Software as a Weapon:

Concepts, Perceptions, and Motivations in Pursuit of a New Technology of Conflict

Jantje A. M Silomon

University College
University of Oxford

A thesis submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

Abstract

This thesis addresses the topic of 'Software as a Weapon' (SaaW) using a mixed-methods approach, bringing together elements of Computer Science, International Relations, and Strategic Studies.

The thesis therefore first addresses the nature of software, malware, and weaponised software via questionnaire-based public solicitation, with three groups of respondents: military officers, academics, and others. The results show that there is consensus among participants regarding the importance of defensive software capabilities for state security. However, depending on the training and background of respondents, questions pertaining to the nature of software exhibit statistically significant differences. For example, when deciding whether software should be treated like a physical object, or whether malware is a weapon. Yet, there is also consensus, such as that defensive software capabilities are vital to a state's security.

The second part of the thesis investigates the factors that contribute to an actor pursuing SaaW. It explores the proliferation debate and examines similarities and differences to traditional weapon groups, including nuclear, biological, and chemical weapons, as well as small arms and light weapons. These factors are then used to create a Bayesian Network model representing an actor's source of impetus. From such a model, it is possible to reason about the interplay of complementary and competing forces. By accounting for restraining and motivating elements, the model introduces objectivity to the debate on actor motivation in the cyber domain, giving a variety of stakeholders a tool to evaluate actors' software weaponisation probabilities.

To showcase and evaluate this model, three different actors are used, representing terrorists, state powers, and generic attackers. Quantitative data is combined with qualitative interviews, populating network nodes with prior probabilities and relative weightings of observed dependencies. The results show that the probability of the generic actor pursuing SaaW is uncertain, which captures the nature of this scenario well. The state actor also shows ambivalence, but in this case high restraints are being countered by almost equally high capabilities, whilst motivating forces are low. The terrorist actor on the other hand has a medium to low probability, driven by a lack of capabilities and limited motivations despite very low restraining factors.

Overall, this thesis emphasises the interdisciplinary nature of cyber security, and provides novel tools and concepts from Computer Science, International Relations, and Strategic Studies to understand SaaW.

Software as a Weapon:

Concepts, Perceptions, and Motivations in Pursuit of a New Technology of Conflict

Jantje A. M Silomon

University College

University of Oxford

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

Acknowledgements

Firstly, I would like to express my sincere gratitude to both of my supervisors, Professors Bill Roscoe and Lucas Kello, for the continuous encouragement and support. I would also like to thank David Hobbs and Maureen York for making the CDT experience unique and truly amazing.

I am very grateful to my Dad for his continued support and belief in me, without which this would have been a far more challenging endeavour. The same is true for my friends and colleagues outside of Oxford, whose moral support has been invaluable – you know who you are! To Jens-Arthur, a tremendous thank you and simply 'RawЯ'.

To my fellow CDT students, thank you for all the fun over the years and making the office such a great place to be, particularly Room 102 in the Robert Hooke Building. I would like to express gratitude particularly to my fellow coffee addict Chris, 'tüftler' Ilias, and cynic Andrew.

Jackie, thank you for being an amazing friend, a great flatmate and an engaging collaborator, your support has been invaluable. My time in Oxford would also have been far less fun without my 'co-conspirators' Busra and Monica, cake-monster Valentin, globe-trotter Kris, and the always-witty Mark.

Lastly, I would like to thank the EPSRC, the CDT in Cyber Security, the Department of Computer Science, and University College Oxford for their generous funding, without which I would not have been able to pursue my doctorate.

Abstract

This thesis addresses the topic of 'Software as a Weapon' (SaaW) using a mixed-methods approach, bringing together elements of Computer Science, International Relations, and Strategic Studies.

The thesis therefore first addresses the nature of software, malware, and weaponised software via questionnaire-based public solicitation, with three groups of respondents: military officers, academics, and others. The results show that there is consensus among participants regarding the importance of defensive software capabilities for state security. However, depending on the training and background of respondents, questions pertaining to the nature of software exhibit statistically significant differences. For example, when deciding whether software should be treated like a physical object, or whether malware is a weapon. Yet, there is also consensus, such as that defensive software capabilities are vital to a state's security.

The second part of the thesis investigates the factors that contribute to an actor pursuing SaaW. It explores the proliferation debate and examines similarities and differences to traditional weapon groups, including nuclear, biological, and chemical weapons, as well as small arms and light weapons. These factors are then used to create a Bayesian Network model representing an actor's source of impetus. From such a model, it is possible to reason about the interplay of complementary and competing forces. By accounting for restraining and motivating elements, the model introduces objectivity to the debate on actor motivation in the cyber domain, giving a variety of stakeholders a tool to evaluate actors' software weaponisation probabilities.

To showcase and evaluate this model, three different actors are used, representing terrorists, state powers, and generic attackers. Quantitative data is combined with qualitative interviews, populating network nodes with prior probabilities and relative weightings of observed dependencies. The results show that the probability of the generic actor pursuing SaaW is uncertain, which captures the nature of this scenario well. The state actor also shows ambivalence, but in this case high restraints are being countered by almost equally high capabilities, whilst motivating forces are low. The terrorist actor on the other hand has a medium to low probability, driven by a lack of capabilities and limited motivations despite very low restraining factors.

Overall, this thesis emphasises the interdisciplinary nature of cyber security, and provides novel tools and concepts from Computer Science, International Relations, and Strategic Studies to understand SaaW.

Contents

Appendices

List of Figures

List of Tables

Acronyms

ABMT Anti-Ballistic Missile Treaty.

APT Advanced Persistent Threat.

BN Bayesian Network.

BTWC Biological and Toxin Weapons Convention.

CBM Confidence Building Measure.

CDE Collateral Damage Estimation.

CFAA Computer Fraud and Abuse Act.

CfPD Child-friendly Parent Divorcing.

CHM Cyber Harm Model.

CIA Confidentiality, Integrity, and Accessibility.

CMA Computer Misuse Act.

CPC Compatible Parental Configuration.

CPT Conditional Probability Table.

CS Computer Science.

CTBT Comprehensive Test Ban Treaty.

CUREC Central University Research Ethics Committee.

CWC Chemical Weapons Convention.

CWMD Cyber Weapon of Mass Destruction.

CWME Cyber Weapon of Mass Effect.

DAG Directed Acyclic Graph.

DBN Dynamic Bayesian Network.

DDoS Distributed Denial of Service.

DoD United States Department of Defense.

DoS Denial of Service.

GC Great Cannon.

GDP Gross Domestic Product.

GFW Great Firewall.

GGE UN Group of Governmental Experts.

GT Game Theory.

HGV Hypersonic Glide Vehicle.

HMM Hidden Markov Models.

IAEA International Atomic Energy Agency.

ICS Industrial Control Systems.

IHL International Humanitarian Law.

INF Intermediate-Range Nuclear Forces.

IoT Internet of Things.

IR International Relations.

JPD Joint Probability Distribution.

KF Kalman Filters.

KMO Kaiser-Meyer-Olkin.

KWH Kruskal-Wallis H.

LAWS Lethal Autonomous Weapon System.

LBN Learning Bayesian Network.

LOAC Laws of Armed Conflict.

MITM Man-in-the-Middle.

MN Markov Network.

MWU Mann-Whitney U.

NATO North Atlantic Treaty Organization.

NBC Nuclear, Biological and Chemical.

NGO Non-Governmental Organisation.

NPT Non-Proliferation Treaty of Nuclear Weapons.

NRC US National Research Council.

NSA United States National Security Agency.

OPCW Organisation for the Prohibition of Chemical Weapons.

PCA Principal Component Analysis.

PLC Programmable Logic Controller.

PPP Purchasing Power Parity.

SaaW Software as a Weapon.

SALT Strategic Arms Limitation Talks.

SALW Small Arms and Light Weapons.

START Strategic Arms Reduction Treaty.

TM Tallinn Manual on the International Law Applicable to Cyber Warfare.

UAV Unmanned Aerial Vehicle.

UN United Nations.

USCC United States Cyber Command.

WMD Weapons of Mass Destruction.

ZDE Zero-day Exploit.

1 Introduction

Contents

Modern life would not function without software: from individuals using smart phones and Internet of Things (IoT) devices, domestic and international commerce, critical national infrastructure, or even military systems. However, cyber attacks, including data breaches, ransomware or other malware have equally become a part of this life. Over a decade ago, the cyber attacks against Estonia in 2007 and against Georgia in 2008 led to the foundation of the NATO Cooperative Cyber Defence Centre of Excellence (CCDCoE) as well as the United States Cyber Command (USCC), with many other countries moving cyber security and strategy up their agendas. In 2010, US Senator Carl Levin gave an opening statement at the Committee on Armed Services in the US Senate, saying that "capabilities to operate

in cyber space have outpaced the development of policy, law, and precedent to guide and control those operations. This policy gap is especially concerning because cyber weapons and cyber attacks potentially can be devastating, approaching weapons of mass destruction in their effects, depending on how they are designed and used" [316, p.147]. Keith Alexander, then head of the USCC and the United States National Security Agency (NSA), stated that there is "much uncharted territory in the world of cyber policy, law and doctrine"[316, p.157]. Likewise, Kello drew attention to the scholarly "inattention toward the cyber issue" [166, p.9], arguing that it will hamper the "intellectual development and policy relevance of the field" [166, p.9].

Since then, progress has been made on these various topics, including academic research in numerous disciplines as well as various initiatives, such as the Paris Call [307] in late 2018. Whilst private actors and industry interest can greatly contribute to establishing responsible behaviour and advancing norm emergence, substantial progress requires interstate cooperation [133]. Furthermore, cyber attacks are also still evolving, "like the cyber-enabled information operations during the US Presidential Elections, [that] have (re)shifted the focus of the field and changed our understanding of what cyber conflict entails" [287]. At the same time, the Council of Europe's Convention on Cybercrime, which requires ratifying states to criminalise certain behaviour, remains the only agreed-upon treaty [305].

An effort was made with the UN Group of Governmental Experts (GGE) process, which sought to address the various perspectives on international norms, measures and rules by bringing together experts from different states, with the overall aim of fostering cooperation and reducing the risk of cyber conflict. Since it was established in 2004 by the UN General Assembly, six GGEs have been convened. Whilst agreement was reached on several recommendations, the process has been hampered by fundamental differences in perspectives and interests. More specifically, the 2015 GGE set out voluntary norms to draw an upper boundary on aggressive cyber activities and to reaffirm the applicability of international law [300], [312]. This process collapsed, partially due to China, Cuba and Russia opposing the inclusion of Laws of Armed Conflict (LOAC) for fear of further militarisation of

the cyber domain. Yet, as Perkovich and Hoffmann argue, this is paradoxical, given not only the development of cyber commands, but also that if cyber means are applied in warfare, there is no reason why they would not have to conform to LOAC [236]. More recently, the US, supported by Australia, Canada and EU states, presented a resolution in 2018, seeking to establish a working group where member states can explore understandings of international law applicability to the cyber domain. There is also an argument that a lack of agreement stems from underlying clashing security interests and power struggles, such as for example in the context of the failed GGE, where "the iceberg underneath is a mix of great power politics, competing and occasionally clashing political systems, and continued uncertainty with respect to a new technology" [199, p.2].

Another example was the advancement of the Wassenaar Arrangement, a voluntary export control regime that shares information on conventional weapon transfers as well as dual-use goods and technologies. It ran afoul of these conceptual and definitional issues in late 2013, when a proposal was made to modify and extend the arrangement to include malware and its related technology [12]. Initial issues were found with the language, limiting beneficial research and information sharing [39]. Though the discussions are still ongoing, the first changes were made in its 2017 version [330]. This showed the importance of including different populations, particularly in this still emerging domain: exploring and sharing concepts across different populations sheds light on potential inconsistencies and misunderstandings, whilst also providing a solid basis to build-upon further. Furthermore, inclusion of a wider community would not only have allowed for greater consistency, but it could also have possibly avoided the issues from the beginning.

Part of the challenge of cyber domain comes from it being both an old and a young field, with a mixed heritage of disciplines to draw upon and include. As Kello argues, "cyber studies, in sum, suffers from two deep structural flaws: different states of preparedness among relevant disciplines; and the absence of a unifying charter to avoid misconceptualization and give their various labors coherence" [167, p.26]. Furthermore, policy makers have to translate "both technical knowledge and

political axioms into policies to defeat relentless threats" [167, p.24], which are not only increasing but also evolving. At the heart of these threats are cyber attacks that are predominantly perpetrated through cyberspace, although they can have origins beyond. The nefarious act can be targeted or non-targeted, with the aim varying from observing to altering, stealing, or even destroying information. Some intrusion events have a single vector or angle of attack, others are multifaceted, pervasive, and persistent; yet others are barely more than cyber graffiti. Whilst cyber attacks between rivals were found to be predominantly restrained and regional from 2001 to 2011 [318], there have been some countervailing examples since, such as NotPetya, Mirai, or WannaCry.

When discussing cyber attacks and capabilities, a wide range of terminology is in use, including software, malware, cyber- or virtual weapons, as well as offensive cyber or information operations. The views on this topic are disparate – not only in terms of terminology but also focus and scope, ranging from strategic thinking and question of cyber war to technical discourses that seek to avoid the cyber-prefix. There are fundamental disagreements about what terminology and labels are used, driven "partly by the technical specificities of cyberspace and partly by the fact that the categorization of military operations is fraught with major policy and legal implications" [36, p.7]. This still new and evolving technology requires a re-examination, or even a re-thinking, of "what have been unequivocal concepts in international law, such as 'weapon', 'attack' or 'armed conflict' " [36, p.vii].

An example is Article 36, 1977 Additional Protocol to the 1949 Geneva Conventions, which aims "to prevent the use of weapons that would violate international law in all circumstances and to impose restrictions on the use of weapons that would violate international law in some circumstances" [154, p.933]. This applies "in the study, development, acquisition or adoption of a new weapon, means or method of warfare" [154, p.933] and was further supported via Rule 110 in the Tallinn Manual 2.0 [266]. However, none of those terms are defined and are thus subject to different interpretations. Moreover, unless a technology or equipment directly contributes to warfare, it is not included in the review process, exempting dual-use technology – a

key element of cyberspace. Take packet analysers such as *Snort* or *Wireshark* as an example. They can be used to troubleshoot networking issues, optimisation, misuse detection, protocol development, and education. At the same time, they can be used for nefarious activities, such as Man-in-the-Middle (MITM) attacks. There is no differentiation apart from the use itself, as the software is functioning nominally in all those cases. Is it then software, malware, a weapon, or possibly even a means or method of warfare? Is the use that makes it into Software as a Weapon (SaaW)?

Consider for example China's Great Cannon (GC) that was first seen in 2016. It is an attack tool that shares infrastructure with the Great Firewall (GFW) but with its own capabilities and design, first used in a large-scale attack on servers of GreatFire.org and GitHub in March 2015 that was analysed by CitizenLab [197]. Whilst the term GFW is often used to encompass any regulation attempt of media and legislation within China, the technical side focusses on Internet censorship achieved by on-path eavesdropping on all traffic coming in and going out of the GFW. In simplified terms, it listens to the traffic and checks if it is banned; if it is, it tells the sender and receiver to abort communication, however, it cannot stop packets already sent. The GC on the other hand can do so, thus being a fully-fledged MITM in-path system, but has so far focussed only on targeted addresses, not canvassing the whole spectrum the GFW does. The attack on GitHub and GreatFire.org saw Baidu infrastructure being used: any user outside of the GFW using a connection to their server enlisted them as an unwitting participant in the Distributed Denial of Service (DDoS) attack. This same tool was used in 2017 against the New York based Chinese news site Mingjingnews, and again more recently in 2019 against the pro-democracy movement in Hong Kong [59], [94]. But is the GC a 'cyber weapon'? Some reports and commentaries used that term following the attacks in 2015 and in 2019 [54], [268], [338]. Others, particularly in the information security community, refer to it as DDoS [94], [214]. Meanwhile, the original analysis of CitizenLab spoke of "an attack tool to enforce censorship by weaponizing users" [197, p.1] and "the weaponization of the Chinese Internet" [197, p.8]. Does the

disruptive as opposed to destructive effect mean it is not a weapon? What if this was part of an attack that distributed various forms of malware?

Or what about NotPetya that pretended to be ransomware but sought to destroy data [29], [60], [129], [136]? Petya, first discovered in 2016, infected the master boot record with a payload that would encrypt the hard drive's file system table, in turn preventing Windows from booting up. The culprits demanded payment in Bitcoins to unlock the systems. A new 2017 variant, NotPetya, used the EternalBlue exploit that had been leaked from the NSA earlier that year, combined with Mimikatz, a proof of concept demonstrating users' passwords remaining in Windows memory [129]. The vast majority of infections affected the Ukraine, but Germany, Russia, and France were also hit [202], believing the attacks were "political operatives seeking to disrupt Ukrainian institutions yet again, using a massive ransom scheme to hide their true motive" [127]. Would NotPetya qualify as a 'cyber weapon', a means or method of warfare? Furthermore, how discriminate can it be – an important element under International Humanitarian Law (IHL) and Tallinn Manual's Rule 105 [323]?

On the other hand, applying the term 'cyber weapon' to Stuxnet is far less disputed, [73], [237], [345]. It was a computer worm discovered in 2010 and believed to be of joint US-Israeli development. It became famous for attacking the industrial Programmable Logic Controllers (PLCs) at the nuclear facility in Natanz, Iran. It utilised several Zero-day Exploits (ZDEs), for example one pertaining to the Print Spooler Service [216], something that had not been seen previously. It had a multilayer operation attacking first the Windows OS, then specific Siemens software for industry, followed by additional Siemens S7 PLCs. Stuxnet is the first instance of known malware to be released into the wild, finding a target, stealthing, sabotaging and deleting itself. An early report by Symantec suggested "five to ten core developers not counting numerous other individuals, such as quality assurance and management" [103, p.3] over a period of approximately six months. Reverse engineering and understanding Stuxnet was equally as challenging a feat and drew on a large number of resources. Experts around the globe worked on various aspects with Langner publishing a detailed technical report in 2013 [183]. A few other

points are noteworthy: firstly, contrary to a majority of attacks, great emphasis was placed avoiding collateral damage, in this case computer and networks that did not meet the specified criteria. The code intruded on a great number of systems worldwide, including personal computers, that it then used to spread further, but did not execute nor cause other damage to those systems; secondly, it had a 'fail-safe' of limiting its spread to three other devices and a self-deletion date of 24 June 2012; thirdly, different versions have since been discovered, for example by having different propagation settings or destructive capability. Lastly, this attack is seen as the first cyber weapon to have become public, one that has also blurred the lines of cyber and physical like none before [157], [187], [345].

Given the difference in the examples above, could or should they all be seen as cyber weapons? An argument has been made that they "span a wide spectrum from specific, highly sophisticated weapons to more generic, less sophisticated ones" [323, p.16] while excluding espionage activities. There has been some effort to use the term 'cyber capabilities' instead of 'cyber weapons' in order to clarify their different nature, and in order to encompass the broad spectrum of activities. Whilst this would alleviate some of the problematic surrounding the term 'weapon', it also implies an extended concept akin to a whole tool-kit. It could, for example, include hardware aspects, such as access to vulnerabilities or the supply chain; it could also include a skilled workforce, such as developers or *in situ* operatives. There are numerous contending definitions, which include Kello, who states that the "crucial definitional criterion of a virtual weapon lies in its intended and possible effects" [167, p.49]; Herr defining cyber weapons as "the combination of a propagation method, exploits, and a payload designed to create destructive physical or digital effects" [141, p. 10]; Brown & Metcalfe on the other hand believe that the same definition of a kinetic weapon applies to a cyber one, an "object designed for, and developed or obtained for, the primary purpose of killing, maiming, injuring, damaging, or destroying" [43, p.135]; or the US Air Force stating that "an Air Force cyber capability requiring a legal review prior to employment is any device or software payload intended to disrupt, deny, degrade, negate, impair or destroy adversarial computer systems,

data, activities or capabilities [313, p.5]. Furthermore, while 'cyber weapon' and related terminology is often used, it is rarely defined, as for example Sanger, who discusses a number of cyber attacks with countless references to these weapons, yet does not explore what they are or how they become a weapon [259].

