its impact may not only be economic (e.g. loss of human life). Some experts also felt that operators might have too narrow of a focus on the direct impact of service delivery failure on their customer base, without fully considering the possible broader or cascading impact on society as a whole.

One expert estimated that 80% of threats might be defeated with basic security measures, but coming up with the resources needed for the remaining 20% was difficult. It was suggested that, while it was difficult to assess likelihood (see above), it was nevertheless important to consider preparing for a worst-case scenario, such as an attack disabling or damaging critical infrastructure in a way that endangered the whole nation. Especially worrisome would be several cyber attacks targeting different critical infrastructures at the same time. One expert thought that there might be a greater risk of such an attack during armed conflicts than in peacetime, as enemies could seek to create widespread effects over a range of sectors with cascading consequences.

Many experts underscored the critical nature of cyber hygiene to overall cyber security, including staff compliance with cyber security protocols. Several structural measures to reduce risks were suggested, in particular:

- Complete network isolation (air gap, with no exceptions, not even for VPNs). While this cannot prevent supply chain attacks, it will affect access and make the enemy's situational awareness and transfer rate of command and control data very low.
- Increasing and simplifying the operator's view (by graphical display) of traffic exchanged between components on the internal networks to help detect anomalous behaviour.

Other measures were suggested, including the use of redundant sensors to increase data reliability; diversifying components; and periodically restoring and reloading configurations. The main

[27] S. Gallagher, "German nuclear plant's fuel rod system swarming with old malware", *Ars Technica*, 27 April 2016.

disadvantage of these measures is that they may significantly increase operational and maintenance costs. One expert noted that the next generation of security analysis would use new technology, such as big data analysis and machine learning algorithms, to develop new anomaly detection tools.

As in the previous session, some experts also underscored the importance of cyber security industry standards, whether self-imposed or required by States or other regulators. Various international standards were mentioned (such as 2017 ISO/IEC 27019 on information security controls for the energy utility industry).[28]

F. Incident notification and response

Several experts described industry-produced cyber security incident reports and the extent to which they reflected reality.

Some of the experts noted that in the U.S., for example, the requirements were limited. Most incidents are self-reported on a voluntary basis by utilities to their regulators, which then share them with the Department of Homeland Security. But the utilities themselves decide what is actually considered a cyber incident. They may not know the exact nature of the incident, and may therefore designate an actual cyber incident as non-cyber. Even for confirmed cyber incidents, the reporting requirements are ambiguous.

The need for national authorities to know about significant attacks in due time was emphasized. Their primary worry is of being unaware of the attack, which means they cannot spread the information so that national operators can take measures to prevent its proliferation. The 2017 WannaCry attack was mentioned as a good case of information sharing. The notifications did not come from companies at the national level: the relevant technical information came through international cooperation mechanisms. Generally speaking, incident reports from public sources may also include pieces of information that help form the overall picture, and they may help to correlate incidents.

One expert described Russia's development and deployment of a nationwide detection network known as GosSOPKA.[29] Composed of hardware systems installed directly in the facilities, the system is designed to detect irregular patterns that might be computer attacks and send this information to the central data processing unit at the national coordination centre. Another expert noted that in other countries, such automatic systems might have to be limited to public companies, though specific requirements might be imposed on private companies, such as contracting a qualified detection service provider. Several experts however recalled that the 2016 European Union NIS Directive required operators of essential services to report incidents that had a significant impact on the continuity of the essential services.[30]

Finally, one expert highlighted that, however good national cybersecurity agencies might be, they could not deal with all crises. There is a need for effective collaboration and the proper allocation of resources. This expert gave the example of the certification process in France, where qualified private operators could respond and directly help the victims.

[28] See International Organization for Standardization (ISO), ISO/IEC 27019:2017, *Information security controls for the energy utility industry*, ISO, 2017.
[29] D. Turovsky, "Moscow's cyber-defense: How the Russian government plans to protect the country from the coming cyberwar", *Meduza*, 19 July 2017.
[30] *EU Directive 2016/1148*, 6 July 2016, Art. 14(3).

Session 4: Cyber attacks on the internet core or that may have other systemic effects

A. Cyber attacks[31] on DNS servers

There was agreement among the experts that significant interference with the global Domain Name System (DNS), which is composed of hundreds of geographically distributed servers, was unlikely. The system is indeed inherently redundant. Even in the case of a successful attack, the internet as a whole would still function. The DNS translates hostnames into IP addresses (for example, the DNS converts the name www.icrc.org to its corresponding IP address). As such, if the DNS suffered prolonged downtime, dependent systems such as email and the World Wide Web, or services whose aim is to provide reliable time (i.e. network time services), would be affected. That said, data would still be sent and received between end points (i.e. knowledge of the ICRC site's IP address would be sufficient to connect). While the internet as a whole would still function, DNS interference would seriously hamper users' ability to communicate and access information easily.

To better assess their likelihood, the experts discussed potential motivations behind such attacks.[32] One expert mentioned that the reasons could include the ability to send spam or hijack internet traffic. Subverting DNS systems[33] could force users to connect to servers they had not intended to connect to, which could allow the attacker to obtain traffic that it would not otherwise be able to get. Such attacks could also facilitate malware infection campaigns, for example when a website to which the user is redirected hosts malicious binaries. Participants also mentioned the Border Gateway Protocol (BGP) rerouting incidents, where traffic was surreptitiously made to transit via unusual paths (such as through cyber infrastructure located in specific countries). In such cases, there is a risk of the traffic's content being tampered with.

However, according to the experts, these risks seemed increasingly limited because most of the internet traffic had become encrypted and protected by the Transport Layer Security protocol. Furthermore, the system of digital certificates and modern browsers would make such attempts visible to the user: web browsers have started to inform users about problems with the digital certificates of visited sites.

After the expert meeting took place, threat actors have been found to apply operational techniques enabling the DNS system to be hijacked and user credentials to be stolen. Though not a global attack, it did demonstrate the feasibility of affecting a large number of organizations.[34]

B. Distributed Denial of Service (DDoS) attacks

The experts then discussed DDoS attacks and botnets. One expert shared a paradoxical observation, noting that botnet operators that controlled millions of machines appeared unprepared to actually use them to their full extent. This reluctance might be explained by the undesired visibility caused by massive attacks. This may even result in abandoning the botnet. Botnet owners also segment bot networks to use them selectively (i.e. not all at once). This allows the botnet owners to decide how many bots to engage to perform a task.

DDoS attacks can be conducted on multiple levels and in several ways. Of particular note are concerted attacks with continually changing targets, as in the multi-week DDoS campaign in Estonia (2007). DDoS activity can use several techniques at the same time. Attacks can take place at the levels of the network layer (i.e. attacking internet infrastructure by sending a large number of data packets) and the

[31] On the notion of "cyber attack" as used in this report, see note 2 above.
[32] Some of the attacks discussed need not affect the actual root servers (which is highly unlikely in practice) and may well apply to regional or even local DNS downtime. Details are always case specific.
[33] In other words, causing the user system to connect to the wrong end points, either by poisoning the DNS cache or by having malware installed in the user system.
[34] M. Hirani, S. Jones and B. Read, "Global DNS Hijacking Campaign: DNS Record Manipulation at Scale", Threat Research, FireEye Blog, 9 January 2019.

application layer (targeting specific service or website components, possibly based on their design). Filtering bogus traffic is the primary protection against DDoS attacks. However, filtering can be challenging in complex cases, such as when network spoofing is used (see below). The common denominator of all DDoS attacks is that the attacker chooses the mid points, the end points and the targets, and whether to switch between targets.

One expert illustrated the ability of attackers to direct DDoS attacks against specific targets, citing the concerted DDoS attacks on specific banks in the U.S.[35] Owing to how the internet is designed and operates, there will always be the risk of other systems or networks being affected. The reason is that volumetric attacks can end up using all available channel capacity, which would affect the networks on the path to the target. This has happened in the past, with powerful DDoS attacks launched against specific targets and affecting entire internet infrastructure providers.[36] Furthermore, attackers may have a limited understanding of the potential consequences of their actions, since they may have incomplete information about the network layout and interconnections. This means that the full consequences may in fact not always be accurately assessed before the attack. One expert said that DDoS was always a targeted attack, although its use implied that the attackers accepted the risk of accidentally causing a broader impact. Another expert noted that if an attacker wanted to take down a specific site hosted at a cloud provider through a DDoS attack, the only way to achieve this objective could end up affecting the whole cloud service infrastructure. Indeed, in order to be successful, the DDoS attack would potentially need to consume all of the available bandwidth that enables the cloud infrastructure to be accessed (see below for the discussion on other types of attacks on cloud providers).

The experts also discussed IP address spoofing, which refers to sending packets of data with a false source IP address. This is an important feature of many network layer DDoS attacks. When using IP spoofing, the apparent source IP addresses are not the real ones, so detecting the attacker is more difficult. It is indeed challenging to filter such data packets at the destination of the attack (this is called ingress filtering and is done by the ISP that controls the network infrastructure targeted by the DDoS attack). The most effective approach would be preventive measures (egress filtering), where internet service providers prevent data packets with spoofed IP addresses from leaving their networks. But this method of collective precaution is not broadly deployed by internet service providers.[37] One expert estimated that over 70% of ISPs do not do such filtering.

C. Attacks against cloud service providers

Experts generally agreed that cloud providers typically offered high security standards because they were centralized and able to devote resources to security and reliability. One expert noted, however, that, although ambitious, a successful attack on a cloud infrastructure would be of high value for the attacker, as it would effectively allow them to attack a large number of cloud users.

One expert remarked that in the case of distributed cloud infrastructure, the attacker would have to focus on the common elements of the cloud infrastructure that could potentially be attacked simultaneously.

Another risk could be supply chain attacks or, alternatively, insider threats, where a well-resourced attacker places insiders at the cloud services provider. One expert remarked that just a few well-placed insiders could have a big impact. Such operations would have significant effects, since cloud providers provide services to a large number of people. Other experts disagreed, noting that it was unlikely that a lone insider would be able to cause systemic effects throughout the cloud due to internal procedures, security layers, and limited access rights. One expert also noted that it could be simpler to bribe someone at the target organization rather than go after the cloud provider.

[35] The concerted DDoS campaign against banks in 2012, sometimes referred to as Operation Ababil.
[36] For example, the 2016 DDoS by Mirai botnet, which affected systems operated by DNS provider Dyn. It affected many websites, including those of the BBC and The New York Times.
[37] Among other reasons, egress filtering has a costly performance impact and offers no evident advantage to the ISP.

D. Practical results of attacking internet services and their dependencies

Finally, the experts focused on the impact of cyber attacks on internet services and their dependencies, namely systems that depended on other, internet-based systems in order to function (such as cloud-hosted systems). While attacking internet services would not directly cause physical effects, there might still be significant repercussions on the delivery of essential services. One expert remarked that the inability to access important software, such as when it was stored in cloud infrastructure, could potentially impair service provision, in cases where such a dependency existed. The consequences could be serious, such as in the case of emergency ambulance services.[38] The internal systems of certain essential service providers might be at risk if they are designed to function only if connectivity is available. Losing internet access could also affect the ability to monitor patients who use health monitoring devices (see Chapter 2(c) above). It is therefore important to design systems in a fail-safe way, so that they will function in such cases.

Even if it were possible, the experts deemed it rather unlikely that the entire cloud system provider would be compromised. But if important services depended on components placed in the cloud infrastructure, and that infrastructure experienced downtime, the service in question would be affected. One expert compared this to supply chain issues: if an important part of the system was affected, the whole system would be impacted.

[38] See for example D. Volz, "Hackers disrupt Baltimore's emergency call system; Atlanta still affected", *Reuters*, 28 March 2018; for another recent illustration of the dependency of emergency services on data centres and their related vulnerability (though apparently without any link to a cyber attack), see C. Cimpanu, "CenturyLink outage takes down several 911 emergency services across the US: Downtime caused by network issue affecting 15 of CenturyLink's data centers", *ZDNet*, 28 December 2018.

Session 5: Cyber operations during armed conflict

A. Peace time, armed conflicts and grey zones

Some experts noted that some States had already crossed the proverbial Rubicon: they had adopted cyber operations as a means of statecraft. This reality was here to stay, there was no turning back.

States, including their military forces and intelligence agencies, use cyber operations not only during armed conflicts, but also – and primarily – outside armed conflicts. Some experts noted that States carried out cyber operations in what they perceive to be grey zones, namely situations below the threshold of armed conflict, but in a more aggressive manner than in their regular peacetime relations. One of these experts deemed that the peace/war dichotomy was confusing in view of the nature of offensive cyber operations. The example of the repeated cyber attacks against Ukraine's power grid was offered as an illustration: part of the challenge was to identify the aim, which in this case might have been to spread fear or undermine the confidence of the population in the government. The issue at stake was how to constrain the misuse of cyber capabilities in such grey zones, which, in the expert's view, was not adequately addressed by international law. Other experts objected, noting that there were distinct legal regimes for each situation: international law applicable during peacetime regulates any use of cyber operations outside of armed conflict, including in times of tension, while IHL regulated them during armed conflict. Just as for any other new technology, new questions can arise about specific legal rules or principles (e.g. with regard to IHL, how to distinguish oneself as a combatant, and the notion of "object"), but these experts did not deem them insurmountable (for more on the application of IHL to cyber operations, see Session 6 below).

Some experts noted that they did not expect an armed conflict to be waged exclusively in cyber space or through cyber means. The notion of cyber war, which was understood to refer to such a hypothetical situation, was therefore considered unhelpful and a misnomer, as no armed conflict has ever remained confined to the domain in which it began. One expert expressed the view that an armed conflict could however be initiated through cyber operations. Assuming a State was able to cause an impact similar to that of a kinetic attack, which this expert deemed difficult but possible, it could constitute a use of force under the United Nations Charter. Another expert noted that the notions of use of force, armed attack and aggression found in the UN Charter[39] and in the United Nations General Assembly Resolution 3314[40] did not refer to cyber means or operations. While some States did provide examples of cyber operations that they would consider a use of force, such as in the U.S. Department of Defense Law of War Manual,[41] there is no internationally agreed interpretation thereof. Therefore, this expert recommended that these notions be updated or interpreted and adapted to the specificities of the cyber domain. It was noted that the ICRC uses the notion of cyber warfare – not cyber war – to refer to the use of cyber operations as means and methods of warfare during an armed conflict[42] in the same way that land warfare, air warfare and naval warfare refer to hostilities conducted in these domains.

B. Cyber space as an operational domain of a predominantly civilian nature

Many States and their military forces treat cyber space as an operational domain, like the land, naval, air or outer space domains. But unlike these natural domains, cyber space is entirely man made. The experts described it as constantly changing in a hyper-dynamic manner: every device that was plugged into the internet or the cyber space changed this domain. Alterations to cyber space also take place on

[39] UN Charter, Arts 2(4) and 51, respectively.
[40] UN General Assembly, Resolution 3314 (XXIX), *Definition of Aggression*, 14 December 1974, Annex.
[41] U.S. Department of Defense, *Law of War Manual*, Office of General Counsel, June 2015 (updated December 2016). See Chapter XVI – Cyber Operations, and in particular § 16.3.1.
[42] "The ICRC understands 'cyber warfare' as operations against a computer or a computer system through a data stream, when used as means and methods of warfare in the context of an armed conflict, as defined under IHL. Cyber warfare can be resorted to as part of an armed conflict that is otherwise waged through kinetic operations. The notion of cyber warfare might also encompass the employment of cyber means in the absence of kinetic operations when their use amounts to an armed conflict, although no State is known to have publicly qualified an actual hostile cyber operation as such", in ICRC, *International Humanitarian Law and the Challenges of Contemporary Armed Conflicts*, ICRC, Geneva, 32IC/15/11, 2015, p. 39.

a constant basis by cyber security updates and by changes to its physical or logical architecture, for example.

Another specific characteristic of cyber space is that it is probably a 90% civilian built and owned infrastructure. This raises legal challenges, and one expert wondered what limits there were, or should be, on military manoeuvring in such a domain. Experts noted in this regard that domestic law varied on the authority of States to require civilian companies under their jurisdiction to establish lawful access (possibly even backdoors) to the IT or otherwise connected devices they built.

It was pointed out that belligerents tailored the means and methods of warfare they used to the specificities of the domain. One expert noted that because of the hyper-dynamic nature of the environment and the enemy's ability to quickly patch vulnerabilities, gaining and maintaining access to information systems and manoeuvring through them were highly dependent on one's camouflage, which led to the use of cover and concealment. Matters of operational security like camouflage were noted to be standard practices in traditional military operations and an operational imperative in cyberspace. However, operating in this manner in a predominantly civilian domain was deemed to create tensions with the requirements of taking passive precautions and distinguishing oneself.[43] Furthermore, while ruses such as camouflage are not prohibited by IHL, the death, injury or capture of an adversary by perfidious cyber operations (such as pretending to be a protected civilian) is prohibited.[44]

Another specificity of the cyber domain is that it crosses over traditional State boundaries. For example, botnets might rely on connected nodes on a global scale, and a belligerent having to defend against a botnet would need to take this into account.

C. Vulnerability disclosure, secrecy and deterrence

The experts offered diverse perspectives on the disclosure of vulnerabilities. They generally considered that disclosing vulnerabilities to the vendor to enable it to patch them should be the preferred option. It was noted that the recommendations that the UN Group of Governmental Experts offered for consideration by States for voluntary and non-binding norms included encouraging the responsible reporting of vulnerabilities.[45]

While recalling that 0-day exploits were only one aspect of cyber operations, some experts emphasized that there could be national security reasons not to disclose vulnerabilities. For example, a vulnerability discovered in enemy weapons or communication systems represents a very powerful piece of information for planning purposes. Furthermore, States are not prohibited by international law from engaging in espionage, an act that can be enabled by exploiting vulnerabilities.

Other experts underlined the risks entailed by a decision not to disclose vulnerabilities in view of the specific characteristics of cyber space. Even if a State is exploiting a vulnerability solely for espionage purposes, until the vulnerability is patched it remains available to other States or actors to exploit it, potentially for more harmful purposes. It is generally unknown whether others have already developed the tools to exploit it. Furthermore, exploits could be reused or repurposed after having been used or leaked.

---

[43] See Art. 58 AP I and Art. 44(3) (first sentence) AP I.

[44] See Art. 37 AP I.

[45] "...[T]he present Group offers the following recommendations for consideration by States for voluntary, non-binding norms, rules or principles of responsible behaviour of States aimed at promoting an open, secure, stable, accessible and peaceful ICT environment... (j) States should encourage responsible reporting of ICT vulnerabilities and share associated information on available remedies to such vulnerabilities to limit and possibly eliminate potential threats to ICTs and ICT-dependent infrastructure", in United Nations, *Report of the Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security*, United Nations, A/70/174, 22 July 2015, para. 13(j). In December 2018, the UN General Assembly "Welcome[d] the...international rules, norms and principles of responsible behaviour of State, enshrined in the reports of the Group of Governmental Experts...", in United Nations, *Developments in the field of information and telecommunications in the context of international security*, Resolution adopted by the General Assembly on 5 December 2018 (A/RES/73/27, OP 1).

States' positions vary on the necessity of disclosing vulnerabilities, and some have put in place equity processes to balance such competing interests and risks and decide whether to disclose a vulnerability that has been found.[46] One expert noted that the decision to exploit a 0-day vulnerability could and should consider the risks of repurposing. Different manners of exploiting a vulnerability might exist, some more discreet than others, and they might entail different risks of the exploit being discovered, reused or repurposed. Another expert objected that any operation carried out in cyber space might leave traces, which meant that exploiting a vulnerability always entailed some risk of the exploit being discovered, including by potentially unrestrained or non-law-abiding actors.