The cyber domain is still a young and rapidly-evolving field — especially from the perspective of national and international security – and has to call upon a 'congress of disciplines' [167]. Questions and meanings surrounding software, malware, and its weaponisation are still open. The capabilities and tools used in cyber attacks are perceived and defined differently, posing great challenges for cyber security practices. There is also value in arriving at unified understandings of 'cyber weapons' from an analytical perspective. Moreover, a greater illumination of the similarities and differences would allow for discussions not only between academic communities working on issues of cyber security but also extend to include other members of society, such as policy makers. For example, fundamental studies of individuals and different populations are lacking, an investigation of which could help elucidate underlying power struggles and clashes of interest. Another reason to draw upon the opinions of individuals is the fact that they play key roles in this rapidly expanding domain, for example as policy makers or members of the defence community. There is also the possibility that they will operate these new capabilities, even as civilians [15]. Furthermore, individuals not only profit from its advantages but are equally affected by its disadvantages, with the 'WannaCry' attacks in 2017 [22] being one example. Decades ago, the importance of the individual was made in the context of nuclear weapons [135], as a player in a nuclear drama, but this is even more true for the cyber domain: many will have experienced some breach of cyber security, for example in the form of malware or phishing attacks, either intentionally, or as mere collateral. Or, they may have been an unwitting provider of resources and connections, thus acting as a key component in nefarious activity. These experiences can directly affect policy, with individuals making their voices heard in debates and elections. One such example is via the defence community, which can affect the context of major political decisions through its doctrinal or strategic understanding.

1.1 Research Questions

The introduction highlighted the challenges faced when addressing the topic of weapons in the cyber domain. In short, when is it software, when malware, when a weapon? Some strands of thinking have begun to emerge as briefly addressed. The most closely related concept is that of cyber weapons, often also referred to as digital or virtual weapons; however, there is no agreement on the definitions or concepts. Research on related technical aspects, such as on vulnerabilities or various forms of malicious software, dwarfs all other fields of scholarship in this area, but that is slowly changing. Some work centred on cyber war to examine these weapons [247] or took a conceptual approach linking cyber to physical weapons [249]; another developed a conceptual framework for understanding the effects of these weapons on the international order [167]; some centred on the process of development, construction and use [237], or breaking down this type of weapon into components [141] and differentiating between military and non-military uses [143]. Similarly, other research focusses on legal and ethical implications [231], [266], [267]. More recently, this type of weapon was titled a *Perfect Weapon* [259], yet the work skirted definitions and did not discuss when or how software becomes a weapon.

To take a recent example of the importance of definitions, consider the expiration of the Intermediate-Range Nuclear Forces (INF) Treaty in August 2019 that the North Atlantic Treaty Organization (NATO) had considered crucial for Euro-Atlantic security for decades [220]. An argument can be made that without a mutual understanding of what ground-launched ballistic and cruise missiles were, the will alone would not have been sufficient to arrive at a treaty that led to the destruction of 2,692 short- and intermediate-range missiles [220]. This understanding is of even greater importance when agreements cannot rest solely on universally agreed upon (technical) definitions, particularly in a dual-use environment such as the cyber domain that is still evolving. Moreover, a rupture in mutual understanding can have negative consequences, as seen in the discussions surrounding the Russian SSC-8/9M729 system that contributed heavily to the treaty's dissolution. The US first alleged Russia's violation of the INF in its 2014 compliance report. It claimed

that Russia was in breach by possessing, producing or flight-testing a ground-launched cruise missile with a range of 500 to 5,500 kilometres, or by possessing or producing launchers for such missiles [317]. However, it did not specify the type of missile in question, nor state the number or timing of those tests [169]. Russia denied its violation and instead accused the US of non-compliance [72]. At the end of November 2017, US National Security Council official Christopher Ford identified the land-based cruise missile Novator 9M729, NATO designation SSC-8, as being the system in question. One year later, Daniel Coats, Director of US National Intelligence, provided details on Russia's alleged treaty violation, including how Russia conducted two different legal tests that led to the development of a "missile that flies to the intermediate ranges prohibited by the INF Treaty" [70]. In January 2019, Russia rolled out its 9M729 land-based cruise missile insisting it conformed to the treaty, claiming a maximum range of 480km and differences being a better precision guidance system and a more powerful warhead [155]. Whilst an argument has been made that Russia's violation provided a pretext for US withdrawal from the treaty [184], or that Russia might have followed the letter of the treaty but not the spirit, the dissolution also highlighted the surrounding diverging interpretations and the importance thereof.

This is even more vital for new technologies, such as Lethal Autonomous Weapon System (LAWS), military Unmanned Aerial Vehicles (UAVs) or Hypersonic Glide Vehicles (HGVs), which have raised debates on their effect on the international order [262]. There are a number of challenges, particularly "capturing and controlling LAWS, in particular the lack of a common definition, the limited understanding of underlying technologies, and an insufficient awareness of security implications" [11]. This is no different in the cyber domain. On top of clashing security interests and power struggles, endeavours such as norm construction [100], regulation or even 'red-lines' in the cyber domain will likely remain hamstrung. Or, in Maurer's words "The fight over norms for cyberspace starts with the terminology" [199, p.5]. Furthermore, there have been fierce debates within the arms control and related communities not only discussing the technology, but also when, how, or even if

to regulate [15], [82], [113], [150]. Unlike nuclear weapons, for example, the cyber domain has very little tangible artefacts, bringing the treatment of weapons into question and greatly increasing the reliance on deeper conceptual understandings.

These numerous issues surrounding the concepts, definitions and understandings motivated the core research question, which contributes to the academic discourse via a fundamental study discussed in Section 1.2.1, and is explored in detail in Chapter 3:

1. *What does it mean for software to be a weapon?*

The definition and conception of SaaW can also shape the strategic purpose of acquiring them. An understanding of where and how a capability can be used affects the motivations, objectives and restraints of weapon acquisition, similarly to how sometimes "relatively tech-specific rules will be preferable to ensure that a particular technology is used (or not used) in a particular way" [81]. For example, an actor may wish to ensure that a certain capability fulfils or avoids a number of measures or controls, such as weapon reviews or potential export controls. Alternatively, such a capability may give an actor an option to engage targets that were previously out-of-bounds: "When a sledgehammer is excluded by LOIAC owing to the expectation of 'excessive' injury/damage to civilians/civilian objects compared to the anticipated military advantage, the availability of a scalpel may open the legal door for an attack at a lawful target" [92, p.169]. The first research question and fundamental study therefore provides a starting point to explore why actors, for example states, wish to pursue SaaW.

Whilst weaponisation of software might be a new element of 21st century warfare, the pursuit and proliferation of weapons is not, nor is the study thereof. The literature, particularly on nuclear weapons, spans over half a century and is drawn upon as a previous major technological game changer. Strategy emerged as "an intellectual field of endeavour in response" [297, p.101], aiming to prevent yet another great power war. What can be learnt from that? To what extend are the driving forces and restraints the same? What about other weapon technologies, such as biological or chemical ones? And what about Small Arms and Light

Weapons (SALW)? How do the different schools of thought in International Relations (IR) impact this topic? These and related questions provide the backdrop for Chapter 4, which examines:

2. *What factors contribute to the pursuit of SaaW?*

Various factors are brought together on the basis of an extensive literature review, encompassing elements of Computer Science (CS) as well as IR, as well as the empirical results of the fundamental study. This provides the groundwork for the conceptual model that contributes not only to the academic discourse but can inform policy debates, as seen in Section 1.2.2. For example on what can be done to curb the pursuit of SaaW more generally, or with an eye to an actor, such as a specific state, as "identifying, understanding, and quantifying the complex interrelationships underlying even seemingly simple situations can help us make sense of how risks emerge, are connected, and how we might represent our control and mitigation of them" [105, p.49]. Academic insight can, for example, help make sense of the globalised world, provide a "counterweight to governmental efforts to manipulate public perceptions", and offer a "useful model of constructive debate" [326]. Part of this is due to the role of theory in the policy process, which involves weighing, choosing and using implicit or explicit theories that are often competing. Here, debates are vital, as "relying on bogus theories can get states into deep trouble", whilst "good theories often produce beneficial policy results" [325, p.34]. In turn, policy makers gain insight into new situations and likely developments, which is particularly important in a still evolving environment such as the cyber domain, despite drawbacks, such as diverging agendas and foci. In this particular context, the conceptual model of SaaW pursuit can act as a further element of discussions on arms control and disarmament or influence norm construction, for example by exploring retraining or motivating aspects and how these align or differ across actors.

Furthermore, the conceptual model then provides a basis for implementing it in conjunction with case studies, seeking to understand the probability of a specific actor pursuing this technology, answering the final research question:

3. *What is the probability that a given actor is pursuing SaaW?*

Whilst understanding tendencies is important, knowing what is most likely to happen in a specific case can support policy makers more strongly [325]. It also provides a starting point of how to manage the situation going forward. For example, if tailored to a specific actor, one can not only tell how likely current probability is for pursuit is but also what shifts in policy, for example increased or additional restraining factors, may effect. It could be explored how a specific actor would react to a stronger promise of retaliatory attacks, or tighter international retaliation. Similarly, changing variables that contribute to the motivation could iterate different starting points to examine their effectiveness.

This question is addressed in Chapter 6 using three case studies implementing the conceptual model created in Chapter 4 with the use of Bayesian Networks (BNs), highlighted in Section 1.2.3 and discussed in detail in Chapter 5. The case studies centre on a *Generic Actor*, a *State Actor*, and a *Terrorist Actor*, resulting in a model with probabilistic information. Whilst similar work has been done in the nuclear field [71], [110], [147], this thesis is the first to apply it to the cyber domain, particularly SaaW. For this purpose, a new data set is created, combining quantitative data and qualitative interviews, populating nodes with prior probabilities and relative weightings of dependencies.

If the empirical analysis of the thesis shows that the probability of a state actor pursuing SaaW when motivation is low sees capabilities balanced by restraints, it means that current efforts to curtail the pursuit are sufficient for the time being. However, if the opposite is true, that is restraining factors having no impact in light of capabilities despite low motivation, it would mean that new approaches and concepts are needed on how to prevent further pursuit. This could, for example, be in the form of stronger discouragement either via deterrent measures or by other means of discouragement.

Using this model, future work could also be to assess what effect changes to specific elements may have on an actor given a scenario. For example, it could be explored what would happen if certain restraining elements, such as ‘fear of

retaliation' were increased, decreased, or removed, whilst keeping all other factors equal. This could similarly be applied to motivational factors or capabilities.

In summary, the cyber domain is still a young and evolving field, with a variety of perspectives and understandings, even of fundamental constructs, such as malware and its weaponisation. This thesis seeks to explore these concepts, followed by creating a conceptual model of SaaW pursuit that is then implemented and can be applied in a variety of situations.

1.2 Research Contributions

This thesis provides four core contributions to cyber security with an interdisciplinary focus:

1. *A fundamental study that examines perceptions and concepts across different populations pertaining to software, malware and its weaponisation, including the creation of a data set* (see 1.2.1)
2. *A conceptual model depicting the various factors that contribute to the pursuit of SaaW, bringing together the disciplines of CS and IR, that can be used by various stakeholders* (see 1.2.2)
3. *A BN implementation of the model that can be evaluated empirically and adjusted* (see 1.2.3)
4. *The creation of a new data set pertaining to actor motivation in pursuing SaaW, which is also used to evaluate the BN model* (see 1.2.4)

1.2.1 Fundamental Study

This study explored the first research question: *"What does it mean for software to be a weapon?"* by carrying out and analysing an online questionnaire sourcing public attitudes towards SaaW. As seen earlier, public attitudes matter because they play key roles in the cyber domain. The members of the public might, for example, be members of the defence community, policy makers, lawyers, or IT specialists. They can also be "a target, a steppingstone to the target or mere collateral; equally, an

individual can provide technical know-how or be merely an unwitting provider of resources and connections, thus acting as a key component" [281]; they might also be called upon to operate these new capabilities. To explore these opinions, the questions were divided into three main sections and mostly used the five-part Likert-scale: weapons, their nature and constitution; software, malware and understandings thereof; as well as SaaW in the context of international security, including capabilities and proliferation. The respondents were also split into three groups: *Academics*, *Military Officers*, and general population (*Others*). These groups were targeted because it was hypothesized that respondents in positions of responsibility or those with experience using other weapon technologies will have diverging opinions.

The results show significant differences in the interpretation and perception of three basic constructs depending on the background and training of respondents, underlining the population heterogeneity in the understanding of the potential of software and its use as malware or a weapon. Understanding this heterogeneity is important to prevent misunderstandings by being clearer about the various standpoints. This can be used to better understand the various actors and potential adversaries, as well as provide another avenue towards norm creation.

Whilst details are discussed in Chapter 3, some key results are presented below. Unsurprisingly, the opinions on what a weapon is varied greatly. Yet, there was also some consensus: around two-thirds consider it to be an object that is designed or can be used to cause harm or damage, and weapons are offensively driven – which is lower than expected prior to the study. When deciding whether malware is a weapon, opinions were split, as they were when asked if a threshold is needed for this to occur – an ambiguity or disagreement that can be seen throughout academia, as well as policy.

Furthermore, *Academics* are more likely to disagree that software should be treated like any physical object, *Military* respondents are more likely to agree and *Others* are indecisive – with the differences being statistically significant.

There is substantial overlap between words used to describe 'software' and 'malware' by the respondents. Whilst 'software' appears to be thought of having a

broad function, 'malware' is characterised by unique expressions. The same overlap is observed between 'weapon' and 'malware', with some unique words used to define a 'weapon'. Overall, it appears that a 'weapon' and 'software' were thought of as predominantly separate constructs, whilst malware is defined as a bridging construct, connecting what is believed to be a 'weapon' and 'software'. This further supports the idea that software can be used as a weapon, but is not necessarily one *per se*. It also helps to establish a basis of understanding, which is needed to explore why any actor would explore pursuing this use of technology.

A clear majority believe malware can cause physical damage but there is disagreement as to what type of potential attack causes it to be a weapon, whether physical damage (to a living being, structure or system) is needed or whether an attack on Confidentiality, Integrity, and Accessibility (CIA) is sufficient, with difference between the *Military* and *Others*.

Whether the intent of the attacker or the damage (effect) caused is what makes software/malware a weapon is inconclusive: on the one hand, most disagree that damage matters regardless of intent; on the other hand, there is no consensus that intent is the decisive factor. In both cases *Academics* deviate more, particularly on the second question where it is statistically significant.

The responses on SaaW created valuable data contributing to the varied opinions on and understandings of weapons in the cyber domain. The work was published in two blindly peer-reviewed papers [280], [281], as well as an invited book chapter [279] that was co-written as lead-author with Mark Patrick Roeling, who contributed the word clouds and performed the Principal Component Analysis (PCA) computations. The latter includes an improvement of methodological aspects, allowing for a future iteration of the questionnaire to be more useful by being more streamlined and easier to analyse.

1.2.2 Conceptual Model

The resulting model was designed on the basis of an extensive literature review and the empirical results of the fundamental study. The literature review was broad,

encompassing elements of CS as well as IR, with a focus on topics related to SaaW. More specifically, this included work on vulnerabilities and malware, including legal perspectives, on the one hand, and a focus on weapon proliferation and its motivation on the other. Aside from furthering the academic discourse, it can be used to introduce the complexities of SaaW to a wider audience by providing a common starting point, which is vital because the interconnected nature of the cyber domain affects everyone, from individuals to state-actors. By seeking to understand what drives actors to the pursuit of SaaW via the second research question: *What factors contribute to the pursuit of SaaW?*, a basis for policy discussion is created. For example, knowing what the reasoning for seeking this technology is could help answer how actors can be dissuaded from doing so, as arms control has done in other domains. Or, more broadly, it can support the discourse of the arms control community by adding to the understanding of the cyber domain.

Within IR, there are competing traditions and schools of thought that seek to provide an explanation and coherent theory of why, and at times how, actors disseminate or acquire new weapons, whether nuclear or other. However, no single theory can account for all the historic scenarios thus far [321]. Instead, each theory has varying degrees of explanatory power, with different strengths and weaknesses for specific scenarios [58]. Whilst fundamental disagreements pose a great challenge, there is also an advantage of drawing on elements from various schools of thought. For example, it can account for elements that might not at first appear relevant, or which go against one's beliefs. This can further the academic debate, raising questions of the relative importance of factors, or even shed light onto interactions not previously considered.

Whilst the data represents a snap-shot in time, the model is flexible to allow for technological change. This aim also influenced the choice of implementation, with BNs allowing for variables to be flexible, for example by adjusting their impact strength, or even adding or removing some. Yet, this also has downsides. For example, focussing on a specific school of thought would have allowed for greater

depth, analysing its strengths and weaknesses more deeply. More scenarios or iterations could have been conducted, building a more solid foundation. Alternatively, two competing theories could have been compared.

Given the nascent nature of SaaW, the speed of development and thus also lack of cyber specific history, as well as the aim of an interdisciplinary work, the decision was made not to focus on one school of thought but to draw on factors throughout. The resulting model uses a set of determinants, or variables, broken into three broad categories of analysis pertaining to weapon development and proliferation: *restraints*, *motivations*, and *capabilities* of an actor.

In short, the *restraints* are essentially reasons for an actor not to develop or pursue SaaW. They could also form part of the *motivations*, albeit in a negative manner. Against a backdrop of increasing and intensifying conflict arising in the cyber domain, it is particularly important to explore and analyse what elements can limit, or even control, further escalation. This could be used to inform policy, particularly discussions involving the GGE or the arms control and disarmament communities, as well as influence norm construction. Without any *motivations* to develop and pursue weapon technologies, the best capabilities will not be utilised and are rendered irrelevant, thus these need to be examined in detail. Lastly, without *capabilities*, the only options to gain SaaW are to steal or buy these which are not in scope of this thesis. Each is discussed in turn, examining interrelationships and sub-elements, both competing and complementing details are presented in Chapter 4.

The conceptual model and elements of the literature review were published in one blindly peer-reviewed paper [277], one invited journal article [278], as well as a related co-authored blog post on the challenges and opportunities of norm construction [101] that was then extended into a journal article [100].

1.2.3 Bayesian Network (BN) Implementation

When deciding to implement the conceptual model, several analytical approaches were considered. The first was Game Theory (GT), which has a long history of being applied to related areas, such as super-powers [38] or conflict resolution [242],

following Schelling's seminal work on bargaining and strategic behaviour [263]. This method essentially requires a conflict or contest between at least two actors, interacting in a game governed by a set of rules. Whilst this simulates behaviour of adversaries well, the focus here is the interaction of factors within an actor, as well as examining the likelihood the actor will pursue SaaW. Furthermore, a lot of information in the context of SaaW is either lacking, incomplete or missing, riddling data with uncertainty, which is a stumbling block for a game theoretic approach, as does the presupposition of rational, gain-maximizing actors.

Another implementation could have used a Markov Network (MN), which is an undirected graphical model that can encode a set of independence relations. They are used to model a variety of phenomena if there is no natural directionality to the interaction of specific variables, i.e. there is no causal relationship. They would have allowed for cyclic relations but they lack of directionality, which is part of addressing factors that lead to the pursuit of SaaW, contributing to the decision to not use MN.

Directionality and the management of imperfect or unknown information, is, however, a particular strength of BNs, providing a convenient framework for reasoning about knowledge that has levels of uncertainty [105]. A BN is a type of probabilistic graphical model that depicts a set of variables and associated conditional dependencies in a Directed Acyclic Graph (DAG) and is based on a formal mathematical framework. The implementation of the conceptual model and the underlying background and theory is discussed in Chapter 5.