While experts agreed that the primary responsibility for the harm caused by repurposing the exploit fell squarely on the actor using it, the question was raised about whether the actor that had developed the original exploit retained some residual responsibility. A suggested course of action was to consider treating used exploits in a manner similar to explosive remnants of war.[47]

One expert feared that ignorance of the cyber capabilities developed by a potential enemy could lead an actor to build up its own capabilities to face potentially inexistent threats. This expert deemed it destabilizing for cyber space and fundamentally different from conventional weapons, while another expert recalled that overestimating enemy capabilities did occur in domains other than cyber. Another expert suggested that a serious conversation might be required in the future in terms of developing arms control through policies of mutual vulnerability disclosure – although the expert acknowledged that currently no State using a vulnerability would be willing to share such information. More generally, this expert felt that the secrecy surrounding the development of cyber capabilities hindered the public debate and the legal and ethical discussions that should take place.

One expert dwelled on the deterrent aspect of cyber operations. Many offensive cyber tools are "one-shot", and cyber capabilities, including knowledge of vulnerabilities and of exploits, are highly classified. Therefore, their development cannot offer the deterrent effect that other weapons might. That said, in their public statements about the development of offensive cyber capabilities, States have sought to have some deterrent effect. However, the expert noted that cyber deterrence could only be achieved by mastering the full spectrum of cyber means, and that cyber resilience was a key factor in this regard.

D. Cyber operations as means and methods of warfare: circumstances of use, aim and expected effects

The experts emphasized that the means and methods of warfare that militaries chose to use depended on the aim and effects sought. Militaries can often assess and monitor the effectiveness of the cyber operation in relation to the effect sought. For example, a belligerent might observe that the targeted enemy's communications have been severed, and assess the success of the operation on that basis.

As noted above, cyber operations are often used for espionage, which as such is not prohibited by international law.

Cyber tools can also be used in information operations. One expert recalled that in the 1991 Gulf war, the US used leaflets as a means of informing Iraqi forces on the battlefield that they would be treated as non-hostile if they fulfilled certain conditions. This prevented physical combat engagement with some of these enemy forces. It was underscored that in warfare, the typical aim was the submission of the enemy, but that physical harm was not the only way to achieve it; breaking the enemy's will to fight could be achieved by psychological means and information operations. Cyber space and the toolset that it offers broaden the choice of means and methods available to carry out information operations, although it might require gaining unauthorized access to the enemy's network to deliver the message. Cyber tools might also be used in psychological operations directed at the civilian population, for example through social media.

[46] See Australian Signals Directorate, *Responsible Release Principles for Cyber Security Vulnerabilities*, Australian Signals Directorate, 2019; UK GCHQ, *The Equities Process*, UK GCHQ, 29 November 2018; United States Government, *Vulnerabilities Equities Policy and Process for the United States Government*, United States Government, 15 November 2017.

[47] See the Protocol on Explosive Remnants of War (Protocol V to the 1980 CCW Convention), 28 November 2003.

The use of cyber means and methods of warfare might also be aimed at deceiving and obfuscating the enemy, creating denial effects or other effects that do not directly cause physical harm to the enemy but that provide the belligerent with a military advantage. For example, if a belligerent can access and spoof (or corrupt) the data that the enemy uses to track the position of its own or allied forces, such data alteration could undermine the overall confidence of the enemy. By increasing the fog of war, it would undermine the enemy's capabilities beyond the possibly limited amount of data that would actually be altered. One expert gave the example of a belligerent knowing that there was a close communication loop between enemy forces such that if subordinate forces did not receive orders from the hierarchy to move troops, they would not do that on their own. This expert also recalled events of interference with drone communication. In such cases, severing the communication link would create a denial effect through cyber means, without causing physical harm.

States and non-State armed groups are increasingly using the internet as the backbone to support command and control communications. The experts shared the view that anything that used computational processes to function, including weapons or weapons systems, was vulnerable to being disabled if the enemy managed to get access to these systems. One incentive to the development of cyber capabilities was therefore to find ways to disable enemy weapons systems. Experts recalled that, conversely, a primary aim of cyber operations, including offensive ones, was to secure the belligerent's own systems and foil or defend against enemy cyber operations.

Finally, cyber operations might be used in support of kinetic operations. One expert recalled the reports that Israel used cyber means to shut down the Syrian integrated air defence system when it carried out air attacks to destroy what Israel suspected to be the construction site of a nuclear reactor in 2008.[48] The operations reportedly interfered with the data inside the air defence system network, although it was not clear how that might have been achieved.

One expert viewed cyber operations as a logical extension of electronic warfare. Cyber operations offer more options than electronic warfare because they are not geographically constrained, as electronic warfare is, although this expert did not expect them to entirely replace electronic warfare operations.

More generally, it was deemed that from a military point of view, the effectiveness of cyber means during armed conflicts depended on their integration with other conventional capabilities.

Several experts underscored an important restraint on choosing to use cyber operations to create effects, namely the risk of compromising future intelligence collection. Once a belligerent has established unauthorized access to an enemy computer for intelligence gathering purposes, using this access to achieve an effect might lead the enemy to discover it. The enemy could then try to remedy, block or patch the vulnerability, preventing future intelligence gathering using the exploit developed for that vulnerability. It was noted that intelligence agencies and military forces might approach such decisions from different perspectives.

E. Potential military cyber operations that take advantage of the medical condition of an enemy.[49]

The experts discussed the likelihood, legality and feasibility of military forces considering tampering with a biomedical device of an enemy commander in order to kill him or her during a conflict.[50] One expert noted that this would require a legal analysis of whether being in need of medical care rendered the target a "protected person" by being sick or *hors de combat* when not taking part in hostilities, or more specifically when in hospital. This expert deemed that the answers to these questions would be nuanced in practice. With regard to feasibility, one expert noted that it was mostly a matter of capability and the willingness of the party considering such a course of action to dedicate sufficient resources to it. Experts recalled that as far back as 20 years ago U.S. Vice President Dick Cheney had his defibrillator

[48] See Part 3(b) in the background document contained in Annex 3.
[49] This discussion took place during Session 2 but was included here as it discusses a potential military use of cyber operations.
[50] M. N. Schmitt (ed.), *Tallinn Manual 2.0 on International Law Applicable to Cyber Operations*, 2nd ed., Cambridge University Press, Cambridge, 2017, Commentary on Rule 104, para. 6.

removed and replaced with one that had no wireless capability owing to fears it could be used in an assassination attempt.[51]

One expert suggested that knowledge of an enemy commander's medical condition could be a piece of information leveraged to mount a kinetic operation rather than a cyber attack that caused harmful effects directly. For example, knowledge of an enemy commander's medical appointment might help locate him in order to capture or kill him on the way to or back from the medical facility. Another expert considered that hard to reconcile with the obligation to respect and protect the medical mission, especially if the information was obtained by hacking into the medical or administrative records of a medical facility. This could unduly impede the facility's medical functioning and hinder the ability of health-care professionals to uphold their ethical duty of preserving medical confidentiality.

F. Cyber operations and expected incidental civilian harm

One expert offered the view that military commanders might feel uncomfortable with the difficulty of anticipating, with a sufficient degree of confidence, the incidental civilian harm expected to be caused by a cyber operation. Commanders are increasingly used to the degree of scientific sophistication reached by collateral damage estimate methodologies currently used by militaries. However, similar methods do not exist yet for cyber operations. Another expert held that not all incidental effects would necessarily amount to legally relevant incidental civilian harm. It was also noted that the acceptable level of incidental harm and the required precautions to be taken to avoid such harm could differ depending on the actor or the type of conflict. Commanders might be more inclined to accept incidental civilian harm in armed conflicts waged for national survival than in less intense hostilities.

Several experts noted, however, that a potential trigger for carrying out cyber operations – and possibly even their development – could actually be to avoid civilian harm. As discussed above, notably with regard to information and denial effect operations, cyber operations might make it possible to achieve military advantages without necessarily causing physical harm. This was deemed important when discussing the potential human cost of cyber operations. History illustrates this trend: for example, graphite bombs were developed to cause short-lived dysfunction instead of long-lasting damage to the targeted electricity grids. In the same way, using cyber means to disable infrastructure that has become a military objective could help put it back in service more quickly than if it was destroyed by kinetic fire. This could even be the case with regard to physical damage directly caused by the cyber operation, which, depending on the situation, could be more precisely targeted and tailored than kinetic operations. Some experts considered that the pressure to use cyber operations to avoid incidental civilian harm could become increasingly powerful over time, and one expert even wondered at what point it could become a requirement.

[51] R. Luscombe, "Dick Cheney feared assassination by shock to implanted heart defibrillator", *The Guardian*, 19 October 2013.

Session 6: The protection afforded by existing law, and possible avenues to reduce the human cost of cyber operations

During this session, experts looked at the protection afforded by existing law against the potential human cost of cyber operations, especially during armed conflicts, and started exploring some potential avenues that could offer added value from a technical, legal, policy or other perspective to avoid or at least reduce this potential human cost.

The experts emphasized that international law applied to cyber operations, and that IHL – including the principles of distinction, precaution, proportionality, military necessity and humanity – regulated the use of cyber means during armed conflict. This was deemed important because various States employ cyber tools in support of kinetic operations in contemporary armed conflicts.[52] While agreeing on the applicability of IHL in principle, the experts debated a number of issues. Several experts underlined the need for further discussion to clarify how IHL applied to cyber operations during armed conflict, and specifically how its rules were to be interpreted, and to assess whether they were adequate and sufficient or whether new rules or a new treaty were needed.

A. Conflict classification and questions of attribution

The questions of whether a cyber operation could trigger the application of IHL and whether the conflict is international or non-international in nature depend on a number of factors. Those factors include the actors involved – which in the cyber realm encompass not only States but also individuals and non-State groups – and the effects caused by a cyber operation. Such effects may include the corruption of computer systems, the leaking of data, or human casualties, and the experts emphasized that many hostile cyber operations did not amount to, or take place, in the context of an armed conflict. It was also recalled that cyber operations could be conducted with varying intentions, ranging from espionage to sabotage, disruption and destruction. While some experts considered the political motivation and the recognition of a state of "war" by States to be important, others noted that under IHL the determination of whether an armed conflict exists was independent of any political recognition by any of the actors involved, and did not require a "declaration of war". The determination of whether a conflict exists needs to be done based on the nature of the operations taking place.

One issue that could complicate the classification of a cyber operation is that attribution of a hostile operation to a particular actor can be difficult in cyber space, especially if multiple chains of proxies are used. While international law provides rules on the attribution of wrongful conduct to States, it can be complicated to prove the required links. Experts recalled that questions of attribution were particularly relevant if a cyber operation was conducted outside the context of an armed conflict: unless the operation could be attributed to a State, it was difficult to hold another State responsible or to consider an operation a violation of the prohibition to use force under the UN Charter. At the same time, one expert emphasized that the difficulty for the target of a cyber attack[53] or an external observer to attribute an operation to a State did not preclude the applicability of international law to this State operation, as that State would know what it was doing. Moreover, it was recalled that, during armed conflict, attribution of acts to another party to the conflict might not be essential: during ongoing hostilities, parties to the conflict may only attack lawful targets, and this requires a determination that is not necessarily related to questions of attribution.

[52] One expert recalled the example of one State reportedly applying cyber means to disable the air defence of another State before launching airstrikes; see the text in relation to note 186.
[53] On the notion of "cyber attack" as used in this report, see note 2 above.

B. The notion of "attack"

During an armed conflict, a number of rules regulate "attacks" as this notion is defined for the IHL rules governing the conduct of hostilities.[54] As reflected in the *Tallinn Manual 2.0 on International law applicable to cyber operations*, there is broad recognition that if a cyber operation results in effects comparable to those of kinetic military operations – such as if, as a result of the operation, the targeted device or some of its components must be replaced for the object to function again – such cyber operations amount to attacks under IHL and are governed by the related rules.[55] Debate continues, however, whether other operations, in particular those that result in a loss of functionality that can be restored without physically replacing some of the device components, amount to an attack as defined in IHL. One expert emphasized that the State from which that expert came considered a cyber operation during an armed conflict as amounting to an attack under IHL when the targeted systems or infrastructures no longer rendered the services they would normally provide. The expert noted that this approach was different from what could be found in the Tallinn Manual and focused on the availability of services. It is also distinct from the question of whether the results of an operation can be equated to a use of force under the UN Charter.

C. Challenges in anticipating the effects of cyber attacks

Key principles of IHL require parties to armed conflict to anticipate the effects of an attack. For instance, the principle of proportionality requires the belligerents to balance the concrete and direct military advantage anticipated against the expected harm to civilians or civilian objects.[56] Likewise, if they have a choice between several targets, parties to armed conflicts must select targets which may be expected to cause the least danger to civilian lives and objects.[57] One expert noted that some States that used cyber capabilities during conflicts had developed specific rules of engagement to govern such use. One question that is analysed during the targeting process is whether the use of cyber capabilities could minimize incidental civilian harm compared to other means or methods of warfare, such as kinetic ones.[58]

A number of experts emphasized that uncertainty about the effects of an operation was an inherent feature of cyber operations. Whereas some cyber tools are operated by humans who will be in a position to halt the operation if unexpected effects occur, other tools – such as self-propagating malware (i.e. worms) – may be designed to operate autonomously and to replicate themselves. Self-replication is a specific functionality that must be deliberately included in the malware. In this context, one expert explained that the operator normally would not retain control over such self-replicating malware, which could have negative effects beyond the system it originally targeted or beyond what the user expected.[59] Even malware that is operated by humans can get out of control if the link between the human and the malware is cut, for instance by attempts of the attacked system to disable the malware. Experts further emphasized that because autonomous types of malware followed algorithms, there was a risk that they could cause unpredictable effects. Moreover, as cyber tools consist primarily of software, there is always a risk of having errors in the code, which can cause unforeseen and unintended effects.

While some experts suggested that that made cyber tools significantly more unpredictable and prone to unintended effects than kinetic weapons, another expert cautioned that conventional weapons had error rates too, that human error could cause unintended effects during kinetic attacks, that environmental constraints, such as weather conditions, affected the accuracy of conventional weapons, and that the more conventional weapons relied on technology, the more they were subject to technical errors. Other experts opined, however, that contrary to conventional weapons, it was not possible to define an error rate for cyber weapons, or the equivalent of a "Circular Error Probability", because the

[54] See, for instance, the prohibition of attacks against civilians (Art. 51(2) AP I), the prohibition of indiscriminate attacks (Art. 51(4) and (5) API), the prohibition of attacks against civilian objects (Art. 52 AP I), or rules on precautions in attack (Art. 57(2) and (3) API).
[55] *Tallinn Manual 2.0*, Commentary on Rule 92, paras 10-12 (see note 50 above).
[56] See Art. 51(5)(b) AP I.
[57] See Art. 57(3) AP I.
[58] See Art. 57(2)(a)(ii) AP I.
[59] One example cited was the malware Stuxnet, which spread well beyond the original target, apparently beyond what the authors had expected, and despite the constraints they had included in the malware to limit propagation.

likelihood and possible effects of an error in the software could not be calculated and the actual effects of a malware became apparent only once it was released. Therefore, a number of experts observed that it was rather complicated to test malware and identify all possible effects before using it. The question was raised whether, against this background, self-replicating malware could be considered indiscriminate and, therefore, unacceptable.

D. The persistence of malware once released

A related issue is that self-replicating or self-propagating malware is likely to continue spreading and causing harm after the initial attack it was created for ended. This can also be the case if the malware is programmed to stop at some point: experts recalled that a worm known as Slammer, developed in 2003, and the Stuxnet worm were still spreading online, consuming power and polluting the cyber environment. While not all worms can be identified across networks, experts reported that initiatives to clean infected computers had been taken. In one State, the national cybersecurity centre constantly alerts internet service providers of users whose PC is infected, asking the service provider to contact the individual user to address the issue by applying a software update. However, internet service providers and individual users do not always follow up on these alerts. In another State, the law enforcement agencies cooperated with technical experts to develop a program to identify computers that were part of a botnet and provide the individual user with the software needed to address the issue. While this approach could work at the national level, having such a programme operate across borders might raise challenges for territorial sovereignty, making international cooperation a prerequisite for success.

E. Potential avenues to reduce or avoid human harm

In international law, rules exist on the responsibility of States for internationally wrongful acts, which apply to acts committed in cyberspace. States also have a responsibility to prevent certain acts committed by non-State actors within their jurisdiction, which may include cyber crime. While these rules provide a basis for international responsibility in case human harm is unlawfully caused in cyber space, one expert cautioned that their application was made difficult because it was not always easy to attribute harmful conduct in cyber space.

When discussing cyber attacks that could affect the provision of health care and attacks against industrial control systems, the experts had already discussed various avenues that could be taken, mostly in terms of enhancing the cyber security posture and resilience of the actors that were potentially affected (see Chapters 2(e) and 3(e), respectively, above). During this last session, the experts presented a number of other ideas to alleviate the risk of human harm caused by cyber operations and discussed some of them in more detail.

Avenues to explore at the legal and policy level

Some experts proposed possible short and long-term measures to help prevent potential human harm. For the long term, some experts suggested that an international convention either banning the development and use of cyber means of warfare or regulating their use be negotiated. One suggestion was to negotiate a fourth additional protocol to the Geneva Conventions, focused on regulating cyber warfare and building on existing law. Acknowledging that any such negotiations would be complicated and take time, one expert suggested that, in the short term, some form of cyber peacekeeping could be established by the United Nations. Cyber peacekeeping could focus on confidence-building measures and norm development to help maintain peace in cyber space, help minimize the impact of cyber operations on civilians if cyber attacks occur, and help rebuild cyber defence and damaged cyber infrastructure following an attack.

Other suggestions were made with regard to industry regulation, such as:

- improving transparency so that customers would know how long a product's cyber security was maintained and updated
- reinforcing supply-chain security
- responsibly managing products' end-of-life.

The commitments undertaken through the Paris Call for Trust and Security in Cyberspace, launched on 12 November 2018,[60] were also mentioned. Finally, the debates and concerns raised by the possibility that private actors would be legally entitled to respond to cyber operations (through "active defence", hack back, etc.) were recalled. One expert described the effort carried out by one State to better frame and constrain what private actors could do.

Avenues to explore at the technical level
The experts focussed particular attention on the ideas of segregating military and civilian infrastructure; tagging malware to avoid escalation; developing a digital watermark to identify certain actors and objects that were protected against attack; and measures to avoid the repurposing of malware.

Demilitarized zones in cyber space
One expert, reflecting on the idea of creating specific areas in cyberspace that would be demilitarized and used only for civilian purposes, opined that that would not be feasible because the military and civilians used the same infrastructure and everything was connected. However, it is already possible to identify and distinguish between civilian and military users or endpoints, although one expert underscored that it could be a challenge to affect one without affecting the other.

Digital marker for operations solely designed for exploitation
Experts further discussed the idea of developing a digital marker to distinguish a malware designed for espionage purposes or other CNE from a malware that would disrupt or destroy computer networks. The rationale for having such marking would be to help the attacked party identify what type of threat it is facing and to calibrate its response accordingly. This could help to avoid an escalation of retaliatory measures. Some experts opined that creating such a digital marker or signature would be technically feasible and could be designed in a way that would not make the malware easily identifiable. The marking could, for instance, be included in the payload and only identifiable by the defender once the malware is discovered. Other experts, however, disagreed on that point, questioning the likelihood that attackers would use such a marker, precisely because it would make the attacker easier to identify. In their view, such a marker would render an attack ineffective from the start because the malware would be identified and rendered useless. Moreover, it was cautioned that such a sign could be used by cybercriminals to disguise an attack that was, in fact, designed to destroy computer networks. One expert also cautioned that malware initially designed for espionage purposes could be easily repurposed for other reasons, making this type of marking unfeasible.

Digital marker for protected objects
The experts also discussed the idea of developing a digital watermark to identify certain actors (for instance, civil defence organizations and their assets) or infrastructure (hospitals and critical infrastructure) in cyber space in order to prevent them from being attacked. Such watermarks could either show that those actors or objects enjoy specific protection under international law (comparable to the red cross, red crescent, and red crystal emblems, or markings for cultural property, goods containing dangerous forces, or civil defence organizations), or simply that they are civilian objects that must not be attacked. Technical suggestions on how to implement this included having a neutral organization create an internationally recognized register of digital assets that were protected. These protected assets could be identified by their IP addresses and/or by including a digital mark in their software, such as a digital signature or serial numbers, to ensure it was in the register. Such assets could include systems, computers and devices, for example. It would of course be necessary to create a well-standardized approach, most likely under the auspices of a neutral international body. While the experts noted that such a marking would not provide comprehensive protection or be immune from abuse, the main rationale for developing such a system would be to make it easier for States and other actors that aimed to avoid civilian harm to effectively exclude the marked objects from attacks. For instance, worms and other autonomous malware could be programmed to spare objects that are marked

[60] "Cybersecurity: Paris Call of 12 November 2018 for Trust and Security in Cyberspace", Ministry for Europe and Foreign Affairs, France, 12 November 2018.

in a certain way. Such markings would not change much for human operators: those actors will normally know whom they are targeting and where they are on the network.

The experts emphasized that in order to determine the usefulness of such an approach, the potentially positive effects of protecting certain actors and objects from unintended harm by law-abiding actors would need to be balanced against disclosing the location or information such as the IP addresses of critical infrastructure to potential adversaries, including criminals. One expert doubted that the overall balance would be beneficial because, compared to criminals, States usually had better capabilities, and in armed conflict they were already under an obligation to verify the nature of the target. Another expert deemed that it could have a positive effect only if attribution became easier. In fact, experts emphasized that States kept lists of critical infrastructure confidential in order not to disclose what infrastructure they considered essential. In their view, if an actor actually sought to target such objects, a disclosure or water-mark would make such an attack easier.

A registry of malware used for legitimate purposes

One expert suggested establishing a list or register of malware used by States for legitimate purposes, such as anti-terrorism operations. This list could be managed by a neutral intermediary. Once internet security companies discover malware, they would check with the neutral intermediary to see if it is a permissible tool on that list. If it is, they would not disclose it. If it is not on the list, they would apply standard procedures, including informing all relevant actors of the exploit and malware, and taking steps to address the threat. That expert noted that if such a list of tools and actors existed, it would also be possible to identify which State was responsible for releasing a tool that was later misused. Other experts, while appreciating the idea, opined that such a registry was not feasible because it would require States to disclose which tools they possessed, which in reality is highly classified information. Another expert cautioned that such a system might not be feasible for cyber security companies, which had contractual obligations to defend and protect clients.

Limiting the proliferation of malware

The experts expressed concern at the proliferation of cyber weapons and the relative ease of using them to carry out cyber attacks. They deemed that that situation was fuelled by the amount of resources channelled into the development of cyber weapons, the fact that many of them had already fallen into the public domain and others could be stolen or leaked, and the fact that tools originally designed for espionage could be repurposed to cause harmful effects to the targeted system.

Several experts raised concerns about States releasing malware to non-state actors, or releasing them without going through the appropriate processes to allow the owner or vendor to patch the relevant exploits. Some experts underscored that, regarding State responsibility under international law, a State using leaked or stolen tools was responsible for its own conduct. Experts wondered, however, about the responsibility of the State leaking or releasing malware that was subsequently used for harmful acts by non-State actors.

Experts deemed it important to find means to incentivize States to reduce the proliferation of cyber weapons or reinforce protections against the weaponization of exploits. Indeed, knowledge of exploits makes it easy for actors to weaponize them and cause harm. It was emphasized that States normally kept the exploits they identify highly confidential because once released, it would be unlikely that they could use them in the future. There have, however, been examples of States' exploits or malware being made public, including to facilitate attacks by other actors. Several experts countered that making vulnerabilities and exploits public was desirable, and was normally done to allow IT vendors to test and improve the level of security. While vendors can usually fix the vulnerability that an exploit uses, this does not necessarily render the exploit entirely ineffective: many users do not update their systems and would therefore not benefit from a patch.

In order to avoid exploits or malware getting into the hands of cyber criminals, one suggestion was to think about how to improve attribution capacities in order to identify which State released the exploit and to hold that State liable for possible damage.

Prevent the repurposing of malware

The experts agreed that it was hardly possible technically to develop a malware that could not be repurposed at all. Malware developers can, however, include significant obstacles in a malware in order to prevent repurposing, or at least to make it very hard and expensive. One expert recalled research conducted in 2012 on Gauss malware, whose method of encrypting the payload had not yet been broken.[61] That said, once the malware is out and gets into the hands of other experts, they will be able to examine it, understand it, and also repurpose significant parts. Nonetheless, some experts emphasized that encryption, in particular of the payload of a malware, and including obstacles in different components of a code, was important because it raised the bar in terms of the expertise required to reengineer malicious tools, or parts of the tool, and therefore prevented at least some actors from doing so. This is equally true with regard to preventing a self-replication function from being added when repurposing an initially narrowly targeted cyber tool.

[61] "The Mystery of the Encrypted Gauss Payload", Kaspersky Lab, 14 August 2012.

Annex 1: Agenda

Summary

14 November 2018	
8:10 – 8:30	Registration and coffee at the ICRC's Humanitarium
8:30 – 9:30	Welcome; objective and scope of the meeting
9:30 – 12:30	**Session 1: Cyber operations in practice**
12:30 – 14:00	*Lunch*
14:00 – 17:45	**Session 2: Evolution and future outlook of cyber attacks against, or that may affect, the delivery of health care (including those affecting medical data, medical devices and hospitals)**

15 November 2018	
8:30 – 12:30	**Session 3: Cyber attacks against critical civilian infrastructure, or that may otherwise affect the delivery of essential services to the civilian population**
12:30 – 14:00	*Lunch*
14:00 – 15:30	**Session 4: Cyber attacks on the internet core or that may have other systemic effects**
16:00 – 17:45	**Session 5: Cyber operations during armed conflict**

16 November 2018	
8:30 – 12:15	**Session 6: The protection that existing law affords against the potential human cost of cyber operations, especially during armed conflicts, and the potential avenues that could offer added value from a technical, legal, policy or other perspective to reduce or avoid this potential human cost**
12:15 – 12:30	**Closing Remarks**

Detailed Agenda with guiding questions

14 November 2018

8:10 – 8:30 Registration and coffee at the ICRC's Humanitarium

8:30 – 9:30 Welcome; objective and scope of the meeting
Introductory presentation on the main notions of international humanitarian law

Mr Dominique Loye and Mr Laurent Gisel

9:30 – 10:30 Session 1: Cyber operations in practice

Short introductory presentations by Mr Serge Droz, Mr Thomas Dullien, Mr Mark ("Magpie") Graham, Mr Vitaly Kamluk and Ms Ella Yu, followed by discussion among the whole group of experts.

10:30 – 11:00 Coffee break

11:00 – 12:30 Session 1 (continued)

Session 1 Guiding Questions:

1. What are the different stages of cyber operations and the challenges they raise?

In particular:
a) What are the resources and time needed to plan, prepare and run various types of cyber operations?
b) When preparing a cyber operation, what are the technical options for assessing whether it can be targeted and executed precisely, and the risk that it might cause collateral damage or that the operator will lose control? What are the limits of such assessments?
c) What categories of actors are there, what are their goals and the type of targets, and to what extent do these factors shape, affect and relate to the points raised in (a) and (b)?

2. How are malware and exploits developed?

In particular:
a) What are the resources and time needed to develop, adapt or operationalize exploits and malware? When is repurposing or reengineering done (as opposed to developing new tools)?
b) What types of exploits and malware are in use in practice? Are they tailored for a specific operation or generic?
c) What is the relationship between sophistication (required time/resources) and the effect of the malware, including the ability to control and limit such effects?
d) What are the options and means of controlling and limiting malware propagation to avoid incidental or indiscriminate effects? Can the risk of incidental or accidental harm be anticipated or tested in advance?
e) Why are self-propagating malware (worms) events comparatively rare?

3. How are cyber operations likely to further develop and evolve?

In particular:
a) How is exploit and malware development changing, and what is its future? Will exploitation remain possible in 10 to 20 years? Is it possible to estimate this with any useful degree of certainty?
b) What are the consequences of high visibility offensive cyber operations (e.g. Stuxnet, WannaCry and NotPetya) on the cyber security landscape? Do the cyber attacks reported publicly represent the bulk of the serious malware in use and vulnerabilities in systems, or are there others? If there are others, how does that influence the threat analysis?
c) Are the risks of high-impact cyber operations increasing (more, new or emerging actors, more capabilities) or decreasing (more resources for cyber security, better security posture)?

12:30 – 14:00 Lunch

14:00 – 15:30 Session 2: Evolution and future outlook of cyber attacks against, or that may affect, the delivery of health care (including those affecting medical data, medical devices and hospitals)

Short introductory presentations by Mr Franck Calcavecchia, Mr Bruce Eshaya-Chauvin and Dr Marie Moe, followed by discussion among the whole group of experts.

15:30 – 16:00 Coffee break

16:00 – 17:45 Session 2 (continued)

Session 2 Guiding Questions:

1. What are the different dimensions of the risks of cyber attacks against or affecting the health-care sector?

In particular:
a) What are the types and likelihood of (i) cyber attacks affecting data availability and integrity, (ii) cyber attacks affecting medical devices, or (iii) other types of attacks that make it difficult or impossible for hospitals to function?
b) What are the risks that cyber systems used in the health-care sector will be incidentally affected by cyber attacks against other targets? What are the factors that increase or decrease such risks?
c) What are the risks that malware developed or used for operations directed at targets other than the health-care sector will be repurposed or reengineered to attack, or will otherwise incidentally affect, health-care infrastructure? What are the factors that increase or decrease such risks, and how could such risks be avoided or the consequences be mitigated?

2. What is the resilience of health-care systems in the face of cyber attacks, and what are the potential consequences of such attacks on the delivery of health care?

In particular:
a) What is the ability of health-care providers to anticipate, defend against and recover from these attacks?
b) What is the ability of a health care provider (e.g. a hospital) affected by a cyber attack to continue delivering medical services to patients? What does it depend on in particular (e.g. the type of attacks, the type and sophistication of the infrastructure being attacked, the overall quality of the infrastructure of the country where the attack takes place, or other factors)?
c) What are the potential consequences of the various types of cyber attacks against the health-care system, and could they lead to a deterioration in patients' health or even to their death? On what factors does the seriousness of the potential consequences depend?
d) What would be the consequences of an entire hospital IT system going down simultaneously?
    - What if this affects hospitals in an entire city, region or country? Or for a long period? Is it technically possible and likely? How would that affect the delivery of health care?

3. How are cyber operations against or affecting the health-care sector expected to evolve in the future?

In particular:
a) What are the worst-case scenarios? What are the most probable scenarios?
b) To what extent could the security and resilience of cyber systems supporting health-care delivery be improved, and what are the technical, operational, financial, commercial or other limits to doing so?

15 November 2018

8:30 – 10:00 Session 3: Cyber attacks against critical civilian infrastructure, including those that may otherwise affect the delivery of essential services to the civilian population

Short introductory presentations by Mr Sergio Caltagirone, Mr Oleg Demidov, Dr Giovanna Dondossola, Ms Marina Krotofil and Ms Susan Lee, followed by a discussion among the whole group of experts.

10:00 – 10:30 Coffee break

10:30 – 12:30 Session 3 (continued)

Session 3 Guiding Questions:

1. **What are the risks of cyber attacks on critical civilian infrastructure other than the health-care sector?**

In particular:

a) What are the potential types, circumstances and likelihood of cyber attacks against, or affecting energy, water, transportation, logistics, dams, nuclear plants or the chemical and biological industry?
b) What are the differences between industrial control system (ICS) malware and non-ICS malware, specifically in terms of operation, design, development and use? What are the differences between cyber attacks on ICS information technology (IT) systems and operational technology (OT) layers?
c) What is the risk of automated cyber attacks on ICS (possibly self-propagating)? Considering the different ICS set-ups between and among sectors and between countries, can automated attacks occur across various sectors (e.g. the energy, water and chemical sectors) or across various countries?
d) What are the security posture and vulnerabilities of critical infrastructure supporting the delivery of essential services to the civilian population? For example, to what extent are air gaps and network segmentation (part of) a solution?

2. **What are the potential consequences of cyber attacks on critical civilian infrastructure in terms of death, injury or destruction and in terms of the delivery of essential services to the civilian population?**

In particular:

a) What are the potential consequences of cyber attacks against, or affecting, energy, water, transportation, logistics, dams, nuclear power plants or the chemical and biological industry?
    - To what extent can the providers of essential civilian services continue delivering their services to the population if they have been affected by cyber attacks, and what are the risks raised by cyber attacks (e.g. blackouts, no access to water)?
    - How resilient are industrial systems? What factors affect their recovery time?
    - What is the risk that cyber attacks will lead to physical destruction or human injury/death (e.g. through an industrial plant explosion) or the release of dangerous forces (e.g. water from dams, or hazardous substances, such as chemical or radioactive material, from industrial plants)?
    - What are the factors that influence such consequences, notably with regard to the type of attacks, the type and sophistication of the infrastructure being attacked, the overall quality of the infrastructure of the country where the attack takes place, the existence of backup analogue systems and the interdependency of critical civilian infrastructure.
b) Can such consequences be caused only by intentional attacks on specific facilities, or could they occur:
    - As an incidental or accidental result of cyber attacks against other targets?
    - On random targets, possibly simultaneously?
c) What is the risk that tools and malware developed for other types of operations will be repurposed or reengineered to attack, or will otherwise incidentally affect, critical civilian infrastructure? What are the factors that increase or decrease such risks, and how could such risks be avoided or the consequences mitigated?

3. How are cyber operations against or affecting critical civilian infrastructure expected to evolve?

In particular:
a) How have recent cyber attacks on ICS evolved?
b) What are the worst-case scenarios? What are the most probable scenarios?

12:30 – 14:00 Lunch

14:00 – 15:30 Session 4: Cyber attacks on the internet core, or those that may have other systemic effects
Short introductory presentation by Mr Monnappa K A, followed by a discussion among the whole group of experts.

Session 4 Guiding Questions:

1. What are the risk and potential consequences of cyber attacks on the internet critical core components?

In particular:
a) What are the risks and consequences of significant (i.e. temporary, prolonged or irreversible) cyber attacks against DNS root servers, the trust system (i.e. Certificate Authorities) and key cloud provider(s)?
b) How resilient are the current internet core components and their elements? Is there a real risk of a significant negative impact?

2. What are the risks and potential consequences of cyber attacks on other important systems that might cause systemic effects?

In particular:
a) What are the risks and potential consequences of cyber attacks against or affecting financial systems such as SWIFT, and of cyber attacks affecting banks more generally?
b) Are there any other core internet components or other systems, the attack or disruption of which would lead to negative systemic effects?

3. What is the risk of cyber attacks incidentally affecting the internet core or having systemic effects?

In particular:
a) What is the risk of the internet core being incidentally or accidentally affected by cyber attacks against other targets?
b) What are the risks that malware developed or used for operations directed at targets other than the internet core will be repurposed or reengineered to attack, or will otherwise incidentally affect, the internet core or will have other systemic effects? What are the factors that increase or decrease such risks, and how could such risks be avoided or the consequences mitigated?

15:30 – 16:00 Coffee break

16:00 – 17:45 Session 5: Cyber operations during armed conflict
Short introductory presentations by Col. Gary Corn and Mr Ewan Lawson followed by a discussion among the whole group of experts.

Session 5 Guiding Questions:

1. What is the current situation with regard to the use of cyber operations during armed conflicts?

In particular:
a) Under what circumstances have cyber operations been used by parties to armed conflicts, and with what aims, methods, military impact and humanitarian consequences?
b) Under what circumstances has the use of cyber operations during armed conflict been considered but decided against, and why?
c) What is the state of play with regard to States' current doctrines and policies on the use of cyber capabilities during armed conflicts?

2. How are cyber military capabilities and their use expected to evolve?

16 November 2018

8:30 – 10:00 **Session 6: The protection afforded by existing law against the potential human cost of cyber operations, especially during armed conflicts, and the potential avenues that could offer added value from a technical, legal, policy or other perspective to reduce or avoid this potential human cost**

Short introductory presentations by Dr Victor Cambazard, Dr Manmohan Chaturvedi, Capt. Anne Laubacher and Dr Longdi Xu, followed by a discussion among the whole group of experts.

10:00 – 10:30 *Coffee break*

10:30 – 12:15 **Session 6: (continued)**

Session 6 Guiding Questions

1. **What protections should the law afford to address the potential human cost of cyber operations, and to what extent does existing law already afford such protections, especially during armed conflicts?**

In particular:

a) Considering the technical characteristics of cyber space and cyber operations, how can we assess the relevance and adequacy of the main principles of international humanitarian law (IHL)? Notably:
   - Are the IHL principles of distinction, proportionality and precaution, along with the special protection regimes, useful in practice to protect civilians and civilian objects from the effects of cyber operations during armed conflicts? How is this analysis affected by the fact that exploits and malware can be repurposed or reengineered, and by the fact that exploits and malware may be developed by private (non-governmental) actors, including individuals?
   - To what extent can the effects, including incidental effects, of malware, exploits and other cyber capabilities be adequately tested or otherwise properly anticipated?
   - What specific aspects of, or tools used during, cyber operations could be understood as weapons, means and methods of warfare?

b) What is the potential to address the challenges that cyber warfare raises through the application and interpretation of existing IHL? Is there a need to clarify and/or further develop IHL?

2. **What avenues could be pursued in the technical, legal, policy or other realms to work towards reducing the potential human cost of cyber operations?**

In particular:

a) Keeping in mind the various suggestions that have been put forward in recent years with regard to new norms or rules on cyber, or to technical developments, what is their added value in comparison to the existing situation, and what is their potential from a technical or practical perspective?

b) Are there technical proposals to which the international community should pay more attention, and, if so, which ones?

c) Does the specific nature of cyber, from a technical or other perspective, call for innovative avenues to be developed, and if so which ones?

12:15 – 12:30 **Closing remarks**

Annex 2: List of experts

Invited experts

- **Mr Franck Calcavecchia,**[1] Information Security Officer, University Hospital of Geneva, Switzerland
- **Mr Sergio Caltagirone**, Director, Cyber Threat Intelligence, Dragos, United States
- **Dr Victor Cambazard**, French National Cybersecurity Agency, France
- **Dr Manmohan Chaturvedi**, Cybersecurity Project Consultant, IIT Delhi, India
- **Col Gary Corn**, Staff Judge Advocate, Cyber Command, United States
- **Mr Oleg Demidov**,[2] Consultant, PIR Center, Russia
- **Dr Giovanna Dondossola**, Leading Scientist at the Department Transmission and Distribution Technologies of the Research Center for Energy Systems, RSE SpA - Research on Energy System, Italy
- **Dr Serge Droz**,[1] Senior Advisor ICT4Peace, Vice-Chair Forum of Incident Response Security Teams (FIRST), Vice-President OS-CERT, Open Systems, Switzerland
- **Mr Thomas Dullien**,[3] Computer Security Researcher, Google Project Zero, Switzerland
- **Dr Bruce Eshaya-Chauvin**,[1] Global Health Institute, University of Geneva, Switzerland
- **Mr Mark ("Magpie") Graham**, Senior Threat Intelligence Analyst, Microsoft, United Kingdom
- **Mr Vitaly Kamluk**, Principal Security Researcher, Kaspersky, Singapore
- **Ms Marina Krotofil**, Senior Technical ICS Security Consultant, Admeritia, Germany
- **Mr Monnappa K A**, Information Security Investigator, Cisco CSIRT, India
- **Capt. Anne Laubacher**, Legal Advisor, COMCYBER, Ministry of the Armed Forces, France
- **Mr Ewan Lawson,**[1] Senior Research Fellow, Royal United Services Institute, United Kingdom
- **Ms Susan Lee**, National Security Analyst, Johns Hopkins University Applied Physics Laboratory, United States
- **Dr Marie Moe**, Research Manager at SINTEF and Associate Professor at NTNU, Norway
- **Dr Longdi Xu**, Research Fellow, China Institute of International Studies (CIIS), China
- **Ms Ella Yu**, Team leader of 360 Advanced Threat Response Team, Qihoo 360, China

ICRC

- **Mr Dominque Loye**,[3] Deputy Director of International Law & Policy
- **Dr Knut Dörmann**,[4] Chief Legal Officer and Head of the Legal Division
- **Dr Lindsey Cameron**, Head of the Thematic Unit, Legal Division
- **Mr Laurent Gisel**, Senior Legal Adviser, Legal Division
- **Ms Netta Goussac**,[5] Legal Adviser, Legal Division
- **Dr Lukasz Olejnik**, Scientific Adviser on cyber, Legal Division
- **Dr Tilman Rodenhauser**, Legal Adviser, Legal Division
- **Mr Alain Ceccato**, Chief Information Security Officer
- **Ms Delphine Van Solinge**, Digital Threats Adviser, Protection Division
- **Mr Georges Baize,** Adviser, Unit for Dialogue with Weapon Bearers
- **Dr Sasha Radin,** Managing Editor, Humanitarian Law and Policy Blog, Department of International Law and Policy
- **Ms Sophie Huve**, Associate, Legal Division
- **Mr Guillem Puri Plana,** Associate, Legal Division

[1] *present on 14–15 November*
[2] *present on 14 (afternoon) and 15 November*
[3] *present on 14 November*
[4] *present on 15–16 November*
[5] *present on 16 November*

Annex 3: Background document

This background paper was prepared by Laurent Gisel, senior legal adviser, and Lukasz Olejnik, scientific adviser on cyber, both of whom work in the ICRC's Legal Division. While it aims to provide relevant material to support the discussions at the expert meeting, it does not necessarily represent the ICRC's official positions.