BNs were also chosen because they have been used in a variety of situations, for example to model and explain a subject area, to find the most probable configuration of variables, or to support decision making and strategies involving uncertainty. They are applied in a large variety of disciplines and scenarios, ranging from medical diagnosis support [209], [273] and device troubleshooting [40], [170], to more related fields of digital forensic evidence hypothesis testing [179]–[181]. Furthermore, they have been implemented to model nuclear weapon proliferation [71], [110], [147], with the aim of balancing the varied political theories and reducing overall bias.

The arrows in the DAG of a BN represent real, casual connections between the variables, unlike in for example neural networks, where this would represent the flow of information during reasoning. Reasoning occurs by propagating information in any direction, with a classic example including a sprinkler, grass, and rain. A prediction would for example be that if the sprinkler is turned on, the grass is probably wet. By reasoning, if somebody were to slip on the grass, it would contribute to the evidence that the grass is wet. If this was turned around, wet grass would mean that there is a high probability that it is either raining, or that the sprinkler is on. If the sprinkler is observed to be on, the likelihood that rain made the grass wet is reduced. This last element, making an explanation 'go away' or at least highly reducing its likelihood is a core advantage of BN that for example neural networks or rule-based systems do not have. In the context of SaaW and pursuit thereof, it can for example be used to find the most strongly driving impetus of a specific actor, supporting decision making on how to act and what risk is posed.

Nonetheless, there are also some drawbacks. Elicitation of expert opinions comes with challenges, from inherent human bias to lack of understanding. Decades of implementing BNs in various fields have explored this weakness, finding several forms of mitigation, including standardised scripts, examples, training exercises or feedback and revision processes [160], [163], [165], [215], [222]. Another challenge is posed by the number of questions that need to be asked to gain information due to the exponential nature of BNs, but there has been work to address this by creating an algorithm for compatible configurations following human reasoning [85]. Lastly, given the reciprocal nature of war, it could be argued that a BN may be too one-sided, not accounting for the opponent. This, however, could be mitigated by another (or several) competing BN run in parallel, or even connected in a future, more complex BN.

In summary, BNs have the ability to 'tell a story' by using "visual and formal mechanism for recording and testing subjective probabilities" [105, p.48]; various forms of data, both quantitative and qualitative can be combined; they excel at

handling uncertainty and forms of reasoning, outweighing the drawbacks leading to our choice of this form of implementation.

The BN created rests upon the conceptual model and was chosen as the method for implementation due to its flexibility of managing various forms of data, as well as uncertainty. The various factors from the conceptual model became the nodes, their relationships, and interactions the connections. Combined with the dataset, it explores the third question: *What is the probability that an actor is pursuing SaaW?*

Given the number of discrete nodes and states and the exponential nature, indexing all the possible combinations became highly problematic, if not impossible. One alternative would have been a re-design of the model itself, attempting to avoid nodes with a large number of parents, either done 'manually' when creating the network, or it could be applied by utilising Child-friendly Parent Divorcing (CfPD) [228], [250], [322]. Instead, an algorithm first developed by Das [85] was adapted and applied. It uses an approach that captures experts' judgemental strategies by weighing relative parent-node's influence strength and using Compatible Parental Configurations (CPCs) to create a linearly growing set of probability distributions.

The BN consists of many parent-nodes that flow into three main sets of nodes representing *restraints*, *motivations*, and *capabilities*, which in turn inform the overall SaaW node and determine its probability. Some parent-nodes in turn have parents, with values determined by Conditional Probability Table (CPT) or priors in case of leaf nodes.

Early versions also included a heavy focus on the technical aspects as part of the *capabilities*, with various sub- and sub-sub-elements, for example breaking down SaaW delivery mechanisms into direct/indirect types, with insiders, externals, bot-based, drive-by-downloads, social engineering and so forth. Similarly, the payload element was broken down into various aspects such as code development, with its different attack vectors and obfuscation techniques. However, getting access to technical experts and specialist data was prohibitively challenging, thus reducing these elements to three nodes (test bed, payload package, and delivery mechanism), with other examples available in Appendix B.

The model structure, encompassing nodes, connections, and states, can easily be implemented by other researchers for future use, as well as adjusted or expanded. The findings, including the part of the dataset, were published in one blindly peer-reviewed paper [275].

1.2.4 Actor Dataset

The dataset was created in combination with the BN model discussed above, relying heavily on interviews discussed in detail in Chapter 6. Aside from the model structure, the various priors and CPT elicited are also available for future use or comparison.

1.3 Scope

The main focus of this thesis are weapons: the concepts thereof, how they apply to the cyber domain and what motivates actors to proliferate these. Several other topics are intrinsically linked, such as security and war, but these are not within scope, nor is the subject of targeted misinformation campaigns (*fake news*). Similarly, although the study of how weapons spread is related and could draw on epidemiological models, this is also not part of this thesis.

1.4 Ethical Considerations

Ethical concerns may be potentially raised with regard to the collection of peoples' opinions discussed in Chapter 3. This has been addressed mainly through anonymisation at the collection level via the *Online Surveys* system (formerly Bristol Online Survey), and by limiting demographic questions to a minimum. Furthermore, the system used is fully compliant with all UK data protection laws and meets UK accessibility requirements.

The other ethical consideration that has to be taken into account is the data collected for the case studies in Chapter 6, as data could be leaked. To mitigate this, the data was stored confidentially in an encrypted folder on the departmental

laptop (located in a locked office), which is also password protected. A backup was stored on a removable encrypted electronic storage device, kept in a locked cabinet (in said office), alongside any notes taken in longhand. Participants were also given the option to remain anonymous and were asked specific permission for any direct quotes and attribution thereof.

All study participants were sent an information sheet beforehand, and informed consent was collected prior to enrolment. As required for all experiments involving human participants, approval was sought and gained from Oxford's Central University Research Ethics Committee (CUREC).

1.5 Thesis Outline

The remainder of the thesis is structured in the following manner:

Chapter 2 places the thesis in the context of related work in the form of a literature review. It peruses the different fields of study, primarily CS, IR, and Strategic Studies, with an emphasis on introducing key concepts and theories of each.

First, weapon and their pursuit are contextualised as a strategic tool in relation to war and warfare, before focussing on cyber weapons and related constructs, comparing their respective environments and weapons. Elements of arms control and specific cyber initiatives are also included, as they contribute towards a foundation for understanding weapons in the cyber domain. This chapter provides a backdrop for exploring '*what does it mean for software to be a weapon*' by examining the current state of thinking. The literature review will show that the notion of cyber war is still contentious. Whilst some believe it to be something new and different, others argue that merely the means of waging war have changed. Similar is true of the concept of 'cyber weapons', ranging from various malware terminology to virtual weapons.

Chapter 3 discusses the findings of an opinion survey on SaaW that solicited the public, particularly three groups: *Military Officers*, *Academics* and *Others*, directly relating to the first research question. Across 46 questions, respondents were asked

to give their opinion on statements regarding weapons, software, malware, as well as the weaponisation thereof. The chapter will show that there is often a great variety of opinions, such as for example when defining a weapon, and it is in many cases statistically significant, for example when asked whether offensive software capabilities are vital to a state's security. In other areas, there is consensus across the groups, particularly when addressing questions of regulation or difficulties in enforceability.

Chapter 4 combines the understanding of weapons in the cyber domain from the previous chapter with proliferation literature by also delving briefly into political schools of thought in order to create a conceptual model of what motivates SaaW pursuit.

The various factors are extracted, analysed, and adapted to the cyber domain in three main categories that explain an actor's impetus: *restraints*, *motivations*, and *capabilities*. These elements and interrelationships create a model that provides a unified starting point for debates across various fields of study. Furthermore, it provides a blueprint for implementation as a BN, which is used to examine three case studies in Chapter 6.

Chapter 5 introduces probability theory and BNs, before discussing how they are applied to the topic of SaaW. BNs have been used across a large number of fields for several decades, providing a convenient framework when elements of uncertainty are involved. They elegantly combine qualitative and quantitative aspects, allowing for a very intuitive graphical depiction, which is supported by a formal mathematical framework. Furthermore, they can be constructed by combining qualitative human expert knowledge with data. Aside from BN theory, this chapter also discusses how the data will be gathered and examined, how the network is created, and how it is evaluated.

Chapter 6 focusses on the last research question and evaluates three case studies using the BN: a *Generic Actor*, a *State Actor*, and a *Terrorist Actor*. Qualitative results relate predominately to network structure, for example such as requesting additional nodes and interconnection, or adjustments of others. Quantitative results on the other hand will for example show that the state actor is ambivalent about pursuing SaaW, with high restraints being countered by high capabilities, whilst the overall motivating forces are low. This means that in this scenario, current efforts to curtail the pursuit are sufficient but they might not be going forward should motivational aspects increase. This would then have to be 're-balanced', for example via new approaches and concepts in terms of restraints, such as deterrent measures or by other means of discouragement. The terrorist actor, on the other hand, has a medium to low probability of pursuit, due to his lack of capabilities despite very low restraining factors. This presents a challenge for the practice of cyber security should the actor increase their capabilities, particularly because restraining options are very limited.

Chapter 7 concludes this thesis, summarising the results and discussing potential avenues of future work.

2
A Literature Survey

Contents

This chapter presents a literature review to situate the topic of SaaW, drawing on a number of related areas, particularly CS, IR and Strategic Studies. Cyber security is still a very young, rapidly evolving field that is increasingly gaining prominence as a topic, and as a field of study [91], [122], [123], yet its implications are already a major global concern, with cyber operations topping the US intelligence community's threat assessments [236]. Despite there having been no reported deaths resulting from direct cyber actions, ideas of cyber war and cyber warfare remain, and with them a lack of specific qualities, boundaries and appropriate responses [236].

This literature review first contextualises weapons and their pursuit as a strategic

tool in relation to war and warfare, before focussing on cyber weapons and related constructs, comparing their respective environments and weapons. It ends with a brief overview and discussion of arms control and specific cyber initiatives.

The literature review will reveal several noteworthy points. The notion of cyber war is still contentious: some interpretations see it as something new and different, whilst others merely believe the means of waging it have changed, for example via the inclusion of cyber capabilities. This varied understanding is also seen when examining cyber weapons, with terminology spanning various denominations of malware to virtual weapons.

These diverging perceptions provide the groundwork for examining '*what does it mean for software to be a weapon*' discussed in Chapter 3, which adds the perspective of different population groups to the debate. It further is explored in Chapter 4, which addresses the factors that contribute to the pursuit of SaaW, as that relies on knowing what SaaW is and how it is different from other types of weapons.

2.1 Cyber: War and Warfare

There are numerous ways of arriving at a definition of war, with core ones being based on international law, IR, and military thought (e.g. via Clausewitz), alongside more colloquial uses of the term. Taking the first approach, the process of arriving at it is defined by international law through *jus ad bellum*, whilst actions within wartime are, in theory, governed by *jus in bello*. The former sets out criteria that need to be met in order for war to be considered just and therefore permissible, and the latter aims to limit humanitarian suffering, for example through the Geneva or the Hague Conventions. However, war itself is not defined. Whilst IR "scholars generally define war as large-scale organized violence between politically defined groups", [191, p.351], there is also enormous variation "in theoretical orientation, methodological approach, ontological assumptions, and empirical domain" [190, p.581]. The type of war that is being studied has also changed, shifting from "the major powers to minor powers, from Europe to other regions, and from interstate war to intrastate war" [190, p.581]. Many of these debates draw upon history,

where the topic of war has been discussed at length in military theory and strategy. Work from antiquity is still being read, such as Sun Zi's *Bīngfǎ – The Art of War* or Caesar's *Commentāriī dē Bellō Gallicō – Gallic Wars*; the Early Modern Era had Machiavelli's *Dell'arte Della Guerra*, followed by Jomini and Clausewitz in the Napoleonic era. The latter's ideas, particularly those captured in *Vom Kriege — On War*, are still drawn upon.

On the other hand, colloquially, the word 'war' has been used far more broadly, such as in the context of anti-drug campaigns or the more recent 'War on Terror', which was the international military campaign launched by the US in the aftermath of the September 11 attacks in 2001. The term is, amongst related ones, used very loosely and often interchangeably, "undermining their value as conceptual tools in the process" [297, p.101].

2.1.1 Cyber Prefix

The prefix 'cyber' was added to war, in the past leading to doomsday predictions of cascading failures and disasters, using language such as 'digital Pearl Harbours' alongside scaremongering media articles. Stevens argues that "cyber security is sufficiently preoccupied with the uniqueness of the present that it tends to ignore both its own history and the historicity of the present" [294, p.126] except to see "'signs' that corroborate apocalyptic narratives and which confirm the likelihood of forthcoming cyber catastrophe" [294, p.126]. Yet, others argue that the past few years have seen a shift towards a more nuanced debate [97].

Another aspect to consider is Libicki's separation of operational and strategic levels of cyber war, with the former addressing cyber attacks supporting other military action, and the latter consisting of a "campaign of cyberattacks launched by one entity against a state and its society, primarily but not exclusively for the purpose of affecting the target state's behavior" [192, p.117].

Focussing on the strategic aspect, there has been a strong group of cyber war-sceptics [27], [114], [247], who in essence claim that the information change is (largely) without effect on the traditional frameworks, or just a mere extension of

practices [193]. For Betz, the term 'cyber war' is "not just a meaningless neologism, but strategically a distracting and nonsensical one" [27]. Whilst Kello agrees on the improbability of cyberwar and the consequences for traditional war fighting, he argues that the emergence of virtual weapons can be likened to the advent of nuclear weapons in the sense of being a revolutionary game changer, scrambling the international order in three ways. The least significant of these is systemic disruption, "which involves the appearance of a new technology that disturbs the regularized interactions of rational state contenders (because, for example, it alters the balance of offensive and defensive forces)" [167, p.13]; the second is systemic revision and "springs from the ascent of a state or a group of states that repudiates the shared basic purposes of the units and rejects the accepted methods of achieving them, in particular restraints on the objectives and means of war" [167, p.90]. Lastly, and most importantly, is the system change, which concerns "not the balance of power but the balance of players" [167, p.251], which "occurs when players alien to the system challenge the supremacy of the dominant units, whose purposes the new entrants may reject or may not even understand, but whose affairs they nevertheless significantly affect" [167, p.252].

Clarke and Knake discussed cyber war in the context of national security [65] and use a loose definition, by first perusing several historic anecdotes of attacks, including the alleged Central Intelligence Agency malware introduction to pipeline technology in the USSR in the 1980s, Israel's disabling of Syrian air defence in 2007, and the attacks on Estonia in 2007. Based on a US perspective, they believe that the authority to wage such operations should reside with the president, and that China poses the greatest threat. Clarke and Knake also caution that deterrence does not work in the way it did for nuclear strategy during the Cold War, and that in cyber war, it is not sufficient to consider an opponent's offensive and defensive capabilities. Instead, a broader view should be taken, addressing the dependence on networked infrastructure, for example, the internet. A core focus of their work is seeking to mitigate attacks along securing three prongs: infrastructure backbones, the power grid, and United States Department of Defense (DoD) networks.

2.1.2 Clausewitzian Links

The ideas of Clausewitz have repeatedly resurfaced over time, including in context of cyber war, particularly by the sceptics.

A reason to explore his work is to examine the driving force behind cyber weapons. For example, do cyber weapons have the utility, in and of themselves, to impose will, or are they merely an 'add-on'? The answer contributes to shaping perceptions of the relative value of these capabilities, particularly in comparison to other domains. Coupled with Clausewitz' defining elements of war, it can affect policy and planning decisions of whether and when to use these new capabilities. For example, his work could be used to argue that a certain capability has not reached the required threshold by lacking an element of violence.

In the chapter on the nature of war, he stated that "war is thus an act of force to compel our enemy to do our will" [68, p.75]. This definition, alongside his concepts of the fog and trinity of war has been called upon, whether to point out military blunders or when refuting changes in international theory. The 1990s and the increase of non-state conflicts gave rise to post-Clausewitzian thought, such as Kaldor's 'new wars' [161], who argued that it was time to move beyond the centuries old framework. Neo-Clausewitzian supporters on the other hand [99], [298], [299] sought to counter that by proving Clausewitz's ongoing relevancy, yet only few step beyond his *magnum opus* to do so [84], [276]. Lesser known work and sources, such as drafts, lectures and letters provide an additional perspective on war and warfare. This goes beyond state-centricity and opponent equality, including 'guerrilla' operations, as discussed in his *Vorlesungen über den kleinen Krieg – Lectures on Small War held in 1810/11* [67] and his *Bekenntnisdenkschrift – Memorandum of Confession* dated 1812 [66].

Clausewitz, his concept of war and its necessity for violence are often cited and applied to the cyber domain. For example, Rid argues that no cyber attack has been war, or an act thereof, based on his reading of Clausewitz [247]. He argues that none of the attacks conducted through malicious code have been potentially lethal, instrumental, nor a political act of force – nor will they ever be. Instead, Rid

addresses various attacks case-by-case, categorising them as espionage, sabotage or subterfuge, all activities as old as warfare, but not war. Nonetheless, he also notes that regardless of that belief, a cyber attack or conflict can result in being a triggering point, or an enabler for war. Betz [26] also argues that war has not changed and remains tied to the Clausewitzian trinity of chance, passion and reason. Yet, how we fight war has changed enormously, and that there "is now 'virtual dimension' to war that is comprised of digital networks and multimedia technologies and that changes many things" [26, p.8]. Since the messy 1990s, the West has "serially tried and failed to use technology to disconnect from war's enduring nature" [26, p.5]. Technological developments have aggravated certain aspects of how war is fought, "raising them to a greater salience than was the case before" [26, p.9], but they are not completely new.

Whilst others agree that cyber war is being 'over-hyped', they disagree with Rid [204], [297]. Stone for example argues this on the basis of force, violence, and lethality, stating that "war demands no necessary causal connection between what are really three distinct phenomena [...] all war involves force, but force does not necessarily imply violence – particularly if violence implies lethality" [297, p.103]. McGraw on the other hand believes cyber war to be inevitable, if security is not built into products given our increased dependency on technology. Furthermore, he challenges Rid's claim that cyber attacks had not yet caused physical damage [204].

Another factor is attribution, which is highly problematic in the cyber domain due to its nature, being physical yet lacking traditional dimensionality [28]. There are a number of perspectives and approaches on managing attribution, ranging from the technical or political to the legal and social, with only a few attempting to span these [49]. The former focusses on the challenges imposed by the underlying design, evolution, and application of technology taking a bottom-up approach [69], [333]. The latter often seek to address the larger picture, including more intangible factors such as motivation, interest, and capabilities, for example arguing that uncertainty can be reduced on tactical, operational and strategic levels simultaneously [248]. The majority do not attempt to connect the different domains. However, attribution

is not only an important step in seeking justice or compensation, but it can also play a key role in the prevention and deterrence of future attacks, though this is disputed [32]. It could also provide legitimate grounds for offensive (retaliatory) action being taken for purposes of self-defence. Without an independent and open group comprised of experts following standardised (and accepted) procedures, apparent culprits can easily be put forward. Any attribution would likely be challenged, undermining any verification and thus validation. Attribution still often remains "what states make of it" [248, p.7].

2.2 Cyber: Weapons, Malware & Co.

Weapons do not exist in a vacuum. In physical space, weapons have been created, improved, and used for attacks, crime, and warfare for millennia, incorporating technological advancement throughout the ages. Time has also affected how and who has the power to wield the weapons, as well as what is considered legitimate and not, regulated by international, norms, declarations, conventions, and laws. Yet, boundaries and legality remain ambiguous, both in cyber [41] and the other domains [175]. There has been some effort to use the term 'cyber capabilities' instead of 'cyber weapons' in order to clarify their different nature, and in order to encompass the broad spectrum of activities. Whilst this would alleviate some of the problematic surrounding the term 'weapon', it also implies an extended concept akin to a whole tool-kit. It could, for example, include hardware aspects, such as access to vulnerabilities or the supply chain; it could also include a skilled workforce, such as developers or *in situ* operatives. However, the focus of the thesis is on the software aspect.

Even if taking a simplistic definition of a weapon being an object centred on intentionality and harm, the cyber component complicates matters greatly [108], [125], [207], [265], with two elements standing out: a lack of physicality, and the nature of harm [295]. Weapons and attacks are intrinsically linked to harm, a concept that informs perceptions and reactions. Here, the concept of harm has to be extended to non-humans and can be thought of two overlapping sets: harm done

to and by digital systems [276]. The first set contains any form of external sources (intentional or not), such as physical damage, electro-magnetic pulses or even cosmic rays, doing harm to these systems. The overlapping section contains damage done by digital systems to digital systems – in essence software vs software. The second set includes damage done by a digital system that creates physical harm (to systems or humans). All three types have the potential to create social harm and 'collateral', including indirect varying degrees of psychological effects. Thus far, aside from proof-of-concept attacks under laboratory settings, such as the Aurora Generator Test at the Idaho National Laboratory in 2007, almost all publicly known cyber-attacks have fallen into the first two categories, with Stuxnet being the exception.

Although Stuxnet set a precedent for avoiding any collateral damage aside from spreading itself from system to system, future attacks may not display the same responsibility, particularly given the difficulty of attribution [41], [248], [333]. Two studies on cyber harm stand out as they focus specifically on the economic and social impact potential. The first is a 2009 study by the US National Research Council (NRC) that addressed ethical aspects, discussing the acquisition, as well as use, of cyber attack capabilities [218]. More recently, Agrafiotis et al. have done work on cyber harm, outlining concepts and creating a taxonomy in order to create a model for national cyber harm assessment and measurement [4], [6]. They consider it to be "the damaging consequences resulting from cyber-events, which can originate from malicious, accidental or natural phenomena, manifesting itself within or outside of the Internet" [4, p.2]. Based on their analysis, they divide cyber harm into six types (physical, psychological/emotional, economic, political/governmental, reputational and cultural) which are applied to four subject groups (individual, organisational, property/infrastructure and national), creating a Cyber Harm Model (CHM). The model can support, contextualise and validate cyber security policy, for example as part of the National Cybersecurity Capacity Maturity Model [120] but cyber weapons or related technologies are not discussed *per se*. Exploring harm in the cyber domain matters because unlike in other types of warfare [219], there is no standard methodology to identify and assess damage, particularly that of collateral.

Yet, due to the interconnected nature of the cyber domain, society as a whole is becoming a stepping stone or target, whether inadvertently or not, making this aspect highly relevant to the discussion on SaaW.

Multiple academic disciplines have taken to examining this topic, encompassing a large variety of concepts and definitions. This section aims to explore the most prominent ones, with a focus on understanding what the literature believe a 'cyber weapon' to be, and it is a direct starting point for the questionnaire in Chapter 3.

2.2.1 Conceptual Approaches

Rid & McBurney discussed cyber weapons conceptually, defining them "as computer code that is used, or designed to be used, with the aim of threatening or causing physical, functional, or mental harm to structures, systems, or living beings" [249, p.7], a subset of their weapon definition. It is also in line with broader weapon definitions that see these as being offensively driven with the design or intent to cause injury, death or damage [34]. Rid & McBurney further argued the importance of intentionality, as well as the difference between low and high potentiality, with the former being represented by, for example, widespread malware that nonetheless cannot fully penetrate the system and has low impact, whereas the latter is very specific and has a high effect on the chosen target. Accordingly, Stuxnet falls into the latter category, and shares similarities to the AGM-88 High-speed Anti-Radiation Missile – something very specific that is to some extent an intelligent agent and has a single purpose. Drawing the line between what cyber weapons are and what they are not is important for security, political and legal consequences. Yet, they did not analyse any cyber weapon's components in detail but merely its purpose, thus coming to the conclusion that Stuxnet is a weapon (it causes/has the potential to cause damage) but Duqu is not (it merely gathers intelligence). This definition is somewhat questionable given the high value of information in modern society, as can for example be seen in the 2018 Facebook–Cambridge Analytica data scandal [306]. The connection to physical weaponry is, however,

appealing, and its applicability is also explored in the questionnaire soliciting public opinion conducted as part of Chapter 3.

Kello on the other hand believes that the "crucial definitional criterion of a virtual weapon lies in its intended and possible effects" [167, p.49] and that the weapon label "is nevertheless appropriate in view of the fact that some forms of code can cause significant political and economic harm even if they are nonviolent" [167, p.50]. An alternative approach centres on specific systems, such as Peterson's work on Industrial Control Systems (ICS) [237]. His discussion focusses on the development, construction and employment of these at a high level. He uses Stuxnet as an example, highlighting the flaws and dangers of the ICS environment, such as legacy systems and lack of focus on security but also does not delve into detail. Peterson uses a three-tiered classification ('simple weapon' / 'moderate attack tool' / 'complex cyber weapon') whilst emphasising the need for experts and/or experience in areas other than CS for the latter two. He concludes that ICS are easily damaged due to the insecure nature of the systems, but that persistence is far more challenging.

Owens et al. [231] on the other hand were amongst the first to comprehensively research policy dimensions and legal/ethical implications, particularly of offensive aspects. Written from a US perspective, their work was informed by the underlying technologies whilst drawing references to the nuclear debate, arguing that "today's state of affairs regarding public discourse on cyberattack is analogous to the nuclear debate of 50 years ago. At that time, nuclear policy issues were veiled in secrecy, and there was little public debate about them. [...] seminal pieces did much to raise the public profile of these issues and stimulated an enormous amount of subsequent work outside government that has had a real impact on nuclear policy" [231, p.xi]. In the decade since, there has been more public debate and research on cyber issues, however, work on fundamental concepts is still limited. The focus of their work was also to differentiate between a cyber attack being a matter of law enforcement, versus one of relevance to national security, which is still relevant to today.

More recently, Herr [141] expanded classification systems for cyber weapons and malware, as well as their components. He addressed the different components in

each before comparing existing definitions to conclude with the proposed PrEP framework, representing 'propagation method' (Pr), 'exploit' (E) and 'payload' (P), which is reminiscent of Owens' 'Vulnerability, Access, and Payload' [231, p.83]. Drawing on previous work on propagation types and payloads, Herr defines cyber weapons as "the combination of a propagation method, exploits, and a payload designed to create destructive physical or digital effects. Unauthorized access to data does not damage integrity, which means software designed for espionage is malware but not a weapon" [141, p. 10] and addresses the shortfall of previous definitions. One criticism is based on the necessity of an exploit in his definition in order to be classified as a weapon, which it not always necessary.

In his 2015 work with Armbrust, they coined the term 'milware' seeking to distinguish state authored malware from non-state authored by reverse engineering several samples in an interdisciplinary effort [143]. They create a 'MAlicious Software Sophistication', short MASS, index with main categories of propagation (to the victim and within a network), severity of vulnerability used, as well as level of tailored payload. Their core set of differentiating features could be summarised as being a matter of sophistication: malware, in comparison to milware, is less tailored, has broader functionality, and is more indiscriminate [143, p.36]. Whilst an interesting pilot project, particularly as it crosses disciplinary boundaries and because it formalises colloquial understandings of 'sophistication', it would be relevant to pursue this further and see if or when a state actor may decide to opt for malware over milware. Similarly, it would be intriguing to explore whether some state actors are limited to producing malware due to their constrains on knowledge and resources – if so, how would that fit into the framework?

2.2.2 Legal Perspectives

Legal scholars also wrestle with defining cyber capabilities, seeking to arrive at explicit norms, regulations or even laws. These additional perspectives further broaden the understanding of what SaaW is and is not. Yet, as will be seen below, there are certain similarities, such as the question of attribution. However,

the perspectives are also diverse, similar to the varying conceptual approaches discussed above.

Tsagourias has explored what conditions have to be met by a cyber attack in order to fall under a state-actor's right of self-defence [308], while the Tallinn Manual on the International Law Applicable to Cyber Warfare (TM) sets out 95 rules to govern cyber conflict, addressing *jus ad bellum*, state sovereignty and responsibility [266], [267]. Although the TM is not binding, it provides the most complete and coherent framework addressing aspects of cyber war. Written from a predominantly western perspective as *lex lata* (and therefore implying that current international law is applicable to cyber warfare), TM attempts to include technical expertise and emphasise consistent terminology. Cyber weapons in particular are defined in Rule 41 as "cyber means of warfare that are by design, use, or intended use capable of causing either (i) injury to, or death of, persons; or, (ii) damage to, or destruction of, objects, that is, causing the consequences required for qualification of a cyber operation as an attack (Rule 30)" [267, p.141-142]. The TM further differentiates between the means, e.g. a computer, and the infrastructure used as a delivery mechanism, e.g. the Internet, whilst only concerning itself with activities above the 'use of force' as defined by *jus ad bellum* – taking an effects–based threshold. This level is also upheld in regards to intelligence activities, therefore rendering it mostly inapplicable to corporate espionage, intellectual property theft or other criminal activities – all which may pose a threat to states but are to be covered by existent law [267, p. 4]. However, due to the approach as legal advice, the document appears to sidestep two important issues: the attribution problem as well as the 'grey-zone', for example state sponsored criminal activities or attacks, although two test cases (*Nicaragua* and *Tadić*) were discussed to understand the issue of 'overall control'. Furthermore, the role and protection of civilians (or non-combatants) can be difficult to manage due to the blurring of lines.

Another legal view is presented by Brown and Metcalf [43], who seek to apply the theoretical discussions into concrete legal advice, stemming from experience of academics colliding with practitioners. Their caveat, however, is that they do not

attempt to differentiate state and non-state actors, due to technical difficulties in distinguishing these – in essence the attribution problem. They present a number of case studies and explain the difficulties in conducting an actual review of cyber weapons, arriving at the conclusion to use the same definition for a cyber weapon as for a normal one: "an object designed for, and developed or obtained for, the primary purpose of killing, maiming, injuring, damaging or destroying" [43, p.135]. A slightly wider view is taken by Blake [31], as he does not limit himself to cyber weapons but seeks to assess whether *all* modern approaches to weapons should be reviewed, first by examining the scope of legal reviews themselves in order to define what "effects, designs or intents must be considered" [31, p.161]. Similarly to the TM, he also addresses *jus ad bellum* but specifically delves into thresholds and concepts of the *use of force*, as well as what implications arise from defining a capability as a weapon. Blake concludes that "those non-lethal bloodless capabilities that can be applied to a military object or enemy combatant should be subjected to a legal review before that capability is used" [31, p.162].

Agreeing with Brown and Metcalf [43], Wilson [337] believes that more clarity and applicability is needed, particularly due to the attribution problem and unknown intentions, as well as due to the potential of cascading effects. He has summarised four common characteristics to cyber weapons, which are based on code analysis and reverse engineering. However, he does not specifically state which code analysis was considered, although he does mention Flame, Duqu, Stuxnet, as well as the alleged steel mill attack in Germany in 2014. Furthermore, he asserts that code can be labelled a cyber weapon if the following characteristics are combined: use of a ZDE; use of a coordinated campaign of malicious programs for espionage, theft and sabotage; use of stealth to prolong malicious operations; intimate knowledge of the targeted system [337]. Although this definition appears to be more specific, the attribution problem is not better defined through his analysis. Additionally, it rules out a large proportion of attacks, for example those that do not include ZDEs or do not (*yet*) have intimate system knowledge.

2.2.3 Technical Angles

Within CS, the term 'cyber weapon' is rarely, if at all, used. Instead, the focus is on various forms of malware and elements thereof, such as worms and viruses, which have been discussed for decades and are well understood. Since the IoT era, some work has centred on various aspects of detection and defence, such as insider patterns [5], botnet taxonomies [168], Denial of Service (DoS)/DDoS research [241], specific vulnerabilities in standards and targeting [130]. Others address characterisation based on code analysis and reverse engineering [337], command and control aspects [309], as well as situational awareness [310]. Attempts have also been made to categorise attacks by examining the various *modus operandi* in order to create a classification and provide a defence base [311]. Alternatively, various data mining techniques and neural network approaches for four main attack classes consisting of DoS, Remote to User (R2L), User to Root (U2R) and Probing [89] have been reviewed. Although a standardised system does not exist, insights like this also further knowledge about technical capabilities, which are an essential element of understanding SaaW alongside of prominent cases, such as Stuxnet.

Many consider the widely discussed Stuxnet worm discovered in 2010 to be the first instance of an actual cyber weapon, blurring the lines of cyber and physical like none before [157], [187], [345]. It "provides a prime example of a cyber weapon expressly designed and used for cyber war" [204, p.114]. It was designed to attack industrial PLCs at the Iranian nuclear facility in Natanz, and whilst it succeeded, some believe that the impact on the uranium enrichment operation has been overstated [19].

Stuxnet utilised several ZDEs, for example one pertaining to the Print Spooler Service [216], something that had not been seen previously, and had a multilayer operation attacking first the Windows OS, then specific Siemens software for industry, followed by additional Siemens S7 PLCs. It is the first publicly known instance of malware to be released into the wild, finding a target, going into stealth, sabotaging, and deleting itself. An estimated 30 people were involved in writing the code over a period of approximately six months. Reverse engineering and

understanding Stuxnet was equally as challenging a feat, and drew on a large number of resources, with experts around the globe working on various aspects, but Ralf Langner being the first to 'join the dots', who has since published a detailed technical report on the subject [182]. A few other aspects are noteworthy: firstly, great emphasis was placed avoiding collateral damage, in this case computer and networks that did not meet the specified criteria. The code intruded on a great number of systems worldwide, including personal computers, which it then used to spread further, but did not execute nor harm the system; secondly, it had a 'fail-safe' of limiting its spread to three other devices and a self-deletion date of 24 June 2012; thirdly, different versions have since been discovered, with earlier ones being less aggressive in their propagation.

Another factor that should be noted is the concept of modularity, as this can affect the cost and the time it takes to create variants. This can, for example, be seen in the similarity of Stuxnet and Duqu: both share the same architecture, which in simplified terms contains a driver file that loads the main module, a separate configuration file and an encrypted block in the system registry, and they are further linkable via drivers used [124]. Other related files have been found in Flame/sKyWIper but the differences are overall vast, including those in later malware, such Gauss/Gödel that included encrypted 'warheads' [24].

Another attack that saw code re-use was that on Saudi Arabian oil company Saudi Aramco in 2012, affecting roughly 30,000 workstations and requiring about ten days to 'clean up' to fully restore services. Whilst the attack received a lot of media attention, it did not reach any of the company's critical systems; it also did not attempt to exfiltrate data, or attempt to cause any physical destruction of their industrial processes, unlike Stuxnet; instead, it overwrote data stored (reportedly) replacing it with an image of a burning US flag and requiring Aramco to disconnect remote access to certain sites. A type of malware discovered just prior to the attack, named the 'Shamoon Wiper', had similar capability: consisting of three parts (the dropper, the wiper, and the reporter) it corrupted files on the computer it compromised, overwrote the master boot record, and sent the information to

a different internal machine on the network. The attack appeared sophisticated, predominantly due to its spread, but analysis showed it was rather basic and did not require extensive skills nor resources [301]–[303].

A third example is BlackEnergy, which in its third iteration attacked the Ukrainian power grid in December 2015, with traces also being found in DragonFly and TeamSpy attacks [232]. Furthermore, malware authors not only use elements of others, but also often use the overall design, such as seen in Figure 2.1: GreyEnergy used an early version of NotPetya in 2016, while also having a similar framework to BlackEnergy in terms of modularity and a stealth emphasis, albeit without specific ICS targeting and a greater focus on espionage and reconnaissance [56].

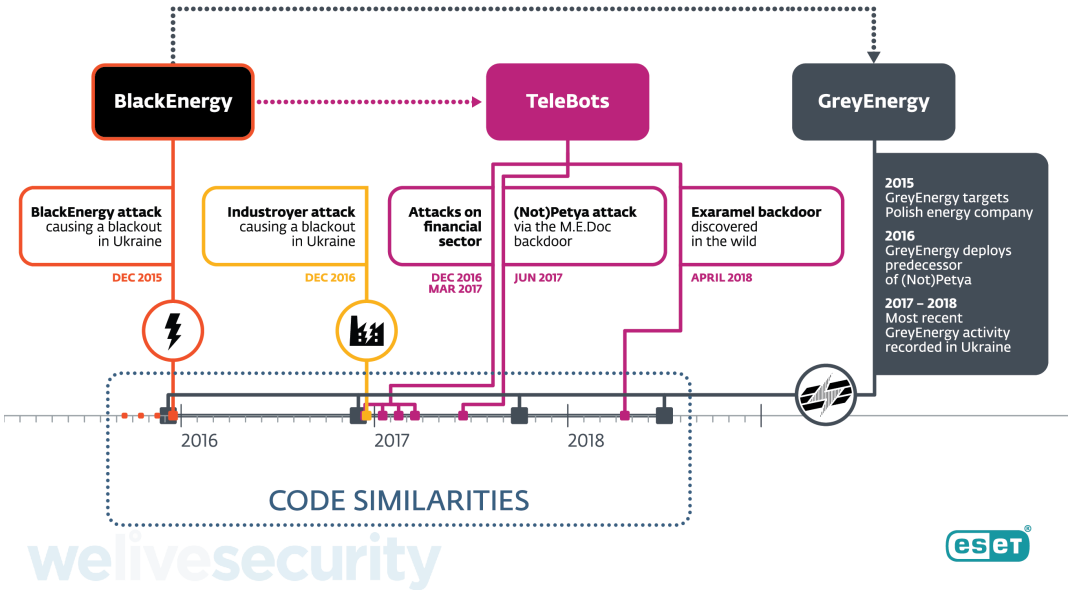


Figure 2.1: Relationship between BlackEnergy, TeleBots and GreyEnergy. Source: [56]

2.3 Cyber: Arms Control

The idea of controlling arms goes back centuries, with example including Charlemagne who sought to limit the possession of high-quality swords and armour produced in the Frankish empire. The modern industrial age saw a treaty between the US and the UK, effectively creating demilitarised zones [17], followed by milestones such as the Hague Convention or the Geneva Protocol. After World War II, various treaties followed, the most prominent including the Non-Proliferation

Treaty of Nuclear Weapons (NPT) 1968, the Strategic Arms Limitation Talks (SALT) in the 1970s and the Anti-Ballistic Missile Treaty (ABMT), the INF in 1987 and the Chemical Weapons Convention (CWC) in 1993, two Strategic Arms Reduction Treaty (START) and the Comprehensive Test Ban Treaty (CTBT).
One of the best known definitions of arms control stems from the 1960s by Schelling and Halperin, who saw it resting "essentially on the recognition that our military relation with potential enemies is not one of pure conflict and opposition, but involves strong elements of mutual interest in the avoidance of a war that neither side wants, in minimizing the costs and risks of the arms competition, and in curtailing the scope and violence of war in the event it occurs" [264, p.1]. Arms control seeks to limit escalatory dynamics by curtailing weapon systems in type and/or number, whilst its preventative form "can be seen as qualitative arms control applied to the (near or far) future [...] cutting off the military-technological innovation process as early as possible" [10]. Furthermore, this type can also "aim for mutual understandings on legitimate or illegitimate uses of emerging technological capacities – be it as complement or substitute to limitations on deployments, designs and numbers" [133]. Given a history of technological innovation or geopolitical dynamics driving states to Confidence Building Measures (CBMs) and arms control for international stability and security, it was only a matter of time before concepts would be applied to the cyber domain, particularly in the wake of Stuxnet that rekindled fears of a cyber arms race [78], [285]. Although some have suggested that cyber CBMs in this area could "blunt some of the factors that contribute to crises and escalation" [35], "an overwhelming majority of researchers in that field have all but given up" [45], [201]. The cyber domain has given rise to additional challenges and exacerbated old ones beyond obvious political obstacles, including trying to define what cyber arms even are, questions of dual-use, and verification issues [269]. Even despite several definitions of cyber weapons, such as Rid and McBurney's [249], none of them are "juridical or legal terms, as there is no consensus definition of either weapons or cyberweapons in international law" [295]. Another key problem is that traditional incentives and reasons of historical agreements do not apply to the cyber domain,

from the expense and accessibility to the potential of sanitising warfare [13]. A key question that also remains is what the exact object of regulation in any type of regime would be [269]. Yet, notwithstanding this hodgepodge of issues, there have been numerous attempts at regulation, which together could be viewed as "an emerging global governance architecture for cyber weapons" [295]. After all, without openness a "system of acceptance, cooperation, technology, and traditional investigation supported by policy, law, and treaty" [152, p.27] cannot be created. This includes a variety of work that focusses on arms control, such as examining the applicability of the CWC [115] or nuclear equivalents [16], conventional arms treaties [201] or arms control more broadly [269]. It also encompasses other processes, such as the UN Charter's cyber relevancy [331], the TM [266], export controls via the Wassenaar Arrangement [330], and the various 2018 multi-stakeholder initiatives, such as the Paris Call [239]. Several initiatives and their relevance to the cyber domain are discussed below.

2.3.1 Dual Use

The concept of dual-use is well understood, with a broad definition being goods, software and technology that can be used for both civilian and military applications. Different agreements and approaches seek to regulate this broad area of dual-use, for example by handling import/export or the create and use of dual-use goods. Similarly, code fragments or whole software suits can be used in various ways. Aside from malware examples discussed above, packet analysers such as *Snort* or *Wireshark* can be used for network troubleshooting, optimisation, misuse detection, protocol development, and education. At the same time, they can be used for nefarious activities, such as MITM attacks. There is no differentiation apart from the use itself. Another example is DDoS, which is primarily known due to its offensive capabilities, such as on Estonia in 2007. Yet, it can also be used by security professionals to stress test an organisation's preparedness, albeit in a controlled manner [270], [304]. Due to this inherent dual-use nature, "it would be nearly impossible to catch a country cheating 'red-handed' " [42, p.57].

In light of this, Herr evaluated export controls in the context of malware [140]. He found three key issues, namely reduction of beneficial research, minimal impact on malware and difficulties of multilateral coordination, which would be better addressed by increasing the incentives surrounding vulnerability research. His further work centred on a malware proliferation definition [142], which is the process of learning and reusing code between different actors, ranging from collaboration and direct support to unintentional sharing. Furthermore, he asserted that state actors not only lack a monopoly but are also not the driving-forces of innovation. Whilst this may be the case for a large proportion of malware, highly targeted and effective attacks still require a large amount of resources, from reconnaissance to test-beds and dissemination. Herr advocates reducing vulnerabilities, however, it will not stop those actors sufficiently motivated and resourceful, nor does it address the stockpile of vulnerabilities governments keep to themselves for future use given the rediscovery probabilities [144].

On the topic of export controls, the Wassenaar Arrangement has to be briefly addressed. This voluntary regime shares information on conventional weapon transfers as well as dual-use goods and technologies, which inadvertently discouraged and stifled research, while seeking to improve security. In late 2013, a proposal was made to modify and extend the arrangement to include malware and its related technology [12] but initial issues were found with the language, limiting beneficial research and information sharing [39], which had a chilling effect on the research community. Positive first changes were made in 2017, though discussions are still ongoing [330].

2.3.2 Treaties and Initiatives

Whilst there are numerous arms control treaties, a number of them have been explored in light of applicability to the cyber domain. They are relevant in light of discussing SaaW because they attempt to deal with weapons of a specific kind, either by defining the weapon or its use. On the one hand, this supports Chapter 3, by highlighting different approaches to dealing with weapon technologies, and

Chapter 4 on the other, as it provides reference to frameworks used to limit the proliferation of other weapons.

The Biological and Toxin Weapons Convention (BTWC) was created with the aim of supplementing the 1925 Geneva Protocol that prohibited the use, but not development or possession of biological and chemical weapons, and it entered into force in 1975. In 2019, 183 states were party to it. Unlike the NPT or the CWC, this convention does not have a formal mechanism to monitor or verify compliance – two elements that are highly relevant to the cyber domain. The BTWC stands out due to its so-called *general purpose criterion* of Article I, which does not prohibit items, such as a specific toxin or agent, but instead their intended use. Thus, small amounts for medical or defensive purposes are allowed, but large quantities or their delivery mechanisms are not. This criterion combined with the convention's unlimited duration make the BTWC future-proof, as all subsequent discoveries or developments are included. Given that this Convention addresses not only dual-use technology but weaponisation potential while supporting peaceful use, it appears to lend itself well to the cyber domain. However, there are several key differences that have been highlighted previously [42], [221]: cyber attacks and associated malware in its various guises have become ubiquitous; bio-weapons and related elements were considered to provide little to no security for state actors, which cannot be said for cyber equivalents; code cannot be as easily destroyed, nor can stock-piled vulnerabilities; imposing licensing and/or regulation is far more challenging given the global spread; civilian and commercial infrastructure in the cyber domain far outnumbers its biological counterparts.

The CWC on the other hand came into force in 1997, after almost a quarter of a century of difficult negotiations and is, similar to the BTWC, an egalitarian treaty. However, unlike the BTWC, it has an invasive verification regime combined with three schedules of clearly defined substances that are treated differently: the first includes chemicals that have few, or no uses outside chemical weapons and they are restricted to very small amounts for research, medical, or defence testing purposes; the second includes chemicals that have legitimate small-scale applications, requiring

declaration and export controls; the last includes the chemicals that have large-scale non-weapon use, and only large production plants have to be declared. Within each schedule it further differentiates two parts, one accounting for toxic chemicals and one for precursors. This convention bans the development, stockpiling and use of chemical weapons, while also requiring members to destroy their existing ones. Regular inspections are managed by the Organisation for the Prohibition of Chemical Weapons (OPCW), akin to the International Atomic Energy Agency (IAEA), and 'surprise inspections' are allowed based on substantiated complaints. The CWC does not lend itself well to the cyber domain, mainly due to its specificity of listed items, which cannot be applied to software, the difficulty of verification and the exhaustive element preventing it from being future-proof in case of new discoveries or developments. Geers [115] examined the CWC with a focus of applying it to the cyber domain. He found five important lessons that need to be considered based on "political will, universality, assistance, prohibition, and inspection" [115, pp.54-549], the first three of which are transferable, the latter two not. He concluded that certain elements are shared, such as acquisition ease and asymmetry, but "the tactics, strategies and effects are fundamentally different. Chemical warfare kills humans; cyber warfare kills machines" [115, p.549].

The number of cyber-specific initiatives is growing steadily, whether bilateral, EU centric or multilateral. However, thus far, the Council of Europe's Convention on Cybercrime, also known as the Budapest Convention, is the only international treaty that requires ratifying states to criminalise certain behaviour. Whilst a summary of various motions can be found in Carnegie's *Cyber Norms Index* [51], several key efforts are discussed below. Pre 2009, the US shunned processes seeking to strengthen cyber security at the global level, such as those by Russia via the UN Committee on Disarmament and International Security. After a shift in policy, this allowed for the GGE process to reach an agreement on several recommendations, which were nonetheless still hampered by fundamental differences in perspectives and interests. The topic of setting voluntary norms to draw an upper boundary

on aggressive cyber activities and to reaffirm the applicability of international law [312] has been contentious, as has the inclusion of LOAC stalling the process.
Several cyber-specific multi-stakeholder initiatives were launched in 2018: the Siemens-led Charter of Trust was signed at the Munich Security Conference (growing to 16 members in 2019), focussing on cyber security standards of supply chains [274]. The Paris Call on the other hand brought together state actors and industry, pledging their support to improve security in the cyber domain, which, as of November 2019, is supported by 74 states (notably absent are China, Russia, and the US), 24 local governments, 333 organisations, as well as over 600 private sector entities [239]. The Paris Call reaffirms the applicability of international law and seeks to bolster numerous forms of resilience measures, ranging from protecting against election inference and intellectual property theft, to securing the whole product life cycle, as well as developing ways to prevent the proliferation of malicious tools and practices [307]. Whilst these initiatives benefit from private actors and show industry interest, substantial progress requires interstate cooperation [133], such as the renewed commitment in the form of the twin dialogues at the UN, focussing on norms and responsible behaviour. Furthermore, whilst this might be an encouraging first step, this effort is voluntary and not legally binding, with many obstacles yet to overcome.

2.4 Conclusion

This chapter presented a literature review to discussing three key areas that are relevant to this thesis in light of the cyber domain: war and warfare, weapons, and arms control. It highlighted the over-arching debate, presenting the key arguments of those that believe cyber war will, or will not, take place. This was followed by delving into the debate surrounding cyber weapons and related aspects, showcasing conceptual, legal, and technical approaches. Finally, it addressed arms control and works specifically pertaining to the cyber domain, including multi-stakeholder initiatives.
Based on the current literature, there are several noteworthy points. Firstly, the concept of cyber war is still debated, not only in its definition but also its

interpretation. For some, it presents a completely 'new' age, whereas others believe that the nature of war itself has not changed, but the means of waging it have. The questionnaire in the following chapter adds the perspective of different population groups to the debate, discussing aspects of (in)security and state centrality. Furthermore, this topic is also picked up in Chapter 4, which explores the reason behind actors pursuing weapons, including a discussion on motivating and restraining factors.

Secondly, the understanding of what a cyber weapon is varies greatly, including different terminology ranging from virtual weapon to malware. Yet knowing, defining, and agreeing upon definition of weapons is vital, not only for scholarship, but also for the design of effective policy and strategy. For example, if there had not been a mutual understanding of what ground-launched ballistic and cruise missiles were, the (now expired) INF treaty could not have come into effect in the first place. For the cyber domain, this common understanding is even more vital, as purely technical definitions are not sufficient: not only is the cyber domain highly interconnected and dual-use centric, but it is also still evolving rapidly. Arriving at, and agreeing upon, certain definitions may create a foundation for communication and acceptable behaviours, which in turn may solidify into a framework of norms. The next chapter seeks to investigate this further, exploring various understandings of what a weapon is in this context. For example, does the view of the general public differ from the literature? What about specific subgroups, such as military officers?

3
Opinions on Weapons

Contents

In as young and rapidly evolving field as cyber security, fundamental studies aimed at describing constructs and understanding the perception of important concepts in different populations are paramount. As seen in the literature review in the previous chapter, not only is the overarching concept of cyber war still greatly debated, but so is the understanding of what a weapon in the cyber domain is.

Yet, this is vital to reasoning about why actors seek to pursue this technology, which is explored in Chapter 4.

This chapter presents the findings of an opinion survey on SaaW, showing significant differences in the interpretation and perception of three basic constructs, depending on the background and training of respondents. This underlines the population heterogeneity in the understanding of the potential of software and malware of SaaW, which in turn contributes to preventing misunderstandings, reaching agreement on definitions and comprehending the potency of adversaries in cyberspace.

Currently, cyber weapons, digital weapons, virtual weapons, Advanced Persistent Threats (APTs), and SaaW are terms that are used frequently and interchangeably [43], [141], [167], [249], [267]. Yet, they remain highly ambiguous, with requirements and definitions varying even further depending on the author. A very loose consensus of criteria seems to be a type of malicious software that is a) employed against a specific target; b) has a very high level of sophistication; and c) is sponsored by an actor (state or non-state). But what is sophistication in this context? Computational complexity of the malware? Accessibility of the target system? Resources used? And who gets to define it as a weapon? Academia? Governments? Military Forces? What about public opinions? Although several papers, books, blogs, and other media articles have been written on the broad area, ranging from technical analysis and case studies to potential consequences for the international system, surveyed public opinion remains absent. However, such a survey is highly important: each individual can be affected, whether as a target, a steppingstone to the target, or mere collateral. Equally, an individual can provide technical know-how or be merely an unwitting provider of resources and connections, thus acting as a key component. Lastly, an individual can also affect public policy via debates and elections. Without understanding the various perspectives, we cannot arrive at definitions or understand the driving forces behind actors seeking to pursue this technology.

Interconnected digital systems are vital to modern society, from personal use for communication and consumer transactions, to national energy supplies and security or international banking and trade. In May 2017, a global cyber attack intruded

upon more than 200,000 systems in 150 countries within the first two days alone. It affected the National Health Service in the UK, Deutsche Bahn networks in Germany, Telefonica in Spain, and the Ministry of Interior in Russia, to name a few. Experts warned of more imminent attacks and Europol's chief Rob Wainwright stated that the "world faced an escalating threat"[22]. The attackers' weapon of choice was ransom-ware called 'Wanna Decryptor', also known as 'WannaCry', allegedly utilising the EternalBlue exploit developed by the NSA for a vulnerability in Microsoft Windows machines. The advent of the IoT sketches out an even more data-driven and interconnected future, encompassing civilian aspects such as smart cities or military applications, with lines blurred far more extensively than before. Given the varied views and groups of people that can be affected, the questionnaire discussed in this chapter is designed to explore opinions on software, malware, and weapons across a total of 46 questions including consent and demographics. It splits respondents by background and training into three groups: *Military*, *Academia* and *Other*. Unsurprisingly, the opinions on what a weapon is varied greatly, for example when deciding whether malware is a weapon, or if a threshold is needed for this to occur. Furthermore, those in *Academia* are more likely to disagree that software should be treated like any physical object, *Military* respondents are more likely to agree and *Others* are indecisive – with the differences being statistically significant. Yet, there is also consensus, for example regarding the need for software/malware regulation, or a strong concern regarding enforceability.

The remainder of the chapter is structured as follows: the importance of public opinion is highlighted prior to a discussion on questionnaire design and analysis methods used. This is followed by a presentation and examination of results, as well as the introduction of a factor model to analyse the questionnaire structure and provide suggestions for future improvement. Lastly, the chapter concludes with lessons learnt.

3.1 Public Opinion

Regarding the importance of the individual opinion of SaaW, an analogy can be drawn to a discussion held in the nuclear domain several decades ago. In the early 1980s, the Harvard Nuclear Study Group published a book entitled 'Living with Nuclear Weapons'. The first chapter addresses issues in the nuclear debate, highlighting the importance of the individual as "each citizen is not just a target of nuclear weapons; each is also an actor in the nuclear drama" [135] before giving examples. This is also very true of the individual and SaaW.

As the nuclear domain has seen the use of questionnaires and surveys since 1945, a wealth of (historic) data has been created. The first series of those were carried out in the USA and addressed far-ranging issues, such as chances of nuclear war, questions about 'first use' or fall-out shelters. They also canvassed aspects of the arms race, weapon development, effects on security and the international balance. The wording, frequency and respondent base naturally varied over time, yet sufficient similarities can be found to ascertain common themes and perform analysis, an example of which can be seen in Kramer et al.'s work [174] from 1983, which evaluates nuclear surveys over 40 years. This has motivated the creation of a questionnaire for software, malware and SaaW, which is presented below.

3.2 Experimental Design

The formation of the questionnaire followed an iterative process, taking literature surveys, supervisor, and colleague discussions, as well as feedback from conversations into account and it was approved under the CUREC of the University of Oxford. A pilot test was conducted with eight people to ensure the questionnaire was well set-up on the Bristol Online Survey system, and that the data was collected, stored, and became accessible as intended.

The questionnaire itself was online-based and was shared through a variety of means, such as social media, direct emails, and word-of-mouth. It had an age requirement of 18, and the participants had to consent to their data being gathered

anonymously. Aside from sourcing attitudes towards a SaaW, there was a contention that the work sector/experience will lead to different responses, in this case groups of *Military*, *Academics* and *Other*. Age and experience within computer science or international relations were also analysed.

Military personnel were chosen for two main reasons: firstly, regardless of their specialism, they will have been trained extensively in the use of weapons, as well as the responsibility and consequences of their use. Secondly, the respondents are all officers, which places them in positions of additional responsibility. *Academics* respondents on the other hand are often considered to have access to, and influence on, decision makers, thus also giving them an important role in forming opinions. Overall, the questionnaire contained 46 questions (Q1 being consent), and only fully completed ones were used for analysis, amounting to a total of N_{Tot}=96. A further 65 began the survey but did not complete, resulting in a completion rate of ~60%. It was divided into three main sections, following demographics and familiarity with the topic (Q2-Q7): the first part centred on weapons, their nature and constitution (Q8-Q16); the second part focussed on software, malware and understandings thereof (Q17-Q34); the final section brought SaaW together against the backdrop of international security, including capabilities and proliferation (Q35-Q46).

3.2.1 Data Analysis

The majority of questions used the common five-point Likert rating scale, allowing for a 'neutral' response. The responses are not interval variables – the strongly disagree category does not, for example, imply a disagreement twice a strong as disagree – and the sample sizes are different. To ensure the robustness of the statistical analyses, Mann-Whitney U (MWU) tests were used to test differences in value ranking between two independent samples, as well as the non-parametric Kruskal-Wallis H (KWH) tests, performed to test median differences for more than two independent sample groups.

The open-ended questions allowed for free-text responses but would pose a greater challenge for analysis. For example, the term 'gun' can be understood in several

different ways: colloquially it is used interchangeably with 'pistol', 'side arm', 'firearm' or 'rifle' and others; it can also refer to pellet, water or other toy guns; in the military on the other hand, it refers to weapons that require a small group of people to use, such as artillery or a heavy machine gun. For the purpose of this thesis, responses were grouped and labelled as 'firearm' to denote the use of an explosive charge to accelerate a projectile. The qualitative analysis conducted per question will be discussed in the relevant section.

A PCA was later performed to better understand the structure in the questionnaire conducted. PCA identifies blocks of highly correlating items and combines these into components that explain portions of the sample variation. Firstly, the factor structure is examined, followed by the dropping of items, leading to a final comparison of groups across the dimensions. A vast number of qualitative surveys, particularly in the social sciences, have used this approach over many decades. This includes the seminal work on depression known as the Beck Inventory [23] and a recent evaluation thereof [320], its slightly less successful counterpart [132], personality tests [76], [203], ageing studies [272] or work by the World Health Organisation [335]. More recently, it has also been discussed a book on correlational research [80] and papers surrounding questionnaire evaluation [145]. The PCA was performed on all 41 items (questions) of the questionnaire including sub-parts but excluding demographics, fitting a model with factors which were allowed to correlate (Oblimin rotation). Sampling adequacy was assessed by evaluating the covariance among variables with the Kaiser-Meyer-Olkin (KMO) test.

3.2.2 Respondent Profiles

All but two respondents provided their nationality, with the largest number of respondents identified as British (59), followed by Norwegian (7), Dutch (6) and German (5). The remaining respondents (19) were two each from Denmark, Greece, Ireland, Saudi Arabia, and the USA; one respondent each was Australian, Belgian, Croatian, Finnish, Pakistani, Spanish, and Swiss. One respondent stated dual nationality and was grouped with his first listed country.

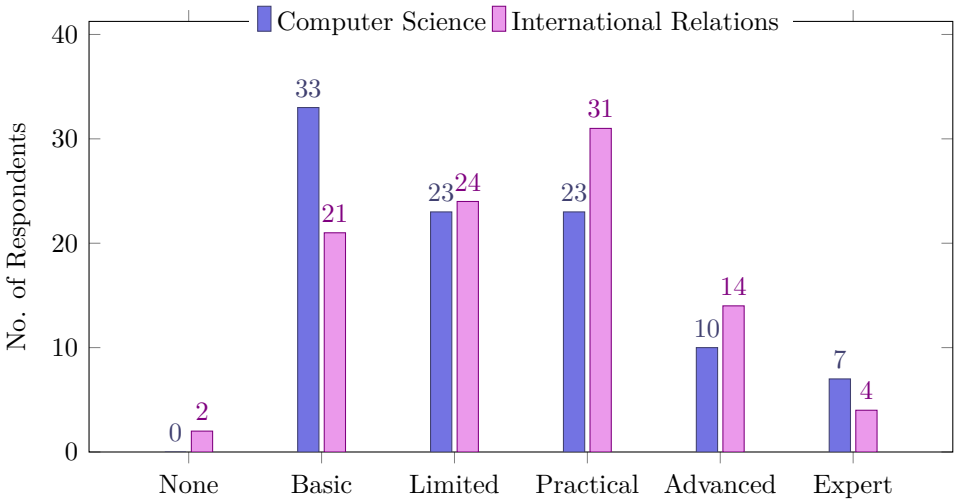


Figure 3.1: Level of Expertise

The respondents were divided into three groups for analysis: *Military* (N_{mil}=38), *Academia* (N_{aca}=19) and *Other* (N_{oth}=39). It should be noted that although the military personnel represent the largest number, their roles vary greatly, ranging from legal, to nursing or intelligence aspects. Within the *Other* group, the IT sector is represented most (N_{it}=10), the remainder is spread widely, ranging from education and finance/banking to hospitality, construction, and Non-Governmental Organisation (NGO).