Table of Contents

Introduction

The evolving nature of cyber operations over the past decade is a pressing concern. Cyber operations can damage objects, disrupt essential civilian infrastructure (including the power grid and hospitals), and otherwise affect institutions and businesses. Cyber attacks, defence and security are constantly evolving, and novel tools, techniques and methods are being developed. High-profile malware, although not a routine occurrence, shows the state of play in this area, how far these capabilities have come, and how sophisticated the actors behind them are.

The expert meeting will focus on some of the specific risks raised by cyber operations, in particular the risk of death, injury or physical damage, and the risk that the civilian population will be cut off from essential services. In this background document, these consequences are referred to as the potential human cost of cyber operations.

The meeting aims to shed light on the potential human cost of cyber operations through a better understanding of the technical issues surrounding these operations. Participants will discuss how both malware and vulnerabilities have changed, IT security, the current and future threat landscape, and the extent to which essential infrastructure relies on cyber systems. The approach is fundamentally multidisciplinary.

In view of the rapid change in cyber operations, both cyber security policies and cyber military capabilities have been developed, and normative developments are under discussion. Given this complexity, a clear understanding of the characteristics of cyber operations and their potential human cost is required. Some of the pressing questions are:

- Are there any technical means or processes that would reduce or even eliminate the human cost of cyber operations?
- Can cyber tools such as malware be designed and used so that their effects are limited to a particular target?
- What are the risks of incidental or accidental damage, and how can they be reduced or avoided?
- Can cyber operations ultimately be controlled and monitored technically?
- What are the risks of cyber tools being repurposed or reengineered, and how can these risks be addressed?
- To what extent can essential service providers continue to deliver services despite being affected by a cyber attack?
- How are the answers to these questions likely to change in the next 10 or 20 years?

Cyber operations raise many other issues and challenges that will remain out of the scope of the expert meeting. In particular, the meeting will not address cyber espionage, intellectual property theft or privacy concerns. It will not look at the use of cyber means as part of information operations – such as by leaking hacked information or through the use of social media or other cyber means for propaganda or disinformation purposes. It will not consider the economic cost of harmful cyber operations borne by businesses and governments, except to the extent that the cost of cyber security is a factor that affects the vulnerabilities and resilience of cyber systems – and therefore the potential human cost of cyber operations.

Part 1 of this background document provides an overview of the ways in which cyber operations are carried out, while Part 2 analyses specific cyber operations that have affected the delivery of health care and other essential services to the population. In Part 3, this document reviews the development and use of military cyber capabilities and the protection afforded by international humanitarian law (IHL). Part 4 contains a summary overview of possible avenues to reduce or avoid the human cost of cyber operations.

1. Cyber operations – an overview

a) Cyber attacks and other types of cyber operations

Governments, media and businesses regularly report that their websites or networks have been subjected to cyber attacks. Indeed, "cyber attacks" and "cyber operations" are widely used terms, yet there is no common definition or understanding of these concepts.

The terms "cyber operations" and "cyber attacks" may be used in a legal sense under IHL (see Part 3 below). In this background document, "cyber operations" and "cyber attacks" will instead be used in an operational sense. Building on definitions adopted by States, "cyber operations" will refer to the employment of cyber capabilities to achieve objectives in or through cyber space.[62] The notion of "cyber operations" therefore encompasses the notion of "cyber attacks", which are understood to be operations that use cyber capabilities (i.e. computers, software and data streams) to disrupt, deny, degrade or destroy information resident in computers and computer networks, or the computers and networks themselves, and/or to affect the physical systems – or more rarely the individuals – that rely on such computers and networks.[63]

These notions cover both the tools and the way they are used. Consequently, cyber operations and cyber attacks should not be understood exclusively as tools or methods, but as a combination of the tools and the methods (or technique) with which they are employed.

The notion of cyber operations includes reconnaissance and espionage activity (sometimes referred to as computer network exploitation), as well as disruption and destruction. In many cases, the tools employed in these diverse operations do not differ significantly on a technical level, although they may be modified and combined with various methods. Furthermore, espionage may be the goal of a cyber operation, it may be part of its groundwork, or it may take place during the initial steps of such an attack – or all three together.

b) Cyber tools - some examples

Offensive cyber activities often employ specialized tools. They can include repurposed or reengineered tools (some of which may have been publicly available), newly acquired tools and custom-developed tools. From a cybersecurity point of view, these tools are called malicious software (malware). They can be designed to facilitate data exfiltration or to disrupt or destroy a target.

The history of malware development spans more than 30 years. The first self-propagating code that resulted in disruption was the Morris worm.[64] Malware has been continuously developed since then.

Recent advanced malware has enabled operations with substantial disruptive or even destructive effects. But there is a huge difference between the first Morris worm, later malware such as Conficker and Shamoon (which targeted IT systems but also affected industrial systems), and tools like Stuxnet and Triton. The Morris worm spread as the result of a mistake; Shamoon resulted in significant disruption; and Stuxnet and Triton, which required significant resources, were specifically designed to attack industrial control systems (ICS).

A selection of recent high-profile tools is listed below. These tools differ in their technical sophistication, their capabilities and even their potential uses. Some of them made use of zero-day (0-day) exploits,[65] others used known vulnerabilities. They targeted IT and/or industrial control systems. Some were automated and self-propagating, while others required human intervention. While there are reports that all these tools (and others) were developed or used by States or State-supported actors, no States have acknowledged as much, and some expressly deny any responsibility.

These summaries include both technical elements and the actual impact.

[62] See U.S. Department of Defense, *Dictionary of Military and Associated Terms*, U.S. Department of Defense, September 2018, Washington D.C., pp. 59–60. This publication provides definitions of cyberspace operations and cyberspace attacks. Cyberspace operations are defined as "the employment of cyberspace capabilities where the primary purpose is to achieve objectives in or through cyberspace", and cyberspace attacks as "[a]ctions taken in cyberspace that create noticeable denial effects (i.e., degradation, disruption, or destruction) in cyberspace or manipulation that leads to denial that appears in a physical domain, and is considered a form of fires".
[63] Such as in the case of an internet-connected pacemaker, or via secondary effects of attacks on physical infrastructure.
[64] See J. Reynolds, "The Helminthiasis of the Internet", Network Working Group, RFC 1135, December 1989.
[65] A 0-day vulnerability is a vulnerability that the party that would be interested in fixing is unaware of.

Stuxnet
Stuxnet is a malware designed to spread in the industrial networks of a chosen site and cause equipment malfunction. It resulted in physical damage to uranium-enrichment centrifuges in Iran.[66] This tool was remarkably different from the most sophisticated malware known at the time, and it used four 0-day exploits. Once the tools were equipped with the capability to cause physical damage, the operation was automated.[67]

Flame
Flame is a malware[68] that was used against targets in the Middle East as part of espionage activities.[69] Its capabilities included taking screenshots of a computer screen, logging keystrokes and recording audio. Flame was capable of spreading and infecting computers automatically. The design team had advanced cryptography skills, as evidenced by the fact that a previously unknown method was used to create a fraudulent digital certificate.

BlackEnergy
BlackEnergy is a trojan horse that, once installed, provides the attacker with remote access.[70] It is a versatile tool: it has been used to form botnets capable of launching DDoS attacks – or to steal financial data. It later became more advanced and even more versatile, with a modular design in which different components had different functions. The third iteration of the malware was used in a prominent attack on a Ukrainian power grid in 2015, where the malware enabled remote human control over the grid by the attacker.[71]

WannaCry
WannaCry[72] is a ransomware[73] that was used in 2017. It was a worm capable of propagating automatically among Windows operating systems using an exploit called EternalBlue – which had reportedly been stolen from a State agency.[74] This attack affected systems in more than 150 countries, rendering many of the systems unusable.

NotPetya
NotPetya is a wiper[75] used in 2017 and designed to cause disruption. Like WannaCry, NotPetya used an exploit called EternalBlue to propagate automatically, yet it also had other, enhanced spreading techniques. When NotPetya infected a computer, it would overwrite the master boot record; when the system restarted, NotPetya would then encrypt the disk and render the system unusable. It was used as a ransomworm and caused significant damage around the world.[76]

---

[66] N. Falliere, L.O. Murchu and E. Chien, *W32.Stuxnet Dossier Version 1.4*, Symantec Security Response, February 2011.
[67] Cyber operations may begin with a reconnaissance operation and then cause physical damage at a later stage. So Stuxnet may have been used initially to facilitate reconnaissance before being updated to deliver a destructive payload. See *W32.Stuxnet Dossier Version 1.4*, p. 4 (see note 66 above): "These design documents may have been stolen by an insider or even retrieved by an early version of Stuxnet or other malicious binary. Once attackers had the design documents and potential knowledge of the computing environment in the facility, they would develop the latest version of Stuxnet. Each feature of Stuxnet was implemented for a specific reason and for the final goal of potentially sabotaging the ICS."
[68] A. Gostev, "The Flame: questions and answers", Secure List, Kaspersky Lab Blog, 28 May 2012.
[69] K. Zetter, "Meet 'FLAME', the massive spy malware infiltrating Iranian computers", *Wired*, 28 May 2012.
[70] Kaspersky Lab, "BlackEnergy APT attacks in Ukraine", Kaspersky Lab. A trojan horse is malicious (yet possibly disguised as legitimate) software that gives the attacker remote control over the target system. This includes the ability to delete, modify or copy data and execute commands locally.
[71] K. Zetter, "Inside the cunning, unprecedented hack of Ukraine's power grid", *Wired*, 3 March 2016.
[72] "What you need to know about the WannaCry Ransomware", Symantec Security Response Team, Symantec Blog, 23 October 2017; W. Smart, *Lessons learned review of the WannaCry Ransomware Cyber Attack*, UK Department of Health & Social Care, London, 1 February 2018.
[73] Malicious programs blocking access to a computer system or data, with the intention of extorting money from the computer owners in return for restoring access.
[74] S. Biddle, "The NSA Leak is Real, Snowden Documents Confirm", *The Intercept*, 19 August 2016; S. Gallagher, "Hints suggest an insider helped the NSA 'Equation Group' hacking tools leak", *Ars Technica*, 22 August 2016.
[75] A wiper is a form of malware that destroys data on targeted systems. See also Cisco Talos, "New ransomware variant 'Nyetya' compromises systems worldwide", Talos Blog, 27 June 2017.
[76] A ransomware with self-propagating (worm-like) functionality. See also A. Cherepanov and R. Lipovsky, "GreyEnergy: Updated arsenal of one of the most dangerous threat actors", *We Live Security*, 17 October 2018. An

VPNFilter
VPNFilter is a malware primarily used against routers to make them part of a botnet.[77] This malware is capable of quickly infecting a large number of vulnerable nodes. It has a modular design, one component of which can listen to a device's network traffic, while another contains some functions that target industrial control systems (e.g. for data exfiltration). When the malware receives an instruction from the controller, it can execute a destructive functionality to render the infected device unusable.

Triton / Trisis
Triton (also known as Trisis) is a very advanced malware designed for use against industrial systems. It targets Triconex safety instrumented systems (SIS) and is capable of reprogramming them.[78] This malware can disrupt or destroy computer systems and in some circumstances could potentially result in physically disruptive or destructive effects. It could even lead to human injuries or casualties, given the critical nature of safety systems at industrial facilities.

Additionally, certain actions may be carried out using standard (i.e. not malicious) programs available on the targeted system. This was the case, for example, of the cyber operations that led to power cuts in Ukraine in 2015. A subsequent operation in 2016 appeared to be more automated.

Exploits developed for specific vulnerabilities may be referred to informally as "weaponized". This term refers to the development of a cyber tool that relies on an exploit to achieve an effect. For example, Stuxnet weaponized exploits for four vulnerabilities. More recently, the EternalBlue exploit was weaponized in WannaCry and NotPetya.

c) The nature of the cyber tools and methods used in cyber operations

Typically, cyber tools using components that exploit a specific system or software vulnerability will affect only systems that use that software. For example, an exploit for a Linux vulnerability will function only on Linux systems and not, for example, on Windows operating systems.

Cyber operations can be tailored technically to specific targets, such as a country, a facility, a type of system or even individual users. It all depends on the intelligence available on the target, the resources used by the attackers to develop their tools, and the care put into it. To some extent, this tailoring is made possible by the fact that the attackers must create or adapt their cyber tools with a specific goal in mind. For example, they may look at the network level (e.g. IP addresses), the operating system version, or other identifying information.

Certain precautionary measures may be included in the tools' design and functionalities. For example, the code may only execute on the targeted systems, and there may be limits to a given tool's ability to propagate (including the option of stopping it entirely).[79] For example, Stuxnet was tailored specifically to the operating systems used in a particular facility; VPNFilter only worked on network nodes that had certain vulnerabilities; and Triton targeted vulnerabilities found only in Triconex safety instrumented systems.

These tools may be designed with limits in terms of their reach and impact, yet they or their components – such as the exploits used – could be repurposed or reengineered for other purposes without these limitations (see Part (e) below for more details).[80] Attacking industrial systems, such as those used by power grids, which are cyberphysical systems, will normally require a highly tailored

early version of NotPetya called Moonraker Petya, which was not equipped with EternalBlue, was used in December 2016.
[77] Cisco Talos, "New VPNFilter malware targets at least 500K networking devices worldwide", Talos Blog, 23 May 2018.
[78] B. Johnson *et al*. "Attackers deploy new ICS attack framework "TRITON" and cause operational disruption to critical infrastructure", FireEye Blog, 14 December 2017; Dragos Inc., "TRISIS Malware: Analysis of Safety System Targeted Malware", Dragos Inc., 14 December 2017.
[79] A "kill switch" was built into the first version of the WannaCry ransomworm.
[80] Vulnerabilities used in Stuxnet and Mirai have been exploited by other tools as well. This does not mean that broad repurposing or reuse is always possible. For example, Stuxnet and Trisis were designed to be effective on specific systems. Reusing these tools as a whole against other targets would be costly and ineffective.

operation.[81] Industrial control systems (ICS) are often based on different designs, even among facilities within same sector. This means that the tools and methods used to attack a power grid or other complex facility (e.g. a manufacturing plant) will not necessarily work on another facility in the same country or a different country. These important technical differences make it difficult to carry out attacks on multiple facilities.

However, designing and automating attacks is possible in many other cases, in particular in pure IT systems (non-ICS). Unlike specifically tailored operations, exploits of a vulnerability in widely used software or computer systems enable widespread and untargeted (or loosely targeted) attacks, as demonstrated by the history of self-propagating malware. Worms are typically very difficult to contain once released, as they can replicate and spread fast and efficiently, with no respect for borders or system perimeters. And worms are persistent: for example, a 2008 worm called Conficker continued to infect systems in 2018.[82] Their persistence and their destructive capability are also illustrated by the recent WannaCry and NotPetya attacks. NotPetya's reach was unprecedented; it was apparently designed and released in the aim of crippling systems.[83]

d) The kill chain model

The cyber kill chain model is used to describe offensive cyber operations[84] and the various phases of attack, in particular in advanced attack campaigns. Each step in the kill chain is tailored towards a specific goal, both in terms of the nature and use of the tools.

Reconnaissance

In this phase, the attacker identifies a target and gathers data. The information may include the nature of the organization and the organization's structure, systems and potential vulnerabilities. This phase can include simple steps such as finding the email addresses and functions of certain employees, or finding the types of software stacks that are used.

Weaponization

This phase is about crafting tools that will be used later to gain access. These tools may use exploits for specific technical vulnerabilities. The attacker chooses the most efficient way of bundling the tool for subsequent delivery. This might be a malicious PDF file, a website that hosts malware (such as so-called watering hole attacks that target the site-specific audience), or an image, for example.

Delivery

The attacker delivers the tool to the target. There are many ways to do this, but they fall into two main categories: those that are automatic and those that require an action by the victim upon delivery (see Part (f) below).

Exploitation

Exploitation is the phase in which a vulnerability in an application or operating system is leveraged to execute the attacker's code. This phase need not be automatic; it may happen when someone is tricked into performing certain actions, or it may even take place through the use of legitimate software installed on the system. The latter was among the techniques used by the NotPetya wiper, which harnessed the (legitimate) Windows Management Instrumentation.

Installation

After the initial tool has been delivered and the vulnerability exploited, this initial tool may enable the installation of subsequent tools that will enable persistent remote access (e.g. trojans, implants or backdoors) and maintain that persistent access. That said, the specific features will depend on the particular circumstances, such as the malware that has been installed. After this phase, the initial compromise is turned into a more persistent one.

---

[81] Cyberphysical systems consist of hardware and infrastructural components (e.g. power lines and physical or electronic switches) and software (e.g. Windows, Linux, and installed programs).
[82] "Forgotten Conficker worm resurfaces to infect systems with WannaCry", *SC Magazine*, 22 May 2017.
[83] A. Hern, "Ransomware attack 'not designed to make money', researchers claim", *The Guardian*, 28 June 2017; A. Greenberg, "The Untold Story of NotPetya, the most devastating cyberattack in history", *Wired*, 22 August 2018.
[84] "The Cyber Kill Chain", Lockheed Martin; E. Hutchins, M. Cloppert and R. Amin, "Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains", *Leading Issues in Information Warfare & Security Research*, Vol. 1, No. 1, 2011.

Command and Control (C2)
When communication with the attacker's infrastructure is established through the installation of a remote-access tool (Phase 5 above), the intruders can issue commands to the malware installed in the victim's infrastructure. A number of techniques can be used to facilitate C2. Most commonly, remote servers are used to communicate with the tool. Alternatively, C2 channels can be established using removable media – such as a pen drive, in which case commands may only be delivered when the pen drive gets connected.

Actions on Objectives
Once the attackers have obtained full access to the system inside the targeted network, they may perform various actions on the controlled system. For example, in NotPetya, the Actions on Objectives included the crippling of the Windows boot system (master boot record corruption). In the attacks on the power grid operators in Ukraine, Actions on Objectives encompassed the various activities leading to propagation within the system and the execution of commands that eventually resulted in power cuts.

The malware implants installed (phase 5) may perform, or be used to perform, additional actions. For example, they may enable attacks on other systems in order to build persistence, exfiltrate data for use against other targets or be used as a bot in distributed denial of service attacks (DDoS attack), if this was the aim of the operation. This phase may thus either be the final step, or it may be a step towards attacking further targets, either inside or outside the network.

For the attacker to complete the desired final step, it is typically necessary to have covered all seven phases. However, the phases do not always need to be executed in exactly the order described in the kill chain model (for example, the exploitation and installation phases may be repeated after command and control has been established). This underscores the dynamic nature of operations, where objectives may change and attackers may have to take steps to maintain access. This is also why an espionage campaign may potentially transition smoothly into a disruptive or destructive campaign. This is possible, both technically and operationally, because of the ease with which tools may be used, reused and modified, and because of their configurable nature.

e) Malware reengineering

If a tool is stolen, leaked, or recovered after being used, it may be reverse engineered, reengineered, or repurposed to obtain effects different from those intended by the original authors.[85] The relative ease with which this can be done is a direct consequence of the way software works. Computer programs can be easily reverse engineered, decompiled and modified; this can save resources (i.e. money and time), and it can be useful in cases of limited access to in-house expertise. When it comes to malware in particular, computer hackers, criminals, and other non-State actors engage in reuse.[86] State agencies are also known to reengineer malware.[87]
For example, the Zlob Trojan, which was detected in 2008, exploited a critical Windows Shell vulnerability.[88] A similar exploit was later employed in one of the variants of Stuxnet,[89] and Flame malware was later found to be exploiting the same vulnerability. Reuse is not limited to exploits: the malware Duqu is similar to Stuxnet in several respects.[90] While the primary aim of Duqu, information theft, was different from that of Stuxnet, the original tool had apparently been repurposed.

[85] Reengineered malware is malware that has been modified before being used again; repurposed malware is malware that is being used for a reason other than its original one (possibly after being reengineered).
[86] M. Laliberte, "Why hackers reuse malware", Help Net Security, 20 November 2017; A. Dinham, "Hackers 'recycling code' to spread worms", *ZDNet*, 1 June 2014.
[87] N. Schmidle, "The digital vigilantes who hack back", *The New Yorker*, 7 May 2018. This article describes how this approach is used by the U.S. government as one of its standard operating procedures. It quotes Lt. General Vincent R. Stewart (then-deputy commander at the U.S. Cyber Command): "Once we've isolated malware, I want to reengineer it and prep to use it against the same adversary who sought to use it against us".
[88] "Vulnerability in Windows Shell Could Allow Remote Code Execution", Microsoft Security Bulletin MS10-046 - Critical, 2 August 2010.
[89] *W32.Stuxnet Dossier Version 1.4*, p. 4 (see note 66 above).
[90] B. Bencsáth, G. Pék, L. Buttyán and M. Félegyházi, *Duqu: A Stuxnet-like malware found in the wild*, CrySyS Lab, Budapest University of Technology and Economics, Budapest, 14 October 2011.

Perhaps the most visible illustration of how efficiently and rapidly malware tools can be reengineered and employed in cyberattacks happened in 2017, when the group Shadow Brokers made tools public. Regardless of the aim with which these tools had been originally designed, they were subsequently repurposed and used in unprecedented malware campaigns (WannaCry, NotPetya, BadRabbit) that had an impact around the world and caused massive, primarily economic, damage. Although patches for the exploited vulnerabilities existed when the attacks took place, many systems had not been updated. This is one of the reasons why these tools were ultimately so effective.

Reengineering can also be done to test for system vulnerabilities. Penetration testing frameworks, for example, employ techniques commonly used in attacks, in an easy, systematic and repetitive manner, in order to test the system's vulnerabilities. The aim is to ensure an appropriate level of security by identifying and addressing any weaknesses found before a malevolent actor can exploit them.[91] For instance, tools designed to test the same vulnerability that Stuxnet took advantage of are readily available.[92] This shows that reengineering in itself is not necessarily a malicious action.

f) Methods of delivery

To carry out offensive cyber activities, the perpetrator typically must first work towards gaining access to the targeted system. Access can be achieved in many ways of varying levels of sophistication. As noted above, the methods may be automated, or they may require action by the victim.

Among the most efficient methods to gain access to the targeted system is to use a tool that exploits a remote code execution vulnerability in that system.[93] Publicly known vulnerabilities of this sort, which are rare in widely used software nowadays, allow automatic delivery. This type of mechanism was at work recently in the WannaCry and NotPetya worms, which spread by exploiting the remote code execution vulnerability in the infected computers.[94] Malware can also be delivered via scripts; one example of this is the hijacking of advertising infrastructures to target specific websites or audiences by exploiting browser vulnerabilities.[95] In similar attacks, attackers place malicious scripts on compromised websites, in order to target specific visitors.[96]

Another widely used delivery method is phishing (and its variants). This is where the attacker plays on the user's trust, baiting the user to perform a harmful action that will, for example, lead to an initial infection. Delivery methods involving human action include phishing emails with malware attached, links to a malware-hosting website, links to malicious applications in a mobile store, messages with links sent via a social network such as Twitter, messages sent via instant messenger, or simple text messages sent to a phone. Malware-loaded pen drives, placed on or near an organization's premises or distributed during conferences, are another vector. Phishing has been used to attack very serious targets, such as power grids operators, government and military facilities, and even political parties.

As technology develops and new authentication standards are drawn up and adopted, phishing methods may become more difficult to use. In that case, attackers may seek other, simpler means of entry. One topic of concern is supply-chain attacks, where the attackers target vendors (i.e. software and hardware developers and suppliers), to gain access to users later on in time.[97] Under some assessments, the number of supply-chain attacks may be much higher than those noticed or publicized.[98] Supply chain attacks work on the assumption that resourced targets may develop a good security posture but still use third-party software and hardware. Those products may be developed by suppliers with low security standards, and the product might be tampered with to undermine its security. The supply chain was the vector used in the NotPetya campaign to first deliver the malware. Given the nature of the targeted supplier in that case (MEDoc, widely used software in Ukraine), all companies that pay taxes

[91] Metasploit is the world's most commonly used penetration testing framework.
[92] J. Vazquez and M. Heerklotz, "Microsoft Windows Shell SMB LNK Code Execution", Packet Storm, 12 March 2015.
[93] An exploit for a remote code execution vulnerability does not require access to the local system (e.g. accounts). It allows code to be executed from outside the system (i.e. from an arbitrary, remote location).
[94] "Schroedinger's Pet(ya)", Secure List, Kaspersky Lab Blog, 27 June 2017. NotPetya could also detect credentials in the targeted system; this feature was used to spread the tool and attack systems that did not contain a vulnerability.
[95] This malicious exploitation of advertising infrastructure is called malvertising.
[96] This is commonly known as a watering hole attack.
[97] CERT-UK, *Cyber-security risks in the supply chain*, CERT-UK, London, 2015.
[98] Kaspersky Lab, *Kaspersky Lab Threat Predictions for 2018*, Kaspersky Lab, 6 December 2017, p. 8.

in Ukraine could have been affected. In general, attacking the supply chain increases the risk of collateral damage and may even lead to indiscriminate attacks. Attackers may have limited or no control over the actual delivery process, resulting in software or hardware with compromised security standards being delivered to consumers.

g) Air gaps

Network segmentation, which refers to the process of isolating sensitive networks and systems, can protect against many types of attacks. Air gapping, where systems and networks are physically separated from others (e.g. from the internet), is the most effective form of network segmentation. It decreases the chances of information passing in or out of the air-gapped networks. Strong network isolation is, however, a challenge. For instance, systems still need to receive upgrades, and this requires a connection to another network or to some form of removable media (e.g. pen drive) that itself had been plugged in a system connected to another network. Isolating networks through air gapping affects the operational design of cyber attacks that target a particular network or network components, in particular phases 1, 2, 3, 6 and 7 of the kill chain. The intrusion of malicious tools could, however, be initiated with the delivery of infected pen drives; Stuxnet and other malware used this method. An extensive body of research documents the many ways in which information can enter and leave air-gapped systems via out-of-band communication.[99] Past examples show that methods do exist for determined attackers, who may find a way to bypass such security measures. Specialized toolkits are being developed to attack isolated networks, and infections have even occurred at military sites.[100] Intrusions might also be made possible by third-party infections before the isolated network has been established, such as in the case of supply-chain attacks. Despite these limits, isolation through network segmentation remains a strong defensive tactic that offers effective protection against many network attacks, and it is most often used in industrial settings.[101]

h) Targeting industrial control systems (ICS)

The cyber kill chain methodology has two stages when applied to attacks on industrial control systems.[102] The first (I) stage is composed of the seven steps of the standard kill chain describing the attack on a traditional information system (see point (d) above).[103]

Stage II involves the specific activities engaged against the industrial systems' parts. These activities are very different from those used in the traditional non-ICS actions. For instance, they might require an understanding of the nature of the ICS system — including the respective physical process it controls (e.g. at a manufacturing plant). Some steps in Stage II, such as understanding the ICS set-up and developing and testing tools, are both time and resource consuming. They typically take place outside of the targeted systems and networks and they may include the use of information exfiltrated during Stage I, as well as any required information gathered from any other source. Attacks may need to be tested in a local set-up, and this may require access to actual ICS components (e.g. equipment, such as pumps, motors). This increases the cost of ICS attacks. When developed and deployed, malicious tools

[99] This could be done via non-standard measures like thermal channels, using HVAC (heating, ventilation, and air conditioning) systems; similar communication can be crafted with power lines or device electromagnetic emissions and magnetic fields. Although such custom techniques would be difficult to set up, they underscore current challenges. See Y. Mirsky *et al.* *HVACKer: Bridging the Air-Gap by Attacking the Air Conditioning System*, Ben Gurion University, 30 March 2017; M. Guri *et al.* *PowerHammer: Exfiltrating Data from Air-Gapped Computers through Power Lines*, Ben Gurion University, 10 April 2018; M. Guri *et al.* *USBee: Air-Gap Covert-Channel via Electromagnetic Emission from USB*, Ben Gurion University, 10 August 2016; M. Guri *et al.* *ODINI: Escaping Sensitive Data from Faraday-Caged, Air-Gapped Computers via Magnetic Fields*, Ben Gurion University, 8 February 2018. See also D. Goodin, "Scientist-developed malware prototype covertly jumps air gaps using inaudible sound", *Ars Technica*, 2 December 2013.

[100] G. Burton, "WikiLeaks reveals details of CIA's 'Brutal Kangaroo' toolkit for attacking air-gapped networks", *Computing UK*, 23 June 2017; N. Shachtman, "Exclusive: computer virus hits U.S. drone fleet", *Wired*, 7 December 2011; J. Vijayan, "Infected USB drive blamed for '08 military cyber breach", *ComputerWorld*, 25 August 2010.

[101] R. Huber, "Exploring the Air Gap Myth", *SC Magazine*, 26 September 2017.

[102] M. Assante and R. Lee, *The Industrial Control System Cyber Kill Chain*, SANS Institute, October 2015.

[103] These steps are typically needed to gain access to the systems that communicate with ICS infrastructure, and they are based on the assumption that ICS are found in internal networks, which is most often the case. The two-stage attack model is simplified when an ICS is connected directly to the internet; indeed such systems are often not isolated, and this increases the risk of attack. See also L. Pietre-Cambacédes, M. Tritschler and G. Ericsson, "Cybersecurity myths on power control systems: 21 misconceptions and false beliefs", *IEEE Transactions on Power Delivery*, Vol. 26, Issue 1, January 2011, pp. 161–172.

may then interact with the ICS infrastructure, possibly interfering with the physical process. Attacks that target ICSs may result in the disruption or destruction of physical systems.

However non-destructive cyber operations can also unintentionally result in physical effects. The operation may have been designed or intended only for spying, or it may have been the first stage ("access" or "exfiltration") of a more destructive operation. This complicates the victim's task of determining the attacker's intentions whenever an attack leads to physical effects, especially when time is of the essence.

i) Threat actors and their reach

Threat actors can be classified by different traits, such as motivation, intent and capabilities. Considering the resources required to carry out operations that might lead to the threats discussed in this paper, low-scale actors such as insiders, opportunistic attackers and hacktivists do not concern us.[104] These three types of actors typically do not have access to sufficient resources.

Organized crime groups ("cybercriminals") may have various goals, including financial gain. Their capabilities vary; many have adequate resources and access to the right skills. Some of these groups may also offer their services for sale.

Groups capable of advanced persistent threat (APT) attacks are unlike the aforementioned threat actors. In many cases, their intentions can only be identified after analysing the various parts of several cyber operations in a broader context — for example the nature of the targeted systems, or the actions taken against different targets. This type of attacker is able to launch high-impact attacks, which raise concerns in view of the potential human cost. In public reports, APT attacks are often linked to the activities or support of a specific State, though such allegations have been systematically rejected by the States concerned.[105] As noted above, the cyber kill chain model is best- suited for describing APT attacks.

2. The potential human cost of cyber operations

The following sections analyse attacks on specific sectors. They focus on operations affecting the delivery of essential services to the population and other attacks that may cause the type of harm that is the topic of this meeting, namely those that may cause death, injury or destruction, or that may deprive the civilian population of essential services. These sections are concerned in particular with attacks against, or that may affect, the provision of health care, energy, water, transport or logistics; manufacturing; and the internet core.

In general, cyber attacks may affect data confidentiality, availability and integrity.[106] Although data availability and integrity are our primary interests here, the question of confidentiality is particularly prominent in typical threat intelligence reports. As a result, the numbers indicated in this part may sometimes also include confidentiality, even if this issue is outside of our primary scope of interest.

a) Attacks against, or that may affect, the provision of health care

Hospitals and other health-care providers host a wide range of computing systems, ranging from individual computers to full-blown systems used to store and process employee and patient data or manage schedules (e.g. surgery) and diagnostic devices (e.g. USG, MRI, X-ray and CT).

Cyber attacks and malware infections in the health-care sector are on the rise. For example, the number of reported cyber attacks in the United States increased by 63% between 2015 and 2016 (93 major

[104] However, it is important to note that there are often few constraints preventing insiders with direct access to control systems from causing damage, either directly or by handing over insight to others. Most organizations are unable to establish meaningful safeguards to prevent such insiders from doing so, and they have to trust their employees.

[105] N. Fraser *et al.* "APT38: Details on New North Korean Regime-Backed Threat Group", FireEye Blog, 3 October 2018; "ESET unmasks 'GREYENERGY' cyber-espionage group", ESET, October 2018.

[106] Attacks on data confidentiality are those designed to steal data; attacks on availability are those designed to make the data or systems unavailable to the user (this was the case of the WannaCry ransomware, but that operation also included a DDoS attack); attacks on data integrity are those designed to modify data.

attacks in total),[107] while breach incidents have reportedly increased by 10% per year since 2010.[108] Attacks on hospitals target computer systems, data and medical devices. Attacks on data confidentiality (which is outside the scope of the expert meeting) and availability are frequent, while attacks on data integrity require skills, tools, effort, motivation and malicious intent. Ransomware infections, which usually affect data availability,[109] are on the rise and are expected to remain among hospitals' top concerns.[110] Due to the weak cyber security posture in the health-care sector, ransomware-based attacks are especially cost effective for criminal groups.

For example, hospitals all around the world were affected by Conficker (2008) and, more recently, WannaCry and NotPetya (2017). WannaCry in particular was a large-scale attack that spread indiscriminately. In the U.K., it disrupted the services of over one third of hospital trusts and about 8% of general practitioners' practices and led to the cancellation of an estimated 19,000 appointments – including for at least 139 patients in need of urgent medical care.[111] It can take weeks to fully restore IT services following an incident, which may reduce the availability of health services for patients.[112] NotPetya also disrupted vaccine production, although existing stockpiles were sufficient to fill the gap.[113]

Medical-related systems (IT infrastructure, servers hosting patient files, medical devices, etc.) run on standard operating systems, and specialized hardware (MRI and X-ray machines, insulin pumps and other medical devices) may too. Yet in many cases, these systems are running outdated software. Patch management in hospital settings can be challenging given the many constraints, such as the limited human and financial resources dedicated to cyber security. In addition, once systems are installed, they may be in place for years and may never get a security update; this may be because the systems or software in question are no longer supported by the manufacturer. The lifecycle of medical devices is typically 5–15 years, or longer.[114]

Attacks that tamper with medical devices could result in wrong doses being administered or distort the result of technical analyses used during the diagnostic process. For example, connected pacemakers could be instructed to issue non-standard and potentially lethal shocks.[115] Computer tomography (CT) systems could be targeted in numerous ways; specific concerns include the ability to tamper with radiation doses during a CT scan, with the subsystem responsible for reconstructing images, or with the subsystems that link image results with actual patients. In extreme cases, these actions could be life threatening.[116] If devices like insulin pumps are tampered with, patients could overdose and be injured or die as a direct or indirect result.[117] Attacks on medical devices, including the ability to administer incorrect drug doses, appear to be technically possible – albeit complex – while the likelihood that they could be carried out indiscriminately on a large scale is unknown. However, they would probably not be extremely difficult to detect given the visibility of the effects.

---

[107] Statistics on such attacks are often general and include attacks on data confidentiality, which are not our primary interest here. These numbers are simply provided as an indication of current trends. See also *2016 Health Care Cyber Breach Report*, TrapX Labs, December 2016, p. 13.
[108] Symantec, *Cyber Security and Healthcare: An Evolving Understanding of Risk, Healthcare organizations and their supply chains are under attack—a review of 2017 and a look ahead*, Symantec, 2018, p. 4.
[109] But they may also be able to exfiltrate data. In addition, ransomware can be used to hide the true intentions of an attack. See A. Dahan, "Night of the Devil: Ransomware or Wiper? A look into targeted attacks in Japan using MBR-ONI", Cybereason, 31 October 2017.
[110] *Kaspersky Lab Threat Predictions for 2018*, p. 29 (see note 98 above); *Cyber Security and Healthcare*, p. 3 (see note 108 above).
[111] U.K. Department of Health & Social Care, *Securing cyber resilience in health and care*, U.K. Department of Health & Social Care, London, October 2018; U.K. Department of Health, *Investigation: WannaCry Cyber Attack and the NHS*, U.K. Department of Health, National Audit Office, London, 24 October 2017.
[112] S. Steffen, "Hackers hold German hospital data hostage", *Deutsche Welle*, 25 February 2016.
[113] E. Sagonowsky, "Hack forces Merck to borrow Gardasil doses from CSC stockpile, slamming Q3 sales", *FiercePharma*, 31 October 2017.
[114] "Renewal of radiological equipment", *Insights into imaging*, European Society of Radiology (ESR), October 2014, Vol. 5, Issue 5.
[115] D. Goodin, "Hack causes pacemakers to deliver life-threatening shocks", *Ars Technica*, 9 August 2018.
[116] T. Malher *et al.* "Know Your Enemy: Characteristics of Cyber-Attacks on Medical Imaging Devices", Ben-Gurion University, 17 January 2018.
[117] J. Finkle, "J&J warns diabetic patients: Insulin pump vulnerable to hacking", *Reuters*, 4 October 2016.

Hindering hospital services can also represent a direct or indirect threat to human life, although establishing clear causality may not be easy. The relatively limited scale of past attacks resulted in a temporary loss of services. Yet patients requiring urgent care can be redirected to other facilities, in theory. A coordinated attack that affected all health care providers in a given region, thus preventing referrals (since no facilities were spared), would be of greater concern. Its broader impact on emergency health services remains unclear.

b) Energy

The energy sector regularly comes under attack. US ICS-CERT reports dozens of serious security incidents in that sector every year, although the severity varies.[118] A 2014 report by Symantec ranked the energy sector as one of the top cyber-attack targets, with an average of 74 targeted attacks per day globally (from July 2012 to June 2013).[119] Kaspersky ICS CERT indicated that, in 2018, the energy sector suffered the most attacks (38.7%) in the ICS industry.[120] Persistent cyber campaigns show how attackers try to gain access to operational systems, which they could potentially use at a later date for more disruptive purposes.[121]

Electricity grids are particularly complex, and they offer numerous potential weak points. Attacks may aim to disrupt all or part of an electricity network, and this may deprive civilians of electric power. Electricity networks include facilities like power plants (supply side); transmission lines, distribution and transmission substations (distribution); and devices on the consumer end (demand side). This complex environment relies on diverse software and hardware, ranging from typical operating systems to ICS.