Two further preliminary questions were asked to self-rate the familiarity with the related topics of CS and IR; results are shown in Figure 3.1, showing a diverse range of backgrounds.

3.2.3 Bias

In this type or research, bias may arise from various sources, such as response, interpretation, or sampling. It is beyond the scope of this paper to delve into detail, particularly as dedicated discussions on the topic exist, for instance [57], [109]. Here, examples of bias can include the (construct) validity of items and their response between different cultures and/or nations, the objectivity of norms, thresholds, and groups (three in this case *Military*, *Academia*, *Other*), the efficiency of the questionnaire measuring our constructs, and the unknown reliability of observed

responses (e.g. over time or split half reliability). Effort has been made to addressing the weaknesses and limitations of this questionnaire, as well as the resulting analysis, in the various sections. An in-depth psychometric evaluation, however, could be an interesting contribution for future work.

3.3 Results and Discussion

A summary of the questions and responses grouped by question can be seen in Tables A.1 and A.2 in Appendix A. These exclude the first seven questions on demographics.

3.3.1 Software & Physicality

The opinions on whether software should be treated like any physical object or tool were divided. Based on background, there is a significant difference, with the *Academics* group tending to disagree more, the *Military* group agreeing and the *Other* group being indecisive. Performing a KWH test underlined this statistically ($H(2)= 6.44$, $p=0.04$); using for *post hoc* analysis, *Academics* to *Military* differences remain statistically significant ($U=249$, $p=0.007$), however *Military-Other* ($U=556$, $p=0.05$) and *Academics-Other* ($U=220$, $p=0.49$) do not.

Given the nature of the cyber domain, such as virtual interaction having physical effects, the response from the academic community is unsurprising. At the same time, however, it is also members of academia that apply physical concepts to cyber ones, for example when trying to establish what constitutes a weapon (see 2.2). This could also be seen as an argument that specific context matters, with 'software' being too broad a terminological category, which might also account for the indecisiveness of the third group.

When suggesting a three-tiered classification of objects and software/malware (tool/dual-use/weapon), there is agreement. No statistically significant differences were found based on background, age or expertise in CS or IR. This is relevant when considering to what extent current handling of weapons, nationally and internationally, is put in place; this includes prosecution of ownership or use, or international agreements or embargoes. However, this sidesteps the conundrum of

identifying 'good' from 'good and bad', 'bad' or 'really bad' code, or what those terms mean specifically, which also would argue for using the term SaaW.

3.3.2 What Constitutes a Weapon?

Prior to seeking answers specifically related to software, several questions canvassed attitudes towards and beliefs held towards weapons to provide a basis for the discussion. Asked what a weapon was, a large proportion centred on harm and damage, with a small number of descriptions including 'amplifying violence', used to 'kill' or 'injure' and a single mention of 'defend' on its own. Overall, the perspectives are offensively driven, which is also reflected globally within the national and international legal system, with only few jurisdictions, such as Queensland in Australia, defining purely defensive measures, such as bullet proof vests as a weapon. To clarify further, respondents were asked to list three different types of weapons each resulting in a great variety, from specific examples of weapon types to general attributes, with a total of 98 different responses including synonyms.
There are four main axes of responses when examining the attributes:

- **harm scale:** intent, lethality, and target number
- **range:** close combat, short- or long-range
- **means:** physical, psychological, digital
- **construction:** bladed, projectile, explosive, bio-chemical, nuclear

Furthermore, responses highlighted the different terminology used by people: for example, computer/cyber-related weapons were referred to as 'computer', 'computer program', 'computer virus', 'computer hacking', 'malware', 'cyber', 'cyber weapon', 'IT' and 'software' (total of 18); combining bio-chemical weapons (specific types and general description) into one category yielded 10 responses; various uses of one's body (fist, foot etc.) accounted for 9 responses. These could further be merged to create larger groupings. Based on the above responses, it could be argued that weapons are part of *Familienähnlichkeit* – family resemblance – a philosophical idea first proposed by Wittgenstein encompassing things that have one essential feature in common but are in fact linked by several similarities without one feature common to all [340].

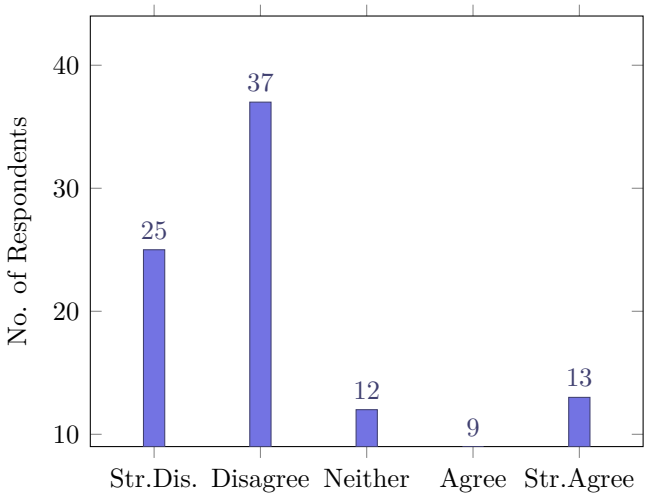


Figure 3.2: Physical damage is necessary for software/malware to be a weapon

A majority of respondents believe that there is a difference between an object being a weapon and being used as one Figure 3.2, and in a follow-up to explain their answer 78 responded, with two groups forming centring on either 'use' defining a weapon (43) or the initial 'design' intent (35). No statistical difference was found based on background, age or expertise in CS or IR.

Yet in the cyber domain, it can be argued that these two aspects are much harder to differentiate given that large amounts of software are inherently dual-use. This would also strengthen the argument for using the term SaaW, combined with the answers of when an object becomes a weapon, with most respondents giving the intent in design and/or use as answers. Keyword frequency varied ('design' (10); 'damage' (10); 'purpose' (10); 'cause' (25); 'harm' (37); 'intent' (42); 'use' (82)), with a greater focus on intentionality, use and harm.

Aside from encompassing software/malware used offensively as a weapon, the use/design intent 'blanket' covers not only offensively used code and attacks, but also exploited vulnerabilities. From a UK legislative perspective, this is in line with the Computer Misuse Act (CMA): Section 1 covers vulnerability exploitation by focussing on obtaining data access regardless of method if not authorised by the data owner; however, knowing that the act is unauthorised is also vital.

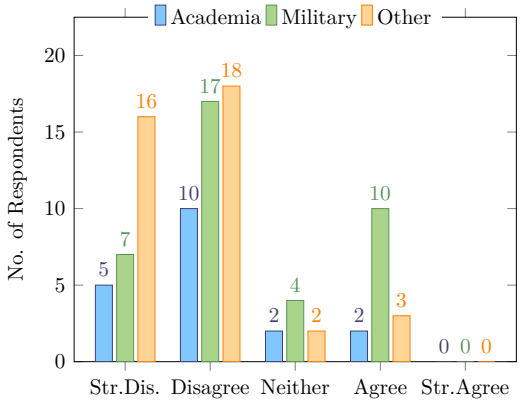


Figure 3.3: Physical damage is necessary for software/malware to be a weapon

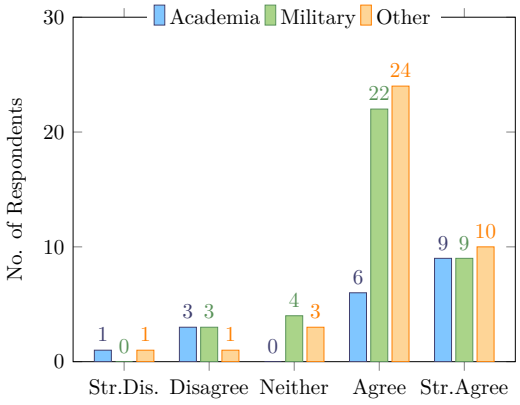


Figure 3.4: Causing physical damage is a weapon regardless of type or severity

3.3.3 Software, Malware and Weapons

Opinions on whether malware is a weapon were split, but not statistically significant based on groupings of age, background, or CS/IR knowledge. The majority agree or strongly agree that malware can be used as a weapon, and that it can sometimes be, or be used as, a weapon. However, there was greater variation in responses to the latter question based on background. More importantly, whether malware needs to cross a threshold to be considered a weapon was inconclusive, with the same number agreeing as disagreeing, indicating uncertainty.

Whilst malware can cause physical damage, with Stuxnet being a prime example, only 83 believe it that be the case, with 7 disagreeing. Of those in agreement, 11 mention 'Stuxnet' or 'nuclear' in their comments referencing the widely publicised attack in 2010. The majority disagree that software/malware is only a weapon if it causes physical damage (see Figure 3.3) – counter to some academic discourse, which would exclude damage done to data CIA.

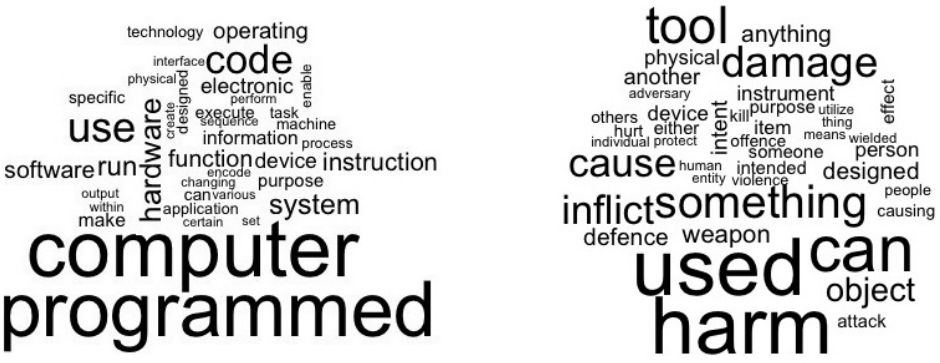
However, there is a difference in responses based on background KWH ($H(2)=7.57$, $p=0.02$), particularly between the *Military* and *Other* group MWU ($U=467.0$, $p=0.004$), with the *Military* being more cautious and tending to require more than a breach of CIA properties for software/ malware to be a weapon. At the same time, there is agreement that software/ malware causing damage to a living being,

a structure or system (a 'cyber-physical attack') is a weapon, regardless of the type or severity of damage – even from the *Military* group, Figure 3.4.

Most disagree that the damage caused determines whether the software/malware is a weapon or not, regardless of the intentions of the attacker; however, there is some difference in opinion, particularly with those of *Academics* background. At the same time, there is no consensus on whether the reverse is true: does the intention of the attacker determine whether the software/malware is a weapon or not, regardless of the damage caused? However, the *Academics* group differs from the other two groups (KWH $H(2)=8.58$, $p=0.01$) by agreeing or strongly agreeing more often (MWU *Academics-Military* ($U=235$, $p=0.003$)), MWU *Academics-Other* ($U=237$, $p=0.005$). This deviates from common and legal perspective on crime, where a conscious decision to deprive or injure is an essential component of *mens rea* – in turn, a basis for establishing guilt.

There is no consensus of whether it makes sense to approach software/malware attacks with (traditional) weapon terminology (e.g. 'warhead', 'trigger', 'payload') and/or measures (e.g. 'calibre', 'yield', 'range'). The responses to the terminology question are surprising to some extent, given that some language of that type is already employed within the literature (e.g. [141]), the IT field and public. Asked for additional comments, (32) responded arguing for and against keeping the terminology; the top reasons can be summarised as keeping those terms if it helps uniformity, awareness or understanding, as language is often 'reused' (12). Those in favour of new terminology (13) suggested new terms centring on effect and output, stating the need for more flexibility and new words for new things. The remaining responses were unsure (5) or using a mixed approach (2).

Three survey questions asked the respondents to define weapons (Q8), software (Q17) and malware (Q19) via free-text. In order to investigate commonalities of words text-mining was used, implementing *tdm* libraries and *wordcloud* in R. The answers were combined for each question into a string, which was then subjected to quality control, such as the removal of uninformative words 'the' or 'it', punctuation marks, standardisation to lower-case and the singular. Derivative words, e.g. 'destruction'



(a) (b)

Figure 3.5: Terms used to define: (a) software and (b) weapons

Figure 3.6: Terms used to define malware (Q19)

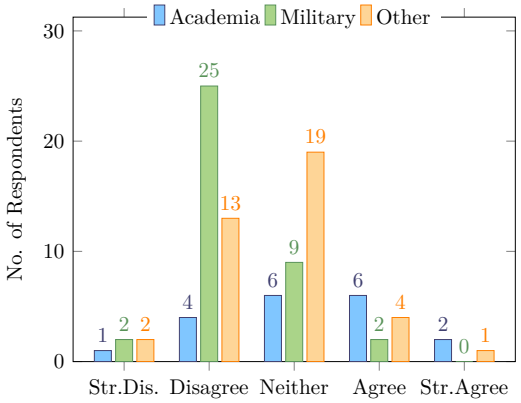


Figure 3.7: Software/malware capabilities have rendered state-centric security models obsolete

and 'destructive', were re-coded into the word with the highest frequency. The frequency of all words was then calculated for each string to create the word clouds seen in the figure below. The size of each word is determined by using the frequency as weight, after excluding any below three.

As seen in Figures 3.5a and 3.6, there is substantial overlap between words used to describe software and malware as they both include 'computer' and 'programmes', with 'code' and 'hardware' that have a 'specific' 'function' or 'purpose'. The words suggest that software can have a broad function, including for example 'various' and 'changing'; malware on the other hand is characterised by unique words, such as 'harm', 'attack', 'malicious', 'extract' or 'steal', 'information', 'negative', and 'damage'. The same overlap is observed between weapon and malware, with both

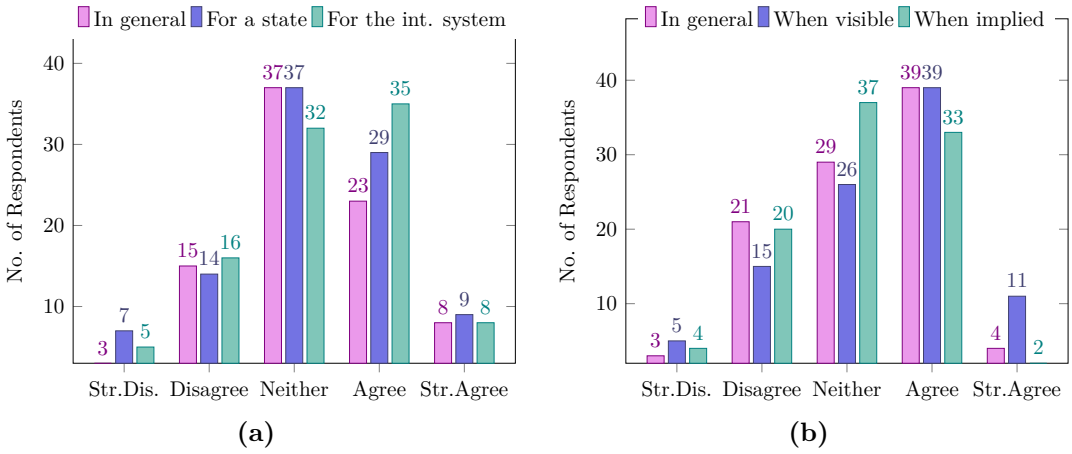


Figure 3.8: Software/malware capabilities lead to more insecurity than security (a), or provide a deterrent (b)

being described by words such as 'damage', 'intent', 'harm', 'attack', and 'person'. Unique words used to define a weapon are 'hurt', 'physical', 'tool' and 'violence'. Overall, the text-mining analysis shows that a weapon and software are defined as predominantly separate constructs, whilst malware is defined as a construct bridging what we understand to be a weapon and software.

3.3.4 (In)Security

No differences were observed between the three groups in the questions pertaining to whether or not software/malware capabilities lead to more insecurity than security in general (Q37), for a state (Q37a) or the international system (Q37b). Figure 3.8a therefore presents the combined responses of the groups to each of the questions. As can be seen, the majority of respondents were split between the neutral option and agreement. A reason for the former could be that the respondents believe the opportunities and threats essentially cancel each other out, or that there is sufficient robustness to accommodate for the changes brought by the new capabilities. The latter on the other hand might be acknowledging the new threat landscape, which disrupts the existing paradigm until some form of normalisation or adaptation occurs, thus leading to less overall security, at least temporarily.

3.3.5 Deterrence

Deterrence behaviour was long practised prior to it becoming a word in its own right, a field of study or a practice of international relations. It has come to be understood, in its most basic form, as preventing someone from doing something, though of course many nuanced forms exist. Within the context of states and international relations, deterrence operates on three different levels (a tactic, a strategy and as a vital element of the global (inter)state system) but a lay person is most likely to be familiar with ideas of military power or nuclear weapons.
When asked whether or not software/malware capabilities provide a deterrent in general (Q38), when displayed/visible (Q38a) or when implied (Q38b), no difference was found between the three groups. The results with combined groups can be seen in Figure 3.8b. A total of 14 expanded on their answers, which covered a wide spectrum and can be found in [281].
Whether it is possible to showcase software/malware capabilities without losing a technological advantage split opinion, but not significantly between the three groups, with the largest proportion being in agreement, followed by those remaining neutral as seen in Figure 3.10(a). The option to comment further was taken by 18 respondents, five of whom stated that they do not feel sufficiently qualified to answer; three suggested that showcasing would be an incentive to replicate the said capability or that it should be possible to show an effect without detail/the code itself; the remainder are single responses, for example stating that it is assumed one has an advantage, that malware requires vulnerabilities/mistakes from others, or that some of the best code development comes from open-source systems.

3.3.6 State-Centricity & Capabilities

There was a high level of agreement that a vast number of actors have been propelled into the security field, both nationally and internationally (Q41), with no difference between the three groups. Yet, the extent to which these new actors have affected the traditional, state-centric model, is highly disputed. The largest divergence in responses overall pertained to the statement that software/malware capabilities

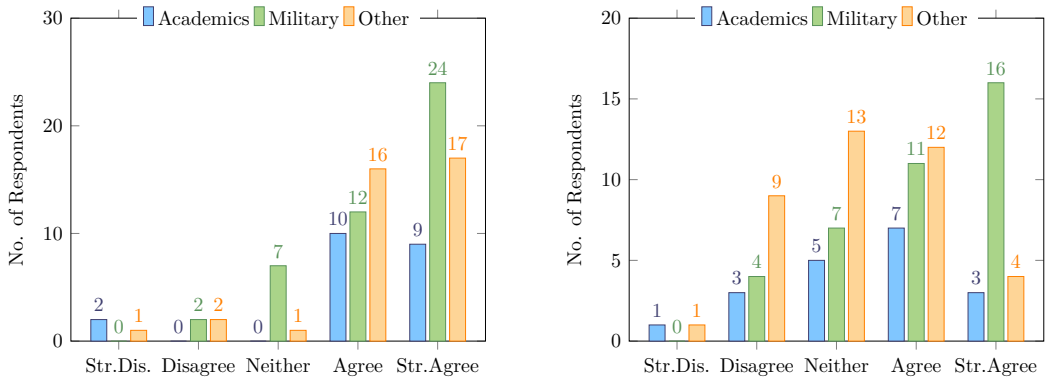


Figure 3.9: Capabilities vital to a state's security: defensive (l) and offensive (r)

have rendered state-centric security models obsolete (Q42) depicted in Figure 3.7, with the KWH test showing significance ($H(2)=13.979$, $p=0.0009$).