The most damaging effects of cyber attacks on electricity grids so far were experienced in Ukraine. Most visibly, 230,000 people living in the Ivano-Frankivsk region lost power for a few hours in 2015.[122] Attacks in 2016 resulted in power cuts affecting one-fifth of the population of Ukraine's capital, Kiev. While only few detailed examples are publicly known, the threat landscape is evolving and attacks are becoming more sophisticated. For example, the attacks in 2016 reflected a higher level of automation than those in 2015.[123]

Most publicly known cases of attacks on power plant facilities did not have a significant impact.[124] An unidentified American power plant was reportedly forced to shut down following a cyber attack.[125] In another case, reported in 2013, the turbine control system of an American power plant was affected.[126] In 2003, the "Slammer" worm interfered with an (offline) Ohio Davis-Besse nuclear power plant. Digital displays, including radiation and temperature sensor readings, were offline for almost five hours; an existing analogue system functioned and still provided safety readouts. In 2017, the NotPetya wiper reached Ukraine's Chernobyl nuclear power plant and affected the systems responsible for radiation monitoring; employees had to resort to manual readings.[127]

Nuclear power plants are designed so they can be shut down safely in response to undesired events, such as disturbances in the electrical grid, and their systems are typically isolated (air gap) as well.[128]

---

[118] There were 79 attacks in 2014, 46 in 2015 and 59 in 2016, according to information provided by the U.S. Department of Homeland Security: *ICS – CERT Year in Review*, National Cybersecurity and Communications Integration Center, 2015, 2016 and 2017.
[119] C. Wueest, "Security Response - Targeted Attacks Against the Energy Sector", Symantec, 13 January 2014.
[120] Kaspersky Lab ICS CERT, *Threat Landscape for Industrial Automation System in H2 2017*, Kaspersky Lab ICS CERT, 2018, p. 22.
[121] "Dragonfly: Western energy sector targeted by sophisticated attack group", Threat Intelligence, Symantec Blog, 20 October 2017.
[122] Power was restored thanks to manual intervention at the electricity distribution facilities; it took up to a year to restore automated control over some of the facilities.
[123] A. Greenberg, "'Crash Override': The malware that took down a power grid", *Wired*, 12 June 2017.
[124] For more examples: PIR Center, *Cybersecurity of Civil Nuclear Facilities: Assessing the Threat, Mapping the Path Forward*, PIR Center, June 2016.
[125] "Cyber attack shuts down power plant", Electric Light & Power, 15 December 2017.
[126] J. Finkle, "UPDATE 1-Malicious virus shuttered U.S. power plant-DHS", *Reuters*, 16 January 2013.
[127] J. Henley and O. Solon, "'Petya' ransomware attack strikes companies across Europe and US", *The Guardian*, 27 June 2017.
[128] European Commission, *Cyber Security in the Energy Sector*, European Commission, Energy Expert Cyber Security Platform (EECSP Report), February 2017, p. 20.

And since some of these plants were built a long time ago, some of their components – including critical systems in particular – do not rely on digital technology. Nevertheless, significant incidents have occurred in the past owing to a failure to follow standard security procedures, as happened at the aforementioned Davis-Besse facility following the Slammer worm infection.[129] In recent years, cyber security at nuclear sites has improved, although much remains to be done and cyber attacks continue to occur.[130]

Our society is highly dependent on electricity, which means that a broad and prolonged power outage represents a real threat. If a single power plant is affected, other suppliers may be able to compensate, because power distribution networks are grids. Yet cyber risks may not have been taken into account when these systems were designed, and this could jeopardize electricity provision in the event of a cyber attack. The safety designs of mechanical components, for example, address the risks of component failure (e.g. simple mechanical wear), but not necessarily those resulting from deliberate attacks, especially if they target multiple facilities at once. The resiliency models may not have factored in the possibility of numerous power plants being targeted simultaneously. One study in the U.S. looked at the hypothetical scenario of a cyber operation taking 50 selected electricity generators off the grid and causing a sudden drop of at least 10% in generating capacity. The study concluded that the ensuing blackout could affect 93 million people and 15 U.S. states. The scenario was based on a cyber attack that triggered rotating circuit breakers to open and close in quick succession; this would cause the generators to catch on fire and lead to their partial or total destruction.[131] Another (future) risk might be a mass exploit of connected devices to create synchronized peaks of energy demand, leading to blackouts.[132]

c) Water facilities

Water treatment facilities increasingly depend on ICS systems. While there are not many publicly known examples of incidents in the water sector,[133] the water industry is vulnerable to the risks of cyber attacks. A recent report highlighted the potential consequences of exposed control systems, available on the open internet, of a water purification plant and a seawater reverse osmosis plant. It noted that "a cyberattack against water treatment facilities will adversely affect the drinking water in that region, leading to supply shortages. Impure water will also help spread waterborne diseases, leading to a public health crisis".[134]

In a 2000 case in Australia, a former employee released raw sewage into the river, local parks, and residential grounds.[135] In a different case, attackers reportedly changed the chemical mix of a water utility control system, though with no effect.[136] In 2016 the internal network of a water utility system

[129] "Infecting the NPP's process control network with a Slammer should not have happened since the plant had an operating firewall in place, that the primitive worm would not had been able to penetrate. As the incident investigation revealed, there was a serious breach in this security policy: an undocumented connection (T1 line) was set up between the consultant's network and the NPP's business network", in *Cybersecurity of Civil Nuclear Facilities*, p. 14 (see note 124 above).

[130] "Although countries have made modest improvements, many remain poorly prepared to defend against cyberattacks. Among the countries and Taiwan that have weapons-usable nuclear materials or nuclear facilities, one-third lack all of the basic cybersecurity regulations measured", in Nuclear Threat Initiative (NTI), *Building a Framework for Assurance, Accountability and Action*, Nuclear Threat Initiative (NTI), Nuclear Security Index, September 2018, p. 7. "In 2016, three publicly known cyberattacks or attempts on information systems at nuclear facilities occurred at the University of Toyama's Hydrogen Isotope Research Center in Japan; the Gundremmingen Nuclear Power Plant in Germany; and one incident that affected both the Nuclear Regulatory Commission and the Department of Energy in the United States. In 2017, the Wolf Creek Nuclear Station in Kansas had its business systems compromised in a series of attacks targeting the energy sector", in *ibid.*, p. 16.

[131] Lloyd's, *Business blackout, the insurance implications of a cyber attack on the US power grid*, Emerging Risks Report, Lloyd's, 2015.

[132] S. Soltan, P. Mittal, and H.V. Poor, "BlackIoT: IoT Botnet of High Wattage Devices Can Disrupt the Power Grid", Proceedings of the 27th Usenix Security Symposium, Baltimore, USA, 15–17 August 2018.

[133] J. Baylis *et al. Water Sector Resilience: Final Report and Recommendations*, U.S. Dept. of Homeland Security, June 2016.

[134] S. Hilt *et al. Exposed and Vulnerable Critical Infrastructure: Water and Energy Industries*, Trend Micro, 2018, p. 38.

[135] N. Sayfayn and S. Madnick, *Cybersafety Analysis of the Maroochy Shire Sewage Spill*, Massachusetts Institute of Technology, May 2017.

[136] J. Leyden, "Water treatment plant hacked, chemical mix changed for tap supplies", *The Register*, 24 March 2016.

in Lansing, Michigan, reportedly shut down after being attacked with ransomware.[137] In 2018, it was reported that a VPNFilter malware infection at a Ukrainian chlorine plant (used at water and sewage treatment facilities) "could have led to a breakdown of technological processes and possible crash", although no details were made available to the public.[138] In October 2018, ransomware infected systems at a North Carolina water utility service (supplying 150,000 people) soon after a hurricane hit.[139] Taking advantage of natural disasters is a growing concern.[140]

d) Transports and logistics

The availability of detailed public accounts and examples of serious cyberattacks targeting specific elements of transportation systems is limited. The security of airplanes and automobiles is a field of active research, with hacking tests being conducted on planes (or their parts, such as avionics) and car systems, although no hostile attacks have been observed so far. Malware could be used to target individual vehicles or control systems. For example, to interfere with airplanes, a cyber operation may have to first access ground systems.

The 2017 the NotPetya wiper affected transportation and logistics, particularly in terms of availability, resulting in significant financial losses (see Part (g) below). The global logistics corporation Maersk, which accounts for around 15% of the world's container traffic, experienced major system downtime. Some of this resulted from the decision to shut down systems as a precautionary measure.[141] One of the impacts was on the company's ability to process orders; the resulting delays caused congestion at 76 ports around the world. The financial loss was estimated at $300 million,[142] and 4,000 servers, 45,000 PCs and 2,500 applications were reinstalled over the course of ten days.[143] NotPetya also affected FedEx subsidiary TNT Express, resulting in delays, at an estimated cost of $300 million.[144]

Attacks on transportation and logistics could result in serious impacts on the increasingly tight design of global supply chains, and this in turn could affect other layers of delivered goods or materials, including in other industries. Beyond the economic costs, attacks that affect the transportation and logistics of food or medical supplies could have serious cascading effects on human health. This would probably depend, however, on a number of variables such as the specific companies and countries affected, the nature and tightness of the supply chain in question, interdependencies among the delivered goods, the scale of the attack, its duration (including the financial resources that those affected can dedicate to restore the functionality of – or replace – the damaged software and hardware), the ability to quickly switch to other suppliers, and so on.

e) Manufacturing

The manufacturing sector is often the target of intellectual property theft, but it also experiences disruptive attacks.

Cyber attacks may disrupt the functioning of manufacturing facilities, leading to downtime and important economic damage, but they are not usually designed to harm humans. Because of NotPetya, the Dacia, Renault and Nissan car-manufacturing plants ground to a halt.[145] The targeting of petrochemical company Saudi Aramco in 2012 by Shamoon, a malware with destructive capabilities, triggered a major disruption of service that affected about 35,000 computers (although production was

---

[137] E. Lacy, "BWL in limbo from cyberattack", *Lansing State Journal*, 27 April 2017.
[138] C. Cimpanu, "Ukraine Says It Stopped a VPNFilter Attack on a Chlorine Distillation Station", *Bleeping Computer*, 12 July 2018.
[139] "Ransomware attack hits North Carolina water utility following hurricane", *CSO*, 17 October 2018.
[140] New Jersey Cybersecurity and Communications Integration Cell, *Cyber Threat Actors Expected to Leverage Major Storms for Fraud*, New Jersey Cybersecurity and Communications Integration Cell, 14 September 2018.
[141] T. Jensen, "Cyberattack hits shipper Maersk, causes cargo delays", *Reuters*, 28 June 2017.
[142] R. Milne, "Moller-Maersk puts cost of cyberattack at up to $300m", *Financial Times*, 16 August 2017.
[143] R. Chirgwin, "IT 'heroes' saved Maersk from NotPetya with ten-day reinstallation bliz", *The Register*, 25 January 2018.
[144] "TNT Express Operations Disrupted, All Other FedEx Services Operating Normally", FedEx, 28 June 2017; John Leyden, "FedEx: TNT NotPetya infection blew a $300m hole in our numbers", *The Register*, 20 September 2017; "The Untold Story of NotPetya", p. 7 (see note 83 above).
[145] A. Ivanov and O. Mamedov, "ExPetr/Petya/NotPetya is a Wiper, Not Ransomware", Secure List, Kaspersky Lab Blog, 28 June 2017; CERT – BE, *Petya/NotPetya Malware Report on worldwide infection*, CERT – BE, 30 June 2017. M. Scott and N. Wingfield, "Hacking Attack Has Security Experts Scrambling to Contain Fallout", *The New York Times*, 13 May 2017.

not actually reduced).[146] Qatari RasGas was also reported to be a victim of that same malware.[147] This particular malware campaign appears to have operated for a long time, as reports of a similar and related malware, notably targeting chemical firms, came out in 2017, although the effect on these firms is less clear.[148]

The first two confirmed cases of physical destruction caused by cyber attacks occurred at manufacturing facilities: the Stuxnet attack in Iran, in 2010, and another one at an unidentified German steel mill in 2014. While detailed information has not been disclosed, the latter attack "specifically impacted critical process components to become unregulated, which resulted in massive physical damage".[149]

More recently, energy companies experienced threats directly targeting their production. In 2017, an attack against a petrochemical plant in Saudi Arabia was probably aimed at sabotaging production and designed to trigger an explosion, posing a risk to human life.[150] This particular campaign expanded and continued into 2018, affecting other facilities in the Middle East and the U.S.[151]

f) Internet core

Attacks able to harm the core internet infrastructure are difficult to achieve in practice due to the internet's inherent resiliency, including the redundancy of nodes and networks. Among the potential critical points are root DNS (domain name service) servers, the increasingly centralized nature of internet services within cloud computing providers, the importance of certain critical software stacks, and trust infrastructure.

Distributed denial of service attacks (DDoS)

Vulnerable internet nodes (such as servers and consumer devices) can be exploited at scale to form botnets. Mirai malware exploiting IoT vulnerabilities demonstrates this threat. The exploitation of protocol weakness (like memcached) is another risk.[152] DDoS attacks can affect the operations of internet infrastructure companies and hosting providers.[153] This is a developing threat, with botnets constantly under construction. Yet, no DDoS attack to date has had a prolonged and global impact.

DNS root servers

In simple terms, DNS root servers are the key infrastructure components needed to resolve IP addresses to hostnames. DNS root servers are among the most critical elements of the internet's infrastructure, and attacks affecting them could have very serious consequences. Attacks directed at root servers have taken place, notably by DDoS, but so far no attack has succeeded, and DNS root servers are perceived as resilient.[154]

---

[146] T. Sandle, "Shamoon virus attacks Saudi oil company", *Digital Journal*, 18 August 2012; "The Shamoon Attacks", Symantec, 16 August 2012; D. Karakanov, "Shamoon the Wiper, Future Details", Secure List, Kaspersky Lab Blog, 11 September 2012; N. Perlroth, "In Cyberattack on Saudi Firm, U.S. Sees Iran Firing Back", *The New York Times*, 23 October 2012; J. Pagliery, "The inside story of the biggest hack in history", *CNN Business*, 5 August 2015.

[147] "In Cyberattack on Saudi Firm", p. 17 (see note 146 above).

[148] R. Shamseddine, "Saudi Arabia warns on cyber defense as Shamoon resurfaces", *Reuters*, 23 January 2017; "[One joint venture] reportedly had to shut down its computer network [...] [although] the downtime had not affected operations at the facility", in "Saudi Arabia again hit with disk-wiping malware Shamoon 2", *CSO*, 24 January 2017.

[149] K. Zetter, "A cyberattack has caused confirmed physical damage for the second time ever", *Wired*, 8 January 2015. R. Lee, M. Assante and T. Conway, *German Steel Mill Cyber Attack, ICS Defense Use Case (DUC)*, SANS Institute, 30 December 2014.

[150] N. Perlroth and C. Krauss, "A Cyberattack in Saudi Arabia Had a Deadly Goal. Experts Fear Another Try", *The New York Times*, 15 March 2018.

[151] C. Bing, "Trisis masterminds have expanded operations to target U.S. industrial firms", Cyberscoop, 24 May 2018.

[152] M. Majkowski, "Memcrashed - Major amplification attacks from UDP port 11211", Cloudflare, 27 February 2018; "IoTroop Botnet: The Full Investigation", Check Point Research.

[153] L. Newman, "What we know about Friday's, massive east coast internet outage", *Wired*, 21 October 2016.

[154] G. Moura *et al. Anycast vs. DDoS: Evaluating the November 2015 Root DNS Event*, Information Sciences Institute, University of Southern California, 2016.

Cloud centralization
The loss of availability of cloud systems affects services all around the world; such incidents are rare and quickly noticed.[155] As Internet services become increasingly centralized, the potential consequences of attacks affecting cloud service infrastructure increase. The probability of a severe event affecting major cloud providers is low, because availability and security are among the priorities of the cloud provider. However, should this occur, the impact would be significant.

Vulnerabilities with global reach
Successfully exploiting vulnerabilities in software and hardware stacks used on a broad scale could reverberate around the globe. The broadly transformative Heartbleed (2014) vulnerability is a good example.[156] The vulnerability affected a widely used implementation of a protocol employed to encrypt internet traffic (OpenSSL); it allowed data to be stolen from the affected systems, which hosted hundreds of thousands of websites.[157] Attacks on vulnerabilities that could have a global impact mostly concern data confidentiality. The extent to which such vulnerabilities could also be exploited to cause disruptive or even destructive effects at a global scale is yet to be determined.[158]

When identical software is installed on important systems, all of the systems are affected at the same time by the same issues – and they can all be exploited by the same or similar tools. Such high-impact events are usually resolved quickly (in a matter of hours or days). But this time-scale gives attackers an opportunity, especially if they have prior knowledge of some of the vulnerabilities. In addition, many systems will not receive updates in a timely manner, or at all, and they may thus be exploited. This was the case with VPNFilter (which targeted routers) and Mirai (which targeted IoT devices).

Trust systems
Trust infrastructures are crucial for reliable communications. They ensure message confidentiality (i.e. encryption), message integrity (tamper resistance), and they protect end-point identity (knowledge about what system the users are connecting to). Digital certificates issued by trusted certificate authorities (CA) are a critical point of these infrastructures.

The cryptography used in modern certificates is hard to break. It is simpler to attack the CA themselves, in order to issue bogus certificates. This would imperil both users and systems that trust the validity of those certificates. In 2011 a Dutch root certificate authority DigiNotar was breached. Attackers issued over 500 fake certificates (notably, for Google). Some of these certificates were used to eavesdrop on network connections, in particular in Iran, and could have been used to attack dissidents.[159] Another CA, Comodo, has also been attacked.[160] Finally, stolen certificates could be used to tamper with traffic, and this could compromise system software or configurations.[161]

g) Economic cost of cyber operations

Economic damage resulting from cyber warfare poses challenges to both countries and private companies. The overall cost of cybercrime alone is measured in trillions of dollars: it was estimated at $3 trillion in 2015 worldwide, and this figure is predicted to double by 2021.[162] Most of the attacks described in previous sections of this paper caused financial damage, either directly or indirectly – as noted above, NotPetya's impact was well above $1 billion, with some estimates as high at $10 billion.[163]

---

[155] C. Bonnington, "Internet outages take down Alexa, Slack, other web services", *The Daily Dot*, 2 March 2018.
[156] *Common Vulnerabilities and Exposures*, CVE-2014-0160.
[157] J. Leyden, "AVG on Heartbleed: It's dangerous to go alone. Take this (an AVG tool)", *The Register*, 20 May 2014.
[158] This has been well-demonstrated by Mirai-like botnets exploiting the same vulnerabilities in mass installed base products (routers, Internet of Things devices, etc.).
[159] H. Adkins, "An update on attempted man-in-the-middle attacks", Google Security Blog, 29 August 2011; C. Arthur, "Rogue web certificate could have been used to attack Iran dissidents", *The Guardian*, 30 August 2011.
[160] E. Mills and D. McCulagh, "Google, Yahoo, Skype targeted in attack linked to Iran", *CNET*, 23 March 2011.
[161] Potential scenarios include modifying data in transit and forging certificates, causing trust in user systems (e.g. for software update integrity).
[162] S. Morgan, "Hackerpocalypse: A Cybercrime Revelation", Herjavec Group, 17 August 2016.
[163] F. O'Connor, "NotPetya Still Roils Company's Finances, Costing Organizations billion in revenue", Cybereason, 9 November 2017; "The Untold Story of NotPetya", p. 7 (see note 83 above).

Lloyd's estimates that the loss generated by a single extreme event could range from $15.6 billion to $121.4 billion.[164]

The financial sector in particular is under constant threat, and there have been high-impact attacks in many countries.[165] In recent years, the SWIFT network has come under a continuous stream of attacks.[166] Tampering with financial transactions is considered a potentially critical risk to financial stability, due to interdependencies within the financial system.[167]

h) Challenges in assessing the potential effects of cyber operations

Given the enormous variety in attacks, potential effects and scale, developing a coherent framework for assessing the risks of cyber operations is complex. The traditional notion of impact relies on a risk assessment, where risk is measured as a function of the likelihood and severity of events taking place. The severity itself will depend on the types of effects (e.g. causing death versus cutting off electricity), the scale of these effects in terms of the number of people or objects affected (including through cascading or reverberating effects), and their duration (for temporary and reversible effects).

For the purposes of the expert meeting, this assessment should focus on effects of primary concern. These are death, injury or other harm to human beings, physical destruction, and deprivation of essential services. The likelihood will therefore be the probability that a certain cyber operation causes such effects, directly or indirectly, and purposefully, incidentally or accidentally.

Establishing how the "likelihood" and "severity" of the potential human cost of cyber operations should be estimated and measured is difficult in practice. For instance, cyber operations targeting the power grid in Ukraine (2015) affected 230,000 users. This is a large number, yet the electricity outage was relatively short and does not appear to have had serious consequences. The severity of the impact on the population will also depend on the resilience of the system, the redundancy of the services (which in turn depends on factors such as the size of the country), and many other aspects. Establishing the likelihood of a prolonged electricity loss affecting a large population is not easy either.

When it comes to the destruction of objects, specialized cyber operations against ICS appear able to cause intended or unintended physical effects (e.g. destruction or explosions), which may lead to a loss of human life (directly or indirectly). When assessing the risk, such effects would be seen as having a high severity, although the likelihood that the operation would succeed on a large scale is not clear and may be low. The difficulty of assessing the potential human cost of cyber operations against ICS is further compounded by the fact that even the malevolent actor behind the cyber operation may not have full knowledge of the facility, including its current layout, operations and staffing. This makes it difficult even for that actor to properly assess in advance the likelihood and severity of the potential harm.

An objective assessment framework should therefore define notions such as "risk", "likelihood" and "severity". For the reasons mentioned above, an initial impact and risk assessment should focus on the actual severity of plausible scenarios first, then on their likelihood (since likelihood cannot be assessed in the abstract, as it depends on the type of attack and the effects in question).

[164] Lloyd's, *Counting the cost - Cyber exposure decoded*, Emerging Risks Report 2017, Lloyd's, 2017.
[165] For recent attacks reported by the media, see for example D. Brown, "Seven UK banks targeted by co-ordinated cyber attack", *Financial Times*, 25 April 2018; J. Stubbs, "Hackers stole $6 million from Russian bank via SWIFT system: central bank", *Reuters*, 16 February 2018; B. Harris, "South Korea in 'emergency mode' over cyber threat to banks", *Financial Times*, 22 June 2017.
[166] SWIFT facilitates the exchange of information for financial transactions.
[167] T. Maurer, A. Levite and G. Perkovich, *Toward a Global Norm Against Manipulating the Integrity of Financial Data White Paper*, Carnegie Endowment for International Peace, 27 March 2017 (especially the table listing attacks on significant entities).

3. Cyber military capabilities and the protections afforded by IHL

a) Development of, and limits to, cyber military capabilities

As noted by the UN Group of Governmental Experts, "it is in the interest of all States to promote the use of ICTs for peaceful purposes and to prevent conflict arising from their use".[168] While some States and State officials have expressed opposition to the militarization of, or an arms race in, cyber space,[169] in 2012 47 States had cyber security programmes that gave some role to their armed forces according to the UNIDIR Cyber Index,[170] and the U.S. claims that by late 2016 more than 30 States were developing offensive cyber attack capabilities.[171] Some States have declared that cyber space is an operational domain of warfare (in addition to land, sea, air and outer space).[172] Cyber operations also appear to be included in some definitions of information war.[173]

It is today undisputed that international law is applicable in cyber space.[174] Although their specific scope of application and content vary, many bodies of international law offer protection against the effects of cyber operations: first and foremost the United Nations Charter, as well as international humanitarian law, international human rights law, the law of neutrality, international intellectual property law, international telecommunication law, and international law on cyber crimes among others. Domestic law also offers protections against the effects of cyber operations.

After the UN Charter, the body of law most relevant with regard to the limits imposed upon the use of cyber military operations is IHL (also known as the law of armed conflict, LOAC). IHL seeks to limit the effects of armed conflict, protects people, such as civilians, who are not or are no longer participating in hostilities, and restricts the choice of means and methods of warfare.

Even if no IHL treaty expressly mentions cyber weapons, means or methods of warfare, there is no doubt for the ICRC that IHL applies to and restricts the use of cyber capabilities as means and methods of warfare during armed conflicts.[175] International organizations such as the EU[176] and NATO[177] have

---

[168] United Nations General Assembly, *Report of the Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security*, A/70/174, 22 July 2015, p. 6 para 2.
[169] See, for example, *Position Paper of the People's Republic of China For the 73rd Session of the United Nations General Assembly*, p. 10; Declaration by Miguel Rodríguez, Representative of Cuba, at the Final Session of the Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security, New York, June 23, 2017, p. 2; *Response of the Special Representative of the President of the Russian Federation for International Cooperation on Information Security Andrey Krutskikh to TASS' Question Concerning the State of International Dialogue in This Sphere*, Ministry of Foreign Affairs of the Russian Federation, 29 June 2017.
[170] United Nations Institute for Disarmament Research, *The Cyber Index, International Security Trends and Realities*, UNIDIR/2013/3, Geneva, p. 1.
[171] James Clapper, Director of National Intelligence, Marcel Lettre, Undersecretary of Defense for Intelligence, Adm. Michael Rogers, USN, Commander U.S. Cyber Command, Director, National Security Agency, *Joint Statement for the Record to the Senate Armed Services Committee, Foreign Cyber Treats to the United States*, 5 January 2017.
[172] NATO, *Wales Summit Declaration* (Issued by the Heads of State and Government participating in the meeting of the North Atlantic Council in Wales), 5 September 2014, para. 72.
[173] The Agreement on Cooperation in Ensuring International Information Security between the Member States of the Shanghai Cooperation Organization (Ekaterinburg, 16 June 2009) defines information war as "a confrontation between two or more states in the information space with the aim of damaging information systems, processes and resources, critically important and other structures, undermining political, economic and social systems, psychologically manipulating masses of the population to destabilize society and the State, and also forcing the state to take decisions in the interest of the opposing party".
[174] See United Nations General Assembly, *Report of the Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security*, A/68/98, 24 June 2013, para. 19.
[175] See, for example, ICRC, *IHL and the challenges of contemporary armed conflicts*, ICRC, Geneva, 2015, p. 39ff (hereinafter ICRC 2015 IHL Challenges report). The application of IHL to cyber weapons, means and methods of warfare concerns the rules whose application is not specifically limited. For example, there are many treaties regulating the use of specific weapons, ranging from chemical weapons to cluster munitions, for example. These treaties do not apply to cyber means and methods apart from specific situations, such as taking remote control of the enemy weapon systems in question through cyber means. Similarly, IHL rules that only apply to specific domains of warfare, such as sea warfare, would generally not apply to cyber means and methods except when they are used to produce effects at sea, for example. However, the general rules discussed below apply to all domains and all weapons, means and methods of warfare, including cyber ones.
[176] E.U. Council Conclusions, General Affairs Council meeting, 25 June 2013, 11357/13.
[177] *Wales Summit Declaration*, para. 72 (see note 172 above).

stated that IHL applies to cyber space, and the Commonwealth Heads of Government "*Commit[ted] to move forward discussions on how […] applicable international humanitarian law […] applies in cyberspace in all its aspects.*"[178] The applicability of IHL to cyber space has also been affirmed in various official documents and declarations of Member States of these organizations, Japan[179] and the Russian Ministry of Defence,[180] as well as in a publication by a Chinese official in a leading international law review.[181]

Of course, any resort to force by a State, whether physical or through cyber space, remains constrained by the UN Charter, and the application of IHL neither justifies a use of force that would violate the Charter nor encourages the militarization of cyber space. The point of applying IHL – in addition to and independently of the UN Charter – is that any State that chooses to develop or use cyber military capabilities for either defensive or offensive purposes must ensure that these capabilities do not violate IHL. In that sense, IHL provides the rules that belligerents must comply with if and when they resort to cyber warfare.

The secrecy that often surrounds military research, development and operations is particularly prevalent when it comes to the development and use of military cyber capabilities.

b) Use of cyber capabilities during ongoing armed conflicts

Cyber operations have been used in conjunction with kinetic operations during ongoing armed conflicts. Such operations appear to be primarily if not exclusively used in direct support of kinetic operations and for spying, disinformation or propaganda purposes. Cyber attacks have also been directed at the critical infrastructure of countries involved in conflict. However, with the exception of Stuxnet (discussed in the next section), there have been no credible reports of military or other enemy assets being directly destroyed by a cyber operation.

U.S. and U.K. officials have confirmed that they have been using cyber capabilities in their conflict against the Islamic State group. While the results of U.S. efforts were at times seen as disappointing,[182] that country has used cyber operations among others to help identify Islamic State group command

[178] *Commonwealth Cyber Declaration*, Commonwealth Heads of Government Meeting, London, 20 April 2018, p. 4, para. 4.

[179] United Nations, General Assembly, *Developments in the field of information and telecommunications in the context of international security,* Report of the Secretary-General, 9 September 2013, UN doc. A/68/156/Add.1, p. 15 (II. Replies received from Governments; Japan).

[180] "[T]he Armed Forces of the Russian Federation follow the international humanitarian law" with respect to military activities in the global information space, in *Russian Federation Armed Forces' Information Space Activities Concept*, Ministry of Defense of the Russian Federation, 2011, section 2.1. Furthermore, Art. 7(2) of the Convention for Information Security (Concept) proposed by the Russian Federation reads: "In any international conflict, the right of the States Parties that are involved in the conflict to choose the means of 'information warfare' is limited by applicable norms of international humanitarian law". Other statements by Russian representatives are less assertive; see *Response of the Special Representative Andrey Krutskikh*, (see note 169 above).

[181] China's declarations at the Asian-African Legal Consultative Organization (AALCO) do not give the impression that China has a fully settled position on the issue. At the Fifty-Fourth Annual Session of AALCO, Special Half-Day Meeting On "International Law In Cyberspace", 15 April 2015, China stated that "Regarding the use of force in cyberspace, lex lata, including jus ad bellum and jus in bello, applies in principle to cyberspace. At the same time, there is a need to adopt new rules on cyber-Wild West" (AALCO/54/BEIJING/2015/VR, Verbatim Record of Discussions, p. 177). At the Fifty-Fifth Annual Session of AALCO, Meeting of the Open-Ended Working Group on International Law In Cyberspace on 19 May 2016, China stated that "Given the absence of international consensus and state practices on cyber warfare, China does not agree with the above interpretation and application of the right of self-defense and the law of armed conflict to cyberspace" (AALCO/55/NEW DELHI (HEADQUARTERS)/2016/VR, Verbatim Record of Discussions, p. 159). Most recently, Ma Xinmin, Deputy Director-General of the Department of Treaty and Law, Ministry of Foreign Affairs of the People's Republic of China, writing in a personal capacity, stated that: "Secondly, the scope of applicability of the rules of IHL has been expanded. […] Second, it has also been broadened to cyberspace. United Nations Group of Governmental Experts (UN GGE) on Developments in the Field of Information and Telecommunications in the Context of International Security confirmed in its 2013 and 2015 UN GGE reports that international law, particularly the Charter of the United Nations, is applicable in cyberspace. IHL should, therefore, in principle be applicable to cyber attacks, but how to apply is still open to discussion ((unofficial and informal translation): "International Humanitarian Law in Flux: Development and New Agendas - in commemoration of the 40th Anniversary of the 1977 Adoption Protocols to the Geneva Conventions", in *Chinese Review of International Law* (special column), p. 8 (available in Chinese only)).

[182] A. Carter, "A Lasting Defeat: The Campaign to Destroy ISIS", Harvard Kennedy School, October 2017.

posts in order to eventually target them with kinetic weapons.[183] For its part, the director of the U.K. GCHQ explained that the U.K. "conducted a major offensive cyber campaign against Daesh...[that] made a significant contribution to coalition efforts to suppress Daesh propaganda, hindered their ability to coordinate attacks, and protected coalition forces on the battlefield".[184]

Cyber operations have been used in other conflicts and/or countries involved in conflicts or where conflicts are ongoing. However, no party typically claims responsibility for these operations (except for some propaganda-driven ones), which makes it more difficult to assess whether such operations were related to the conflict. For example:

- Cyber attacks were used for propaganda and disinformation purposes, among others, through website defacement, in particular during the Georgia-Russia conflict in 2008 and, since 2011, by a group named the Syrian Electronic Army.
- There have been reports of cyber operations carried out by Israel and Hamas against each other, for purposes of propaganda, disinformation and spying.[185]
- Open sources suggest that the Israel Defense Forces may have used cyber-electromagnetic capabilities during the bombing raid on 6 September 2007 on Syria's Al-Kibar facility, which Israel suspected of being a nuclear reactor, although whether this attack really included a cyber component is unclear.[186]
- In recent years, reports have been made of cyber attacks in support of kinetic operations against the Ukrainian military[187] and of cyber attacks against Ukrainian power grids, electoral infrastructure and other cyber systems.[188]
- Several cyber attacks have been directed at targets in Saudi Arabia in recent years. For example, a wave of wiper attacks against various key critical economic sectors started in November 2016,[189] and in 2017, a cyber operation designed to be destructive was directed at a petrochemical plant.[190]

There is no reason to rule out the possibility that cyber operations have been used in other conflicts or conflict-affected countries and not reported.

The use of cyber operations was reportedly debated and rejected for a number of other conflicts. These include during the NATO campaign in Kosovo against the former Socialist Federal Republic of Yugoslavia, in 1999;[191] before the 2003 Gulf conflict, in order to cripple Iraq's financial system;[192] and at the outset of NATO's operations in 2011 against the Libyan military, to disrupt or disable its air-defence system.[193]

c) Use of cyber operations that would amount to an armed conflict

The use of cyber operations by military forces (or another government agency) when it is not related to an ongoing armed conflict raises questions with regard to the UN Charter. Is the cyber operation lawful? And what lawful responses are available to affected States? In view of the ICRC's mandate, the debates arising from the application of the UN Charter to cyber operations – which raises similar

[183] M. Cox, "US, Coalition Forces Used Cyberattacks to Hunt Down ISIS Command Posts", Military.com, 25 May 2018.
[184] J. Fleming, *Director's speech at CyberUK18*, U.K. GCHQ, 12 April 2018, p. 5.
[185] M. Cohen, C. Freilich and G. Siboni, "Israel and Cyberspace: Unique Threat and Response", *International Studies Perspectives*, Volume 17, Issue 3, 1 August 2016, Pages 307–321.
[186] S. Weinberg, "How Israel Spoofed Syria's Air Defense System", *Wired*, 4 October 2007; L. Page, "Israeli sky-hack switched off Syrian radars countrywide", *The Register*, 22 November 2007.
[187] "Use of Fancy Bear Android Malware in Tracking of Ukrainian Field Artillery Units", Crowdstrike, updated 23 March 2007.
[188] See for example R. Sprang, "Russia in Ukraine 2013-2016: The Application of New Type Warfare Maximizing the Exploitation of Cyber, IO, and Media", *Small Wars Journal*, undated.
[189] Kaspersky, *From Shamoon to StoneDrill, Wipers attacking Saudi organizations and beyond*, Kaspersky, 2017.
[190] "A Cyberattack in Saudi Arabia Had a Deadly Goal" (see note 150 above).
[191] B. Graham, "Military Grappling With Rules For Cyber Warfare Questions Prevented Use on Yugoslavia", *The Washington Post*, 8 November 1999.
[192] J. Markoff and T. Shanker, "Halted '03 Iraq Plan Illustrates U.S. Fear of Cyberwar Risk", *The New York Times*, 1 August 2009.
[193] E. Schmitt and T. Shanker, "U.S. Debated Cyberwarfare in Attack Plan on Libya", *The New York Times*, 17 October 2011.

threshold questions for notions such as the use of force and armed attack as those discussed below for IHL – fall outside the scope of this meeting.

Beyond the issue of compliance with the UN Charter, the use of cyber operations that is not related to an ongoing armed conflict raises the question of whether such operations are governed by IHL. This is to be assessed on the basis of Article 2 common to the four Geneva Conventions of 1949[194] (and Article 3 common to the Geneva Conventions, for non-international armed conflict). In the ICRC's view, there would be no reason to treat a cyber operation resulting in the destruction of civilian or military assets or in the death or injury of soldiers or civilians differently from equivalent attacks conducted through more traditional means and methods of warfare.[195]

States are increasingly vocal in attributing cyber operations to other States. When such operations have gone beyond espionage or information operations, they may have disabled cyber systems or interfered with the functioning of infrastructure or processes relying on such systems. But they have not caused human casualties or, with very few exceptions, physical destruction. No State is known to have publicly qualified a hostile cyber operation outside an ongoing armed conflict as triggering the applicability of IHL.[196] Even in the case of Stuxnet, experts are divided on the issue.[197] It remains to be seen how cyber operations that solely disrupt military or civilian infrastructure will be treated in this regard.

d) General principles on the use of weapons

The main rule that underpins all the rules on weapons and on the conduct of hostilities is that the choice of means and methods is not unlimited.[198]

In particular, the use of means and methods of warfare which are of a nature to cause superfluous injury or unnecessary suffering is prohibited.[199] Probably more relevant for cyber operations, weapons that are indiscriminate by nature are prohibited.[200] This includes weapons (including cyber tools) that cannot be directed at a specific military objective, and weapons whose effects cannot be limited as required by the law of armed conflict. The latter category includes weapons that escape the control of the operator or cause effects that cannot be controlled.[201] For example, a destructive cyber tool that self-propagated automatically and affected civilians and military targets without distinction would be inherently unlawful.[202] Such weapons are absolutely prohibited because they cannot be used without affecting civilians or civilian objects indiscriminately.

e) IHL principles governing the conduct of hostilities

The use of weapons that are not inherently unlawful is governed by the rules on the conduct of hostilities, whose main principles are those of distinction, proportionality and precaution.

According to the principle of distinction, attacks, whether by kinetic or cyber operations (see below on the notion of cyber attack under IHL), may only be directed at military objectives and must not be directed at civilians or civilian objects.[203] Essential civilian infrastructure is a civilian object and

---

[194] Excerpt from Common Article 2: "[T]he present Convention shall apply to all cases of declared war or of any other armed conflict which may arise between two or more of the High Contracting Parties, even if the state of war is not recognized by one of them." M. N. Schmitt (ed.), *Tallinn Manual 2.0 on International Law Applicable to Cyber Operations*, 2nd ed., Cambridge University Press, Cambridge, 2017, Rules 82 and 83.

[195] ICRC, *Commentary on the First Geneva Convention: Convention (I) for the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field*, 2nd ed., ICRC, Geneva/Cambridge University Press, 2016, paras 255–256 on Common Article 2.

[196] On the scarcity of instances where States publicly qualified specific cyber operations under international law, see for example D. Efrony and Y. Shany, "A Rule Book on the Shelf? Tallinn Manual 2.0 on Cyberoperations and Subsequent State Practice", *American Journal of International Law, 112*(4), 2018, pp. 583-657.

[197] *Tallinn Manual 2.0*, para. 15 on Rule 82, p. 384 (see Note 194 above).

[198] Art. 35(1) of Additional Protocol I of 8 June 1977 (AP I).

[199] Art. 35(2) AP I; J-M. Henckaerts and L. Doswald-Beck (eds), *Customary International Humanitarian Law, Vol. I: Rules*, ICRC, Cambridge University Press, Cambridge, 2005 (hereinafter ICRC Customary IHL Study), Rule 70.

[200] ICRC Customary IHL Study, Rule 71 (see note 199 above).

[201] *Tallinn Manual 2.0*, para. 4 on Rule 105, p. 456 (see note 194 above).

[202] See, for example, *United States Submission to the U.N. Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security 2012-2013*, p. 3; U.S. Department of Defense, *Law of War Manual* (U.S. DoD Law of War Manual), June 2015 (updated December 2016), para. 16.6.