More specifically, whilst the vast majority of the *Military* disagree, there is far less consensus within the remaining two groups: *Other* appears to favour the neutral response, but still has many disagreeing; *Academics* on the other hand has more in agreement than not, although also a large proportion of neutral responses. Statistically, significance was found *post hoc* using MWU, with *Military-Academics* ($U=200.5$, $p=0.0003$), *Military-Other* ($U=482.5$, $p=0.0052$), and *Academics-Other* ($U=280.5$, $p=0.0326$).

Additional comments (Q42a) were included by 16 respondents, six of whom can be summarised as believing that it has affected the model but not sufficiently yet to render it obsolete; five see states as still being the dominant element, and another four did not feel qualified enough to answer well. The last respondent asserted the model is constantly changing.

There is clear agreement across the groups that defensive software capabilities are vital to a state's security (Q35) as shown in Figure 3.9 on the left. However, this is not the case when asking about the offensive element (Q36): the *Military* group most strongly agrees that it is vital, *Academics* less so, and the *Other* group the least, as seen on the right hand side of Figure 3.9. A KWH test underlined this finding showing a difference ($H(2)=11.167$, $p=0.0038$) between the groups; using MWU for *post hoc* analysis, *Military-Other* ($U=409.5$, $p=0.0007$)

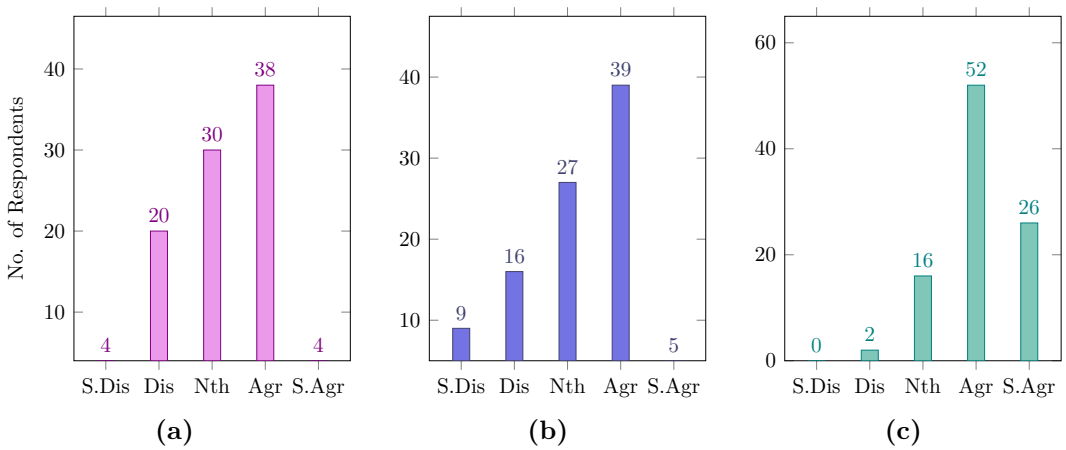


Figure 3.10: Software/malware capabilities: can be showcased without losing technical advantage (a), need to be regulated globally (b), and have propelled a vast array of new actors into the field (c)

and *Military-Academics* (U=273, p=0.0192) variation remain significant, however *Academics-Other* (U=339, p=0.2048) do not.

3.4 Factor Model

The PCA revealed a 15-factor model with *eigenvalues* > 1, together explaining 74.17% percent of the variance. The KMO test was close to the 0.5 adequacy threshold (KMO=0.497, p<0.001). Initially, most factors included a very small number of items, with sub-parts of questions adding little information and thus being removed for subsequent iterations. One factor included Q10 – '*An everyday object or tool (e.g. a cup, a chair, a shirt) can be a weapon*' – and Q12 – '*There is no difference between an object being a weapon and an object being used as a weapon*'. To reduce the number of estimated components and because these questions do not directly assess attitudes towards software and/or weapons, they were also removed from subsequent analyses.

As items were removed due to their relatively small explained variance, the factor solution became more constrained and many items assessing overlapping constructs converged. Ultimately, an eight-component model emerged based on the analyses of 33 items explaining 54.30% of the variation, as seen in Tables 3.1 and 3.2. It assesses the following continuous dimensions:

Table 3.1: Component interpretation and analysis

#	Component Interpretation	% var	Mean (SD)	F	*p*
1	Defining software/malware thresholds and capacity as a weapon	13.37	29.77 (6.43)	0.193	0.825
2	Perspectives on software/malware regulation, terminology, nature and effects	8.55	42.64 (5.71)	2.693	0.075
3	Weapon potential and role of software/malware	7.22	46.2 (4.64)	0.846	0.434
4	Nature, intent and potential of software/malware as a weapon	5.91	39.54 (4.94)	4.847	0.01
5	Topic familiarity and effects of software/malware capabilities on (inter)national security	5.44	23.14 (3.75)	6.654	0.003
6	Weapons use and software/malware capabilities pertaining to state security	4.88	43.25 (4.97)	1.521	0.229
7	Physicality of software and its effects on security	4.61	40.48 (5.12)	2.556	0.09
8	High level perspectives and attitudes	4.33	23.86 (3.18)	2.266	0.113

The table shows the proposed component interpretation, the variance explained by each according to the factor analyses, and the results between groups *(Academic, Military, Other)* testing with one-way ANOVA; Mean shows the group mean and standard deviation; F shows the ratio of between and within subject variation; *p* indicates whether F is significant.

Mean group differences (*Academics*, *Military*, *Other*) were tested in the eight-factor model with ANalyses Of VAriance (ANOVA). Two of the eight scales (components four and five) contained significant mean differences between the three groups. *Post-hoc* comparisons revealed that in component four (concerning the nature, intent and potential of software/malware as a weapon) these effects derived from *Academics* (mean=42.71, SD=4.77) reporting higher values compared to *Military* (mean=38.55, SD=4.16) and *Other* (mean=38.91, SD=5.2) with $F_{(2,81)}$=4.847 and p=0.010. In component five, (assessing topic familiarity and effects of software/malware capabilities on (inter)national security), the significant group differences ($F_{(2,41)}$=6.654, p=0.003) existed because *Military* respondents provided higher values (mean=25.05, SD=2.72), compared to the *Other* group (mean=22.44, SD=3.72) and *Academics* (mean=20.33, SD=3.81).

Although the eight-factor model explained over 50% of the variation and resulted in interesting/understandable/logical components, the KMO test indicated that the options with complex dimensionality reduction techniques are limited in this sample. The relatively low respondent/item ratio (96/33) could partially explain that estimate, as with a larger sample, correlations are estimated with higher accuracy. Another step in instrument development could be to select the most informative items, thereby reducing the number items and the questionnaire length, potentially increasing its usefulness without losing information.

Table 3.2: Proposed eight-factor solution item loadings, after removing Q10, Q12, and suppressing factor loadings below 0.2 for clarity

Question		Component							
		1	2	3	4	5	6	7	8
27	Malware needs to cross a threshold to be considered a weapon.	-0.731		0.232					
29	Software/malware causing physical damage to a living being, a structure or system (a 'cyber-physical attack') is a weapon; software/malware causing damage to data integrity, accessibility and confidentiality is not a weapon.	-0.602		-0.335					
30	Software/malware causing damage to a living being, a structure or system (a 'cyber-physical attack') is a weapon, regardless of the type or severity of damage.	0.55			0.205				
31	The damage caused determines whether the software/malware is a weapon or not, regardless of the intentions of the attacker.	-0.546	0.205		-0.249			0.237	0.217
24	Malware is a weapon.	0.515	0.26			-0.273		0.233	
44	Software/malware capabilities can be clearly differentiated between civil and military applications	-0.394	0.235	-0.245		-0.346		0.312	
40	Software capabilities should be regulated globally, similar to the production and/or proliferation of other weapons.		0.679						
28	It makes sense to approach software/malware attacks with (traditional) weapon terminology (e.g. 'warhead', 'trigger', 'payload'):		0.636	0.202					
38	Software/malware capabilities provide a deterrent in general.		0.629				0.428		
36	Offensive software capabilities are vital to a state's security.		0.485			0.446	0.291		
39	It is possible to showcase software/malware capabilities without losing a technological advantage.	-0.403	0.424		0.305				
25	Malware can be used as a weapon.			0.722					
20	All malware is equally dangerous.	0.236		-0.684			0.2		
22	Software and/or malware can cause physical damage.			0.628					
26	Malware can sometimes be/be used as a weapon.	-0.442		0.585					-0.385
32	The intention of the attacker determines whether the software/ malware is a weapon or not, regardless of the damage caused.				0.83				
23	All software has the potential to become malware.		0.394		0.54		-0.29		
15	There are three types of objects: 1) those created for the sole purpose of being used as a weapon (offensively or defensively) 2) those created for dual use, either as a tool or a weapon, depending on the situation 3) those created to be used as a tool				0.376	-0.354	0.295	0.207	
7	Please rate your familiarity with the topic of international security:					0.748			
42	Software/malware capabilities have rendered state-centric security models obsolete.				0.375	-0.479		-0.443	
14	Weapons are only used offensively.						-0.722		
45	Software/malware exemplify asymmetric warfare.		-0.203				0.637		
35	Defensive software capabilities are vital to a state's security.		0.248	0.201		0.404	0.423	0.218	
43	Software/malware is a vital component of modern warfare.			0.292			0.348		0.231
18	Software should be treated like any physical object or tool.	0.342					-0.207	0.601	
33	Software/malware can be separated into three types: 1) created for the sole purpose of being used as a weapon (offensively or defensively) 2) created for dual use, either as a tool or a weapon, depending on the situation 3) created to be used as a tool				0.4			0.57	
6	Please rate your familiarity with the topic of computer security:			0.253	0.257			-0.539	
37	Software/malware capabilities lead to more insecurity than security in general.							0.482	0.43
34	Most attacks by software/malware fall only into three categories: 1) Mere nuisance 2) Summary offences / misdemeanour / petty crime 3) Indictable offence / felony							0.464	
16	If an object with an un-precedented destructive potential is created, but nobody is aware of it, it is still a weapon.		0.324						-0.546
41	Software/malware capabilities have propelled a vast array of new actors into the security field, nationally and internationally.			0.284	0.303		0.245		0.521
11	An everyday object or tool (e.g. a cup, a chair, a shirt) can be used as a weapon.	0.226			0.303		0.348		-0.499
46	An attack by software/malware can be considered an act of war.	0.229	0.294						0.295

Extraction Method: Principal Component Analysis; Rotation Method: Oblimin with Kaiser Normalization; a. Rotation converged in 90 iterations.

3.4.1 Future Iterations

In as young and rapidly evolving field as cyber security, fundamental studies aimed at describing constructs and understanding the perception of important concepts in different populations are paramount. This study presented significant differences in the interpretation and perception of three basic constructs depending on the background and training of respondents, underlining the population heterogeneity in the understanding of the potential of software and malware of cyber weapons. Understanding this heterogeneity is important to prevent misunderstandings, reach agreement on definitions, and comprehend the potency of adversaries in cyberspace. The aim is that the standardised and objective nature of the assessment procedure proposed in this study contributes to more efficient and valid study of basic cyber security concepts.

On the one hand the groups and responses are very interesting, and the data is therefore valuable. Whilst the descriptive aspect of the chapter therefore provides a contribution to the field, methodological aspects could be improved to present better way of quantifying and assessing the constructs. In retrospect, the questionnaire should have had a more extensive pilot study, for example by utilising a large student base. Thereafter, a PCA would be performed, in turn narrowing the questionnaire down to roughly 15 questions. Resources permitting, there could even be two or more waves prior to final questionnaire creation. The resulting questions could for example have three definition-based ones, three each on software, malware, and weapons, followed by two on objects or overlapping ones. The selection would be based on the informative questions with the highest factor loading in their own component. Once created, this would become far more attractive: for participants due to the reduced time constraint, and analytically due to its empiric construction and basis for statistical evaluation. It could then be administered in waves, giving further data for reliability analysis, for example in a process as in Figure 3.11. It could also possibly include elements from the ‘Danger Assessment Instrument’ [50] or similar sources.

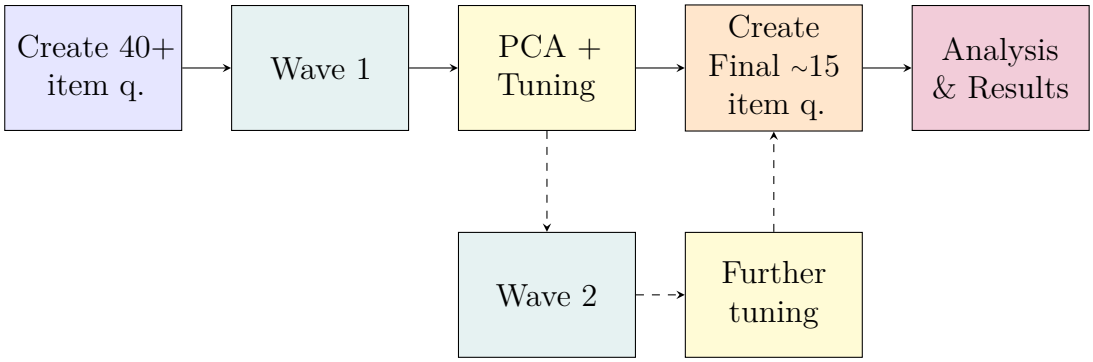


Figure 3.11: Improved questionnaire construction process

Aside from narrowing the questionnaire down as discussed above, the sources and effects of bias may also lead to an interesting contribution, for example in the form of an in-depth psychometric evaluation. Furthermore, it will be interesting to conduct this questionnaire, or a streamlined version, over time to gain longitudinal data, and to observe the changing nature and attitudes.

3.5 Conclusion

This chapter presented the results of an online questionnaire carried out to source public attitudes towards SaaW, supplementing the academic discourse. A total of 46 questions were asked, including consent and demographics, divided into three sections: weapons, their nature and constitution; software, malware, and understandings thereof; as well as SaaW in the context of international security, including capabilities and proliferation. Results included text-mined comparisons between software, malware, and weapons, as well as a structural analysis of the questionnaire itself via PCA.

The responses totalled to N_{tot}=96, split into three groups *Military* N_{mil}=38, *Academics* N_{aca}=19 and *Other* N_{oth}=39. The common five-point Likert rating scale including a 'neutral' response was used for the majority of questions. As the responses were not interval variables and the sample sizes differed, the non-parametric KWH tests were performed to test median differences for more than two independent sample groups; MWU tests were used to test variation in value ranking between two independent samples. The PCA resulted in an eight-factor

model explaining over 50% of the variation, with interesting and understandable results. However, the low KMO test indicated that the options with complex dimensionality reduction techniques are limited in this sample, possibly due to the relatively low respondent/item ratio (96/33).

3.5.1 SaaW

Unsurprisingly, views on the meaning of 'weapon' varied greatly – yet there was also some consensus: around two-thirds consider it to be an object that is designed or can be used to cause harm or damage and is offensively driven. Despite this result matching initial thoughts on the topic, the harm or damage could be caused through defensive measure, particularly in light of active countermeasures. Whilst passive elements, such as firewalls, intrusion detection systems and logging mechanisms, are still essential, more and more defensive measures include active components, such as 'benign' versions of viruses and related technologies that essentially seek out and destroy attacking malware [88].

When deciding whether malware is a weapon, opinions were split, as they were when asked if a threshold is needed for this to occur. *Academics* are more likely to disagree that software should be treated like any physical object, *Military* respondents are more likely to agree and *Others* are indecisive – with the differences being statistically significant. Given the nature of the cyber domain, the response from the academic community is unsurprising, yet the same group often applies physical concepts to cyber ones, for example when trying to establish what constitutes a weapon (see 2.2).

There is substantial overlap between analogies used to describe 'software' and 'malware' by the respondents. Whilst 'software' appears to be thought of having a broad function, 'malware' is characterised by unique expressions. The same overlap is observed between 'weapon' and 'malware', with some unique words used to define a 'weapon'. Overall, it appears that a 'weapon' and 'software' are thought of as predominantly separate constructs, whilst malware is defined as a bridging construct, connecting what is believed to be a 'weapon' and 'software'. It appears as if malware is a stepping stone to software becoming a weapon, yet something 'more'

is required. This 'more', however, remains elusive and will require more research: a clear majority believe malware can cause physical damage but there is disagreement as to what type of attack causes it to be a weapon, whether physical damage (to a living being, structure or system) is needed, or whether an attack on CIA is sufficient, with difference between the *Military* and *Others*. This diverging opinion reflects the debate surrounding the nature of harm in the literature [4], [295], but it also suggest that opinions are shared across specific populations, and future work in this area could explore this further, as well as endeavour to include mixed populations. Whether the intent of the attacker or the damage (effect) caused is what makes software/malware a weapon is inconclusive: on the one hand, most disagree that damage matters regardless of intent; on the other hand, there is no consensus that intent is the decisive factor. This is particularly interesting given that intent is highly important in other situations, whether legal matters or inter-state relations, and has been used to define cyber weapons. For example, based on this, Rid & McBurney's definition of a cyber weapon "as computer code that is used, or designed to be used, with the aim of threatening or causing physical, functional, or mental harm to structures, systems, or living beings" [249, p.7] would need to be modified, adding further leeway to the element of intent.

3.5.2 (In)Security

Whether or not software/malware capabilities lead to more insecurity than security in general, for a state, or the international system resulted in an overall split with most respondents remaining neutral or agreeing (regardless of groups). The neutral responses could be seen as an argument that there is no advantage gained by this technology; in essence, opportunities are being balanced by threats. Alternatively, the view could stem from the belief that there is sufficient robustness to accommodate these new capabilities and/or that the security dilemma does not apply – an interpretation running contrary to the perceived asymmetry that is heavily weighted to the offensive [44]. The second view, however, could be used to argue for two opposing sides: on the one hand in support of expanding cyber commands (or

equivalent institutions) to combat further insecurity, or, on the other hand, to not do so because it acts as a disrupting force fuelling the security dilemma.

In other domains, such as the realm of nuclear conflict, the idea of deterrence is often drawn upon in the context of proliferation. It has also been discussed in the context of the cyber domain [9], [158], [192], [223], particularly given that capabilities cannot be openly showcased. Yet, only about one fifth disagreed that software/malware capabilities provide a deterrent, whether in general, when displayed/visible, or when implied. Furthermore, 38 respondents believe that software/malware capabilities can be showcased without losing a technological advantage, whilst 20 disagree. The results favour the view that deterrence or a concept thereof applies to this domain, regardless of contrary academic discourse [106]. There was a high level of agreement across the groups that a large number of actors have been propelled into the security field, both nationally and internationally, which corresponds to the literature. To what extent these new actors have affected the traditional, state-centric model, however, is highly disputed, and it will be an interesting question to revisit in the future. The responses supported the choice of including state and non-state actors, with an emphasis placed on the former, for examining what factors motivate SaaW pursuit.

Given these various questions of exploring SaaW and the (in)security it brings, what motivates an actor to pursue these capabilities? Whilst weaponisation of software might be a new element of 21st century warfare, the pursuit and proliferation of weapons is not, nor is the study thereof. The results of this questionnaire provide the backdrop for Chapter 4, which moves on to explores the factors that contribute to the pursuit of SaaW.

4
Conceptual Model

Contents

The previous chapter explored the understanding of what it means for software to be a weapon, which this chapter combines with weapon and proliferation theories to explore what factors drive an actor into pursuing SaaW. The purpose of the analysis is to create a group of determinants that act as variables, which are then interconnected to create a conceptual model. This forms the basis for an implementation using BNs discussed in Chapter 5, followed by case studies being applied in Chapter 6. Whilst these elements are naturally simplified in order to

create a model, they nonetheless introduce objectivity to the discourse on actor motivation in the cyber domain and provide a variety of stakeholders with a unified starting point.

First, a summary is given on how the various theories and bodies of previous work are used to create a set of determinants, or variables. This includes a brief introduction to the history of proliferation and the importance of power, followed by a discussion on overarching weapon categories in this context: Nuclear, Biological and Chemical (NBC) weapons and SALW. The later part of this chapter takes the theories discussed and breaks the various elements into three broad categories of analysis pertaining to weapon pursuit: *restraints*, *motivations*, and *capabilities*. Each is discussed in turn, examining interrelationships and sub-elements, both competing and complementing, creating a conceptual model.

4.1 Proliferation to Diffusion

Whilst weaponising software might be a new element of 21st century warfare, the pursuit and proliferation of weapons is not, nor is the study thereof. Weapons have gone hand-in-hand with human development, utilised as tools and for security with the earliest images depicted in petroglyphs. This rich history helps to explain what factors contribute to the pursuit of weapons, which can then in turn be applied to the cyber domain and SaaW.

Within academia, security appears in various fields and guises, from politics and IR to cryptography, with aspects of each of these examining weapons, whether conventional, non-conventional or digital. Within IR, security theory was dominated by realist thought of power as a key element throughout the Cold War, which affected perceptions of nuclear proliferation. The temporary apathy after the dissolution of the Soviet Union seemed to ring in the *End of History* [112], with the flames of idealism and liberalist notions being rekindled. Some hailed this as a final nail in Clausewitz's coffin, rendering his notion of rationality and war obsolete [79], [161], [164] and bringing realist assumptions even more into question. Yet shortly afterwards, civil wars and conflicts erupted across the globe, ending that period.

At the same time as the international system experienced an abrupt transformation in the 1990s, another but different revolution began to take shape and gather momentum, namely that of information, with parallels recently being drawn to the 'nuclear age' [226]. New technological developments, the spread of the Internet and globalisation connected states and people like never before in history, with its pervasive nature ever increasing. Similarly to the political realm, the technological realm followed suit and sought to extend those liberalist thoughts, with strong sentiments towards and belief in freedom and rights, most pointedly identified by Barlow in his *Declaration of the Independence of Cyberspace* in 1996. However, the technological advancement also brought about new threats and vulnerabilities, with an increasing number and variety of actors, ideologies, abilities and resources, changing the very nature of the challenges faced by states and societies, changing even Barlow's views [93]. Whilst malicious software and attacks are nothing new from a technical perspective, the ubiquity of interconnected digital systems, and the dependency thereon have vastly increased the potential attack surface. These attacks range from generic malware that can be compared to petty theft or vandalism, to sophisticated, multi-stage persistent attacks requiring resource intensive reconnaissance, development and testing touted as a next generation weapon, with Stuxnet often cited as an example.

The problem surrounding cyber weapons or SaaW is often portrayed as something new and different, with elements being "easily available to any and all actors, nefarious or legitimate, the spread of these weapons appears particularly complex but banal, and not so much uncontrolled as uncontrollable"[37, p.14] – yet that phrase was used to describe SALW, not software. But are the factors motivating development and spread the same? Can a comparison be drawn to traditional proliferation theories pertaining to NBC or the more recent SALW discourse given the vastly different nature of weapons?

Originally applied to the nuclear domain, and separated into horizontal and vertical aspects, proliferation has over time become synonymous with the spread of weapons or technology deemed dangerous to (international) security. It assumes a few-to-few

relationship, with the actors traditionally being states. It is inextricably intertwined with concepts of security, power, and war. The literature spans a myriad of different interpretations and sub-schools, with one of the most used definitions of 'security' (within the political context) still remaining Wolfers' from the 1960s – an absence of threat or fear to acquired values [341]. The more formalised concept of security developed and traditionally focussed on the three-dimensional area of our world pertinent to a peoples' survival, or later, a countries' defence, covering aspects of land and sea protection before technological advances introduced the "air domain". Over the course of time, international bodies were created to oversee treaties and laws, governing conduct between, and to some extent, in states, though their reach may often be limited for a variety of reasons, such as ratification or support forming what was to become our current international system. Regarding nuclear weapons, a change in perception occurred in the 1960s, when security beliefs turned to those of insecurity, allowing for the creation of the NPT. The space race led to a codification of space security in the form of the *Outer Space Treaty*, which in essence excludes that domain, ensures neutrality and prevents signatories from military activities, although not the placement of conventional weapons.

Since the end of the Cold War, the literature has shifted to what motivates actors to pursue nuclear weapons, in line with more practical definitions of security [46]. Most proliferation literature has been devoted to Weapons of Mass Destruction (WMD) – NBC weapons – often grouped together for convenience to distinguish them from conventional weapons technologies. However, a sizeable group seems to imply that the pursuit of one may be related to that of another, partially fuelled by US Presidential speeches by Bush and Obama [217] or by books encompassing all three categories [339]. Whilst the cyber domain is in many ways different to the others, the decades of research on the motivations behind the spread of other weaponry cannot be ignored. Similarly, nor can the concept of diffusion, which involves the spread of technology (or more recently weapons) in the form of many-to-many, often including organisations and groups at various levels in addition to state actors.

4.2 Power

Power is a concept that has been studied since Aristotle, with one commonly used definition being based on Max Weber's work, seeing it as the ability to extend one's will – although many variations and nuanced interpretation exist. In the context of global politics, power is projected within a specific scope and relates to specific topics, for example, military actions, economic sanctions, or the ceasing of diplomatic relations. Whilst power is not the focus of this thesis and a detailed discussion is out of scope, concepts need to be introduced as they provide a background to discussing not only weapons but also more subtle forms of influence and coercion, which are vital to the debate regarding SaaW and provide connections between variables.
Since the Peace of Westphalia in 1648, the predominant powers have been sovereign states, and, although power transition has been seen, the recent growth in complexity and diffusion of power are new, a great proportion of which can be ascribed to technological developments and globalisation. It is precisely this increasingly pervasive reliance on cyber that nourishes the use of cyber power at all levels, from strategic to tactical. SaaW could not only tangentially affect dynamics of global power, but also be a core driver and game changer, as the cyber domain is often at the heart of key industrial, as well as social processes. Additionally, Chapter 3 saw that participants from the questionnaire agreed, regardless of background, that a large number of actors have been propelled into the domain, but to what extent they affect the traditional, state-centric model, was highly disputed.
Although concepts of power are ambiguous, varied and defy unequivocal measurements, many have sought to classify and measure power, with some focussing on use of language whilst others seek indisputable quantification. Within political science, power began to be formalised in the middle of the 20th century, originating first as the concept of influencing others to exert behaviour they would normally not. Over time, the focus shifted to measuring power [146], with early approaches[1] seeking to quantify power. These include national income [86] or combining national economy,

[1]It should be noted that measuring power of a state or an empire was also considered in previous centuries, for example by 1741 Johann Peter Süßmilch in *Die Göttliche Ordnung*.

land, population and military power [117], some of which were then developed further [111], [284]. The Composite Index of National Capability [282], [283] is still in use today. This was followed by embracing greater abstraction addressing control over resources, control over actors, and control over events and outcomes [134]. Power can broadly be seen as a push and pull function, relating to 'hard' and 'soft power' respectively, the latter a concept formalised by Nye [224] that has to be briefly introduced. In Nye's terms, hard power is core to realist thought and is based upon a strategy of coercion/punishment and reward to influence behaviour of others. Although this type of power is often reflected by conventional military capabilities, it can also include more resource-focussed elements, such as economic resources, with the US currently being a prime example of a state having the capability to exert both at an almost exclusive level. Soft power, on the other hand, is more subtle and relies on the use of resources to either attract key people who disseminate their thoughts, in turn influencing decisions to the point they become favourable; alternatively, it can be used to attract the greater public, which in turn pressures its own government. This approach has gained strength predominantly due to technology and globalisation, diversifying the number and type of actors, their relationships and means of communication.

4.2.1 Cyber Power

The concept of cyber power has emerged alongside the cyber domain, with varying perspectives. Nye first published on cyber power in 2010 stating that "power based on information resources is not new; cyber power is" [225, p.3], which he then expanded on [227]. He notes that there are many differences between cyber space and the other domains, with the most important ones being: a human created domain; the combination of real and virtual, as well as speed of advancement and change is unprecedented; the cost of entry is low, resulting in a large variety of actors; lastly, the interaction and reliance of private and public resources has also not been seen previously. Moreover, Nye sees cyberspace being a perfect example of a broader trend of power diffusion away from governments [226], which is also

supported by the results of Chapter 3. Yet, his assessment of certain aspects is questionable: for example, he did not perceive China as a threat to United States' supremacy in the near future, nor does he address the role or ability of the United Nations (UN) Security Council in regards to enforcing international treaties. In that work, he does not address the power gained by governments from monitoring their own population, nor the effect this may have on spreading democratisation. This different focus is picked up by Betz [26], encompassing the "power to shape people's perceptions, beliefs and ideals in ways that structure their responses to events in ways that are congenial to one's own ends" [26, pp.9].

Sheldon on the other hand believes that its purpose revolves around "the ability in peace and war to manipulate perceptions of the strategic environment to one's advantage while at the same time degrading the ability of an adversary to comprehend that same environment" [271, p.96], building on Kuehl's work. Whilst some aspects are analogous to Nye [225] (e.g. the low cost of entry or the ability to be stealthy), others share rudimentary similarity but are branched differently (e.g. cyber power's complementary nature), yet others are very distinct, such as cyber space's electromagnetic spectrum reliance. From his perspective, there are two main limitations to understanding cyber power: firstly, due to the greatly varied schools of thought, the perspectives and thus conclusions on the domain itself are greatly different; secondly, the literature gap between a state's capabilities and their actions. Meanwhile Gomez [121] attempted to identify what cyber strategies states may use against each other based on historic cyber conflicts, aiming to provide a basis establishing controls as well as effective deterrence. In his analysis, he groups states by their perceived cyber power, examining their inter-relationships, resulting in a view of the cyber domain that is divided, and at times, even conflicting.

Rowland et al. published two papers on cyber power [252], [253], in which they discuss this new domain's properties and actors. In their first work [252], they seek to clarify what it entails to be a cyber power and what components are necessary to achieve this, addressing both, state and non-state actors. They focus on what cyber space and a corresponding cyber entity is, including themes such as ideology, profit

motivation and infrastructure (physical and cyber), as well as human components of these entities, such as citizens, shareholders, or casual participants. In their second, more expanded paper [253], they explore where cyber power is heading and whether it will trump traditional sea, air, land and space power. They use power theories employed in the physical world to examine cyber power and states wielding it, addressing state as well as non-state actors. After a brief introduction of the history of power they settle on using the diplomatic, informational, military and economic power (DIME) model created during the Carter and Reagan administrations [53], [243]. In contrast to Gomez, they group national powers into five categories based on domination, building on Kokoshin's work. Non-state actors are also covered, in particular multinational organisations, terrorist groups and organised crime.

Whyte [336] published an alternative framework for understanding the sources of national security and cyber power, particularly questioning whether cyber weapons can be used to influence global power dynamics. He asserts that although there has been an emergence of analytic and technical work regarding cyber and national security, the focus has been too greatly on the state, its power, and the cyber domain. Furthermore, he goes beyond Nye's and others' beliefs of power diffusion and seeing cyber space merely as empowering lesser states: he argues that digital advances not only mobilise but also create power in latent and societal terms, with wide-scale regular attacks on non-CNI having a large effect on states, diminishing potential and creating a deficit as a result. Addressing cyber weapons specifically, Whyte differentiates between Cyber Weapon of Mass Destruction (CWMD) and Cyber Weapon of Mass Effect (CWME), to then implement a predator-prey model. It represents international affairs and explores the potential capacity-altering ability of CWME" [336, p.102] from the perspectives of state power and agents. For 'mass destruction' in the context of cyber security, he "refers to the targeting of particular critical systems with sophisticated payloads at opportune moments" [336, p.104]. Furthermore, although he acknowledges the difference between cyber attacks and exploitation, he merely sees them at the functional and not the strategic level. He states that "[...] digital instruments lack a singular function. Unlike nuclear

weapons [...] the shape of digital methods [...] depend very acutely on the technical environment in which they are deployed", which is reflected in the nature of policy allowing for "situation-specific cyberweapon deployments rather than massive ones" [336, p.102-103]. As an example, he references Stuxnet, and less destructive but more information gathering and exfiltration software, including loss of intellectual properly. The importance of CWME for cyber security and their relevance comes from them being designed for the purposes of sabotage with a massive ripple effect affecting a country's 'refresh rate' following the predator-prey model that describes the dynamics of two species interacting in biological systems.

Regarding the utilisation of these weapons, particularly CWME, he hypothesises that it may be predominantly used in private civil and industrial society, followed by governments in classified operations. Whyte also raises a concern that often the debate of cyber capabilities turns on the character and quality of a state, as opposed to the nature of the security environment. This leads to the issue of governance that will remain particularly challenging from his perspective, not only due to the gaps between technological change and speed of policy making but also as it may not be in the interest of a state to govern cyber weapons — any regulatory success is contrary to motivations of pay-off. Furthermore, Whyte sees difficulty in the verification of implementation of frameworks, as well as neutrality, credibility, attribution, and trust: all obstacles that followers of this discussion are familiar with. Yet there is also hope that more data collection or modelling might allow for pattern detection and a quantitative analysis of impact. In summary, Whyte insists on the importance of revisiting relationships and interactions in cyber space to fully comprehend the power dynamics, particularly in relation of cyber weapons. He further warns of low-impact but long-term operations, particularly theft of intellectual property, as that undermines a state's growth potential and resources, in turn limiting leverage as well as credibility.

4.3 Nuclear Weapons

Since 2010, there have been numerous comparisons between nuclear and cyber technologies, ranging from academic work for example on deterrence aspects [61] to statements by various military personnel and ministers around the world [62]. Whilst main criticisms of comparing nuclear and cyber focus on the sheer destructiveness, the assuredness of destruction, as well as a lack of common understanding, lessons can nonetheless be learnt, particularly when taking the problems faced in the early stages of the nuclear age into account, when nuclear learning was "slow, halting, and incomplete" [226, p.36].

In the decades since the introduction of nuclear weapons, a large repository of literature has been built, resting on competing traditions and schools of thought that have sought to provide an explanation and coherent theory of why, and at times how, states proliferate nuclear weapons. The late 1990s saw the publication of Sagan's now seminal article outlining three models, or core motivators: security, prestige and domestic politics, contrasting the then near-consensus that "states will seek to develop nuclear weapons when they face a significant military threat to their security that cannot be met through alternative means" [255, p.54].

However, no single theory can account for all the historic scenarios thus far [321]. Instead, each theory has varying degrees of explanatory power, with different strengths and weaknesses for specific scenarios [58]. Whilst the two domains – nuclear and cyber – are very different, certain comparisons can be drawn based on elements from the various schools of thought discussed below.

4.3.1 External factors: Flavours of Realism

Although a detailed discussion on the various schools of thought is beyond the scope of this thesis, IR's rich tradition of thinking provides valuable context and thus it is briefly summarised. Within the political security theory and IR discourse, the realist tradition has been one of the most prevalent schools of thought, with seminal works by Carr [52] and Morgenthau [213] laying the modern foundations. Classical realism focussed on the brutal reality of the world rather than its ideal

state. Taking the pessimistic view of constant conflict across human history, the main analytical tools became security and power [47], with state actors as the protagonists. In this tradition, Waltz first pursued the origins of war in his famous work '*Man, the State and War*' [327], before contesting aspects in his next work [329] formalising what became to be known as neo-realism.

Also referred to as structural realism, neo-realism aims to explain the relative peaceful period during the Cold War – but not weapon proliferation *per se*. Neo-realism embodies the idea that the structure of the international system itself influences and affects international politics, as well as outcomes, with power being the most important factor. There are also some differences which are more philosophical in nature, such as that realists seeing the flawed human nature as a root cause of conflict, whereas neo-realists ascribe these to the anarchic system. In further work, Waltz also addressed the origins of war [328], with an expansion by Mearsheimer at the beginning of the new millennium adding 'real-world' applicability [206].

However, observed changes and power shifts, such as the rise of non-state actors [324] or technology were not quickly absorbed or accepted into the theoretical framework. Weak predictive power led to a call for re-introduction of various elements, resulting in a far more complex and dynamic version of neo-realism, driven by Buzan, Jones and Little [48] – but without application to nuclear proliferation. Waltz however did, combining it with rational deterrence theory [256] whilst remaining cautious regarding its explanatory power. More nuances are offered by splitting power into economic and political aspects amongst others [87], or the inclusion of domestic issues and organisations that affect various actor types differently [25]. Waltz developed his theory further in a discussion with Sagan [257], allowing for an actor to have organisational or even individual levels.

4.3.2 Internal Factors: Domestic and Organisational Elements

Roughly a decade after Waltz's development of neo-realism, a new approach sought to analyse the issue of nuclear proliferation from the bottom up [245]. Reiss introduced

economic cost, political opposition, international disincentives, environmental risks as well as global public opinion, all of which affected leadership attitudes towards proliferation but cautioned against generalisations.

Another major counterbalance is the liberal tradition, which in its neo-liberal form agrees on the same epistemological approach as that of neo-realism (rational nature, state-centricity) but has a more optimistic perception of human nature. It includes internal elements as core factors [55], [289], seeking to adjust the security concept to reflect the changing nature of world politics, society and power distribution. Moving away from the Hobbesian state of survival, it includes non-state actors, individuals and NGOs, as well as, *inter alia*, economics. This explains why actors with liberal democracies move towards shared norms and values, creating a security community, and it accounts for actors with a nuclear option but without a large arsenal. This approach might be applicable to the cyber domain, once idealism is put aside [102]. As part of this, some argue that the political system has a direct effect on nuclear proliferation [55], [210], [261], with democracies appearing to be less likely to pursue it. Yet others disagree [153], [211], [212], with empirical studies existing for both arguments [148], [149], [159], [286], [332]. Similarly, the belief that public opinion can be a direct influence (and thus stronger in democracies) swings both ways, with arguments for negative [63] and positive [178] correlation based on an actor's political system.

Nonetheless, these theories also do not fully explain the proliferation challenge, with Meyer [208] adding a three-step decision-making process: a latent capacity decision leads to an operational proliferation capability, followed by an operational program. This process can be used to explain externally perceived inconsistency in actions regarding nuclear proliferation, as hurdles and setbacks are faced internally. Sagan [254], [255], a fierce critic of rational deterrence theory and Waltz, focusses on the role of organisational entities in the process of decision making, using many of the same assumptions as outlined above. Believing that nuclear weaponry is a great destabiliser, he tackles the central beliefs of states being rational unitary actors.

4.3.3 Cognitive, Sociological & Interdisciplinary Theories

Psychological approaches attempt to fill the gap left by the other theories when addressing irrational behaviour or decisions by looking at belief systems. Drawbacks however include the view of others being shaped by their own beliefs with projected perceptions causing misinterpretations. Applied to nuclear proliferation and creating the 'myth-maker' model in the process sought to counter this [189]. Another theory centres on cross-national expert groups, sharing common goals beyond national interests, and has been used to address instances of nuclear proliferation (or lack thereof) [3], but it does not explain when or how one group can topple another, nor does it allow for aspects found in learning models which address doubt and belief transitions.

A downside shared by these theories is the difficulty of quantifying psychological elements, as well as their limited scope, inviting sociological approaches. These seek to examine the relationship between elements as varied as technology and beliefs or events. Mac-Kenzie [194] connected these elements in his ground-breaking analysis of nuclear missile guidance, which was applied to nuclear proliferation [107], reintroducing technical, historical and sociological factors without necessitating deterministic approaches in the form of social construction of technology theory. The large number of dependent variables, however, often prevents prediction.

4.4 Biological and Chemical Weapons

The mid 1980s saw a rise in concern regarding