[203] See Arts 48–57 AP I, Arts 13–16 of Additional Protocol II of 8 June 1977 (AP II) and ICRC Customary IHL Study, Rules 1 to 21 (see note 199 above).

therefore protected against attack, unless it has become a military objective.[204] Civilians lose their protection against attack when they take a direct part in hostilities.[205] Taking direct part in hostilities can be done remotely, including through cyber means, in which case the programmers, software engineers, threat intelligence analysts, hackers and so on who are doing so (which could be challenging to determine), and only those people,[206] would no longer be protected as civilians. The same is true for civilian objects: they will be lawful targets if they become military objectives because of their use in hostilities.

Objects that simultaneously perform a civilian and a military function are often referred to as "dual-use objects" (which is not a term defined under IHL). This could be the case, for example, of a power station that provides electricity to a military command post and a hospital. Dual-use objects may become military objectives when their military function is such that it meets the definition of military objective. A strict application of this understanding could lead to the conclusion that many objects forming part of the cyberspace infrastructure would constitute military objectives and would not be protected against attack, whether cyber or kinetic. Due to its consequences, such a strict interpretation would be a matter of serious concern.

Furthermore, indiscriminate and disproportionate attacks are prohibited. Indiscriminate attacks are those which are not directed at a specific military objective, resort to means and methods of combat that cannot be directed at a specific military objective, or whose effects cannot be limited as required by IHL, and consequently, in each such case, are of a nature to strike military objectives and civilians or civilian objects without distinction.[207] Disproportionate attacks are attacks expected to cause incidental civilian harm that would be excessive in comparison to the concrete and direct military advantage anticipated.[208]

In terms of precautions in attack, parties to an armed conflict must take constant care in the conduct of military operations to spare the civilian population, civilians and civilian objects. All feasible precautions must be taken to avoid, and in any event to minimize, incidental loss of civilian life, injury to civilians and damage to civilian objects.[209] It is generally agreed that "feasible precautions" are those precautions which are practicable or practically possible taking into account all circumstances ruling at the time, including humanitarian and military considerations.

Finally, in terms of passive precautions, parties to the conflict must take all feasible precautions to protect the civilian population and civilian objects under their control against the effects of attacks.[210]

f) The notion of attack under IHL for cyber operations

The protection afforded by these rules applies whether the attack or other operation is carried out through cyber or kinetic means. However, there is a debate on the meaning of the notion of "attack" in cyber space.[211] This debate is particularly relevant to the principles of distinction, proportionality and precaution, because the most detailed and onerous rules apply specifically to operations that amount to "attacks" (see Part (e) above). It is less relevant, if at all, for the specific protection afforded to

[204] According to Art. 52(2) AP I: "Attacks shall be limited strictly to military objectives. In so far as objects are concerned, military objectives are limited to those objects which by their nature, location, purpose or use make an effective contribution to military action and whose total or partial destruction, capture or neutralization, in the circumstances ruling at the time, offers a definite military advantage." See also ICRC Customary IHL Study, Rules 7 and 8 (see note 199 above).
[205] Art. 51(3) AP I; Art. 13(3) AP II; ICRC, *Interpretive Guidance on the Notion of Direct Participation in Hostilities under International Humanitarian Law* (Nils Melzer), ICRC, Geneva, 2009.
[206] It has to be noted that most cyber operations are not linked to an armed conflict, so IHL does not even apply. Even in armed conflict, most programmers, software engineers, threat intelligence analysts and hackers would remain civilians protected by IHL against direct attack as long as they do not take a direct part in hostilities.
[207] Art. 51(4) AP I; ICRC Customary IHL Study, Rules 11–12 (see note 199 above); see also Art. 13 AP II.
[208] Arts 51(5)(b), 57(2)(a)(iii) and 57(2)(b) AP I; ICRC Customary IHL Study, Rules 14, 18 and 19 (see note 199 above); see also Art. 13 AP II.
[209] Art. 57 AP I; ICRC Customary IHL Study, Rules 15 –21 (see note 199 above).
[210] Art. 58 AP I; ICRC Customary IHL Study, Rules 22–24 (see note 199 above); see also Art. 13 AP II.
[211] See ICRC 2015 IHL Challenges report, p. 41 (see note 175 above); and *Tallinn Manual 2.0*, para. 4 on Rule 92 (see note 194 above).

particular categories of persons or objects, which are afforded more stringent protection also against operations that do not constitute attacks (see Part (g) below).

The debate centres on the notion of loss of functionality of an object, given that in cyberspace it is possible to render objects dysfunctional without physically damaging them.

The most permissive approach is to consider that cyber attacks are only those operations that cause violence to people or physical damage to objects. A second approach is to make the analysis dependent on the action necessary to restore the functionality of the object, network or system. A third approach is to focus on the effects that the operation has on the functionality of the object.[212]

It is submitted that all operations expected to cause death, injury or physical damage constitute attacks, including when such harm is due to the foreseeable indirect or reverberating effects of an attack, such as the death of patients in intensive-care units caused by a cyber attack against the electricity network that then cuts the hospital electricity supply.

The ICRC also considers that an operation designed to disable an object – for example a computer or a computer network – constitutes an attack under the rules governing the conduct of hostilities, whether or not the object is disabled through kinetic or cyber means.[213]

So far, few States have taken a detailed stance on how the notion of attack under IHL applies to cyber operations. Australia considers that the IHL rules governing attacks will apply to a cyber operation that "rises to the same threshold as that of a kinetic 'attack under IHL'".[214] The US Department of Defense Law of War Manual considers that "a cyber attack that would destroy enemy computer systems is an attack" under IHL,[215] while "defacing a government webpage; a minor, brief disruption of internet services; briefly disrupting, disabling, or interfering with communications; and disseminating propaganda" are not attacks, and therefore "may be directed at civilians or civilian objects".[216] The Manual notes, however, that even such operations "must not be directed against enemy civilians or civilian objects unless the operations are militarily necessary [...] [and] should comport with the general principles of the law of war".[217]

g) Specific protection

Some categories of people and objects enjoy specific protection. The protection is not limited to the use of kinetic means; it covers all means and methods of warfare, thus also cyber operations. Such specific protection is additional and complementary to any general protection that such objects may enjoy as civilian objects under the principles of distinction, proportionality and precaution (see Part (e) above). Among objects enjoying specific protection, hospitals must be respected and protected,[218] and cyber attacks against the availability or integrity of medical data, for example, would be prohibited.[219] Moreover, attacking, destroying, removing or rendering useless objects indispensable to the survival of the civilian population is prohibited,[220] and this includes cyber infrastructure required for their functioning.[221] Drinking water installations are an example of objects that are indispensable to the survival of the civilian population.[222] The natural environment also enjoys specific protection.[223] Finally,

[212] See *Tallinn Manual 2.0*, paras 10–12 on Rule 92, pp. 417–418 (see note 194 above).
[213] ICRC 2015 IHL Challenges report, p. 41 (see note 175 above).
[214] Commonwealth of Australia, Department of Foreign Affairs and Trade, *Australia's International Cyber Engagement Strategy*, October 2017, Annex A, p. 91.
[215] U.S. DoD Law of War Manual, para. 16.5.1 (see note 202 above).
[216] U.S. DoD Law of War Manual, para. 16.5.2 (see note 202 above).
[217] U.S. DoD Law of War Manual, para. 16.5.2 (see note 202 above).
[218] See Arts 19–23 of the First Geneva Convention (GC I); Arts 18–19 of the Fourth Geneva Convention (GC IV); Arts 12–14 AP I, Art. 11 AP II; and ICRC Customary IHL Study, Rule 28 (see note 199 above).
[219] See the discussion on data in ICRC 2015 IHL Challenges report, p. 43 (see note 175 above). See also *Tallinn Manual 2.0*, Rules 131 to 134 (see note 194 above).
[220] Art. 54 AP I; Art. 14 AP II; ICRC Customary IHL Study, Rule 54 (see note 199 above).
[221] *Tallinn Manual 2.0*, para 5 on Rule 141, p. 533 (see note 194 above).
[222] Art. 54 AP I.
[223] Art. 55 AP I; ICRC Customary IHL Study, Rules 43–45 (see note 199 above); Convention on the prohibition of military or any hostile use of environmental modification techniques, 10 December 1976 (ENMOD Convention).

particular care must be taken if works and installations containing dangerous forces, namely dams, dykes and nuclear electrical generating stations, are attacked.[224]

Depending upon conditions that vary for each specific protection regime, persons and objects enjoying specific protection may lose this protection – including possibly the protection against attack – when they perform military functions or are used for military purposes. For example, hospitals lose their specific protection if they are being used, outside their humanitarian functions, to commit acts harmful to the enemy. Protection may, however, cease only after a due warning has been given, naming, in all appropriate cases, a reasonable time limit and after such warning has remained unheeded.[225]

h) Legal review of new weapons

States that may use, develop, acquire or adopt cyber-warfare capabilities, whether for offensive or defensive purposes, must assess their lawfulness under IHL, as specifically required by Article 36 of Additional Protocol I.[226] However, the legal review of cyber weapons, means and methods of warfare presents particular challenges, including in deciding:

- which cyber capabilities must be reviewed (considering in particular that military cyber capabilities might require more tailoring for a specific operation than kinetic weapons, and that they might be subject to constant adaptation, for example when the cyber security of potential targets is enhanced);[227]
- at what stage of the development process to conduct a legal review;
- what kinds of procedures, expertise and standards should be employed in conducting a legal review;
- how to appropriately test cyber tools (including anticipating effects that will not occur until the tool reaches the internet).[228]

While several States are known to have put in place mechanisms to conduct a legal review of means and methods of warfare,[229] few are known to have put in place a specific procedure to review the legality of the cyber capabilities they develop or acquire.[230]

[224] Art. 56 AP I; Art. 15 AP II; ICRC Customary IHL Study, Rule 42 (see note 199 above). The *Tallinn Manual 2.0* gives the following example: "Consider malware intended to reduce enemy electrical supply by targeting a hydro power facility. Paying insufficient attention when planning the attack to the effects on the facility's associated gates, and thereby risking destructive downstream consequences, would violate this Rule", in *Tallinn Manual 2.0*, para. 3 on Rule 140, p. 530 (see note 194 above).
[225] Art. 21 GC I; Art. 19 GC IV; Art. 13 AP I; Art. 11 AP II; ICRC Customary IHL Study, Rule 28 (see note 199 above).
[226] Art. 36 AP I reads as follows: "New weapons: In the study, development, acquisition or adoption of a new weapon, means or method of warfare, a High Contracting Party is under an obligation to determine whether its employment would, in some or all circumstances, be prohibited by this Protocol or by any other rule of international law applicable to the High Contracting Party." All States have an interest in assessing the legality of new weapons, regardless of whether they are party to Additional Protocol I. As underscored in the ICRC guide to the legal review of new weapons, assessing the legality of new weapons contributes to ensuring that a State's armed forces are capable of conducting hostilities in accordance with its international obligations. Carrying out legal reviews of proposed new weapons is of particular importance today in light of the rapid development of new technologies (see ICRC, *A guide to the legal review of new weapons* (Kathleen Lawand), ICRC; Geneva, 2006, p. 1).
[227] See also the distinction made in the U.S. Department of the Air Force, *Air Force Instruction 51-401, The Law of War*, 3 August 2018. This instruction distinguishes "any device, computer program or computer script, including any combination of software, firmware or hardware intended to deny, disrupt, degrade, destroy or manipulate adversarial target information, information systems, or networks", which needs to undergo a legal review, from "a device, computer program or computer script developed or acquired [...] that is solely intended to provide access to adversarial and targeted computers, information systems or networks", which does not (p. 13).
[228] See ICRC 2015 IHL Challenges report, p. 44 (see note 175 above). For a detailed discussion of this issue, see, for example, G. Brown and A. O. Metcalf, "Easier Said than Done: Legal Reviews of Cyber Weapons", *Journal of National Security Law & Policy,* 2014, vol. 7, p. 115ff.
[229] See *A guide to the legal review of new weapons*, footnote 8 (see note 226 above), and V. Boulanin and M. Verbruggen, *SIPRI Compendium on Article 36 Reviews*, SIPRI, December 2017.
[230] *Air Force Instruction 51-401* (see note 227 above).

4. Possible avenues to reduce or avoid the human cost of cyber operations

a) Introduction

Concerns about cyber attacks and cyber security are high on everybody's agenda, and there have been many attempts and suggestions, with regard to possible ways forward, to curb the number, extent and effects of cyber attacks. This chapter does not aspire to describe every single suggestion or proposal in an exhaustive manner. Instead, it aims to provide an overview of some of these proposals. This chapter mentions ideas proposed by a range of sources, including the ICRC. To avoid prejudicing the merits of a proposal based on its author, and to encourage discussion on its substance, the origin of the suggestions is not mentioned.

The following overview focuses on the legal, policy and technical realms, and it includes suggestions regarding the development and use of cyber weapons and other cyber military capabilities. This overview does not address confidence-building measures, international cooperation surrounding the development of cyber capabilities, cyber governance, or issues related to espionage (economic or other) or privacy.

The proposals and suggestions mentioned below have not necessarily been endorsed by the ICRC, and those omitted have not necessarily been dismissed by the ICRC. The aim is to stimulate discussions during session 6 of the expert meeting, where an attempt will be made to identify the best avenues that could possibly be pursued to reduce the potential human cost of cyber operations. Other suggestions are welcome.

Although the suggestions below have been grouped for the sake of clarity, they may fall into more than one category.

b) Proposals with regard to norms or rules for cyber space

Rules and norms that apply specifically in cyber space, and in particular to State behaviour therein, have been proposed, or even adopted.[231] To be legally binding, the rules would need to be adopted through a new international treaty or become customary international law.[232] Non-legally binding norms could be agreed upon through various means, such as United Nations General Assembly resolutions, declarations or other outcome documents of State meetings. Some of these norms and proposals were made for peacetime exclusively or primarily, while others were meant to apply at all times, including in armed conflict. ¨

The following are among the norms that have been adopted or the ideas that have been put forward:

- Ensure respect for existing law, and in particular:
    - refrain from engaging in or supporting activities that would violate international law, and in particular the UN Charter, IHL and international human rights law (IHRL)
    - not knowingly allow one's territory to be used for internationally wrongful cyber operations.
- Clarify the understanding of the restrictions and limits imposed on the use of cyber operations by existing law, including but not limited to the UN Charter and IHL.
- Agree upon or reassert prohibitions to attack specific objects, processes, or specifically identified targets, such as:
    - essential civilian infrastructure such as hospitals, other health infrastructure and electricity networks
    - the internet core and other systems, such as financial transactions systems and cloud-based systems, whose disruption could have global effects
    - electoral processes, institutions and infrastructure
    - CERTs (computer emergency response teams).
- Prevent the militarization of cyber space.

[231] "Norms" are usually understood in these debates as not being legally binding; this is how the term is used in this chapter.
[232] Customary international law is based on evidence of a general practice by States accepted as law. While "instant custom" is technically possible, it is rare in practice.

- Make self-spreading malware expressly illegal.

c) Suggestions of technical set-ups linked to international policies or normative frameworks

- Create specific areas in cyber space that would be demilitarized.
- Equip civil defence organizations with the ability to expand their tasks into cyber space, or assign institutions or organizations that already have such capabilities to civil defence tasks in cyber space; create a digital water-mark to identify in cyberspace those objects and infrastructure assigned to civil defence.
- Create a digital watermark that would identify, in cyberspace, objects or traffic belonging to facilities that enjoy a specific protection and for which an emblem, sign or flag protected by international law already exists. Such emblems, signs or flags include the red cross, red crescent and red crystal emblems; the signs for cultural property, works and installations containing dangerous forces and civil defence; and the flag of the United Nations.
- Create legitimate ways for attackers to identify infrastructure that support the delivery of essential services to the population, such as through a global system certificate database, which could, for example, be based on system hardware ID.
- Agree upon a specific digital marker for cyber operations that are solely designed to exploit (espionage, computer network exploitation) and not disrupt or destroy (computer network attack), to limit the risk of escalation that could ensue from the misinterpretation of the initial operation.

d) Other technical suggestions

- Segregate military and civilian networks.
- Segregate computer systems on which essential civilian infrastructure depends from the internet.
- Develop strategies to protect critical cyber infrastructure,[233] and in particular:
    - make advance arrangements to ensure the timely repair of important computer systems against foreseeable kinds of cyber attacks
    - back up important civilian data
    - put in place strong network segmentation
    - create strong and enforceable regulatory frameworks requiring and incentivizing businesses to improve the security of their products, systems and services.
- Work towards reducing the time window when exploitation is possible by incentivizing rapid fixes and designing software and systems to make updates quick and seamless in all possible settings (consumer and corporate sectors, but also regulated ones such as the industrial, medical, aviation, government and military sectors).
- Require or incentivize regulated industries to make system upgrades improving security a priority and to prohibit the use of outdated or otherwise vulnerable systems and software.

e) Suggestions related to the weaponization of vulnerabilities and the development and transfer of cyber weapons

- Refrain from implanting vulnerabilities via the creation of backdoors, whether during the software or hardware design phase, or following a cyber operation.
- Refrain from tampering with the integrity of the supply chain; share information regarding unknown vulnerabilities with those in a position to remedy them.
- In the study, development, acquisition or adoption of a new cyber tool, and in particular when developing tools that spread automatically, determine whether its employment would, in some or all circumstances, be prohibited by international law.[234]
- If designing cyber tools aimed at disabling, destroying or disrupting a specific target, refrain from designing them without having obtained sufficient intelligence and reconnaissance on the target; in particular, ensure on the technical level that the tool focuses only on specifically identified

[233] Such as those recommended in United Nations General Assembly Resolution 58/199, 30 January 2004.
[234] For States party to the 1977 First Additional Protocol, this is required by Art. 36 AP I when the cyber tool constitutes a weapon, means or method of warfare. See also part 3(h) above.

targets or systems and does not affect other systems, even those to which it might accidentally spread.

- Prevent the proliferation of exploits, malware and other cyber tools; limit their transfer and, in particular, do not transfer them to parties that may be expected to use them against civilians or otherwise in violation of IHL or other bodies of international law.
- Explore ways to limit the ability of exploits or malware to be repurposed or reengineered, for example by:
    - obfuscating payload (such as through encryption)
    - refraining from developing malware that can be easily reconfigured, with its purpose or aim significantly changed (such as to prevent the switch from "access" operations to "disrupt or destroy" operations).
- Include technical safeguards to minimize the self-propagation of malware.

f) Create appropriate legal frameworks, processes and tools for international cooperation

- Conclude bilateral or multilateral agreements for international cooperation in cyber space, such as the Shanghai Cooperation Organisation Agreement on Cooperation in Ensuring International Information Security, the Budapest Convention on Cybercrime and the proposal for a Draft United Nations Convention on Cooperation in Combating Information Crimes.
- Ensure cooperation between government CERTs, including by responding to requests for support from States that are victims of cyber attacks, in particular when the cyber attack emanates from or transits through one's territory; refrain from supporting activities harmful to CERTs, and to this effect ensure that the activities of CERTs do not amount to cyber attacks.
- Cooperate between government and private sector CERTs and/or between government CERTs and the private sector more generally.
- Create an organization (whether international or non-governmental) specifically dedicated to the issue of cyber attacks, which could include the following functions:
    - to help victims of cyber attacks to recover
    - to research and publish technical assessments that could support accountability and possibly technical attribution.

MISSION

The International Committee of the Red Cross (ICRC) is an impartial, neutral and independent organization whose exclusively humanitarian mission is to protect the lives and dignity of victims of armed conflict and other situations of violence and to provide them with assistance. The ICRC also endeavours to prevent suffering by promoting and strengthening humanitarian law and universal humanitarian principles. Established in 1863, the ICRC is at the origin of the Geneva Conventions and the International Red Cross and Red Crescent Movement. It directs and coordinates the international activities conducted by the Movement in armed conflicts and other situations of violence.

facebook.com/icrc
twitter.com/icrc
instagram.com/icrc

ICRC

International Committee of the Red Cross
19, avenue de la Paix
1202 Geneva, Switzerland
T +41 22 734 60 01
shop.icrc.org

4391/002 05.2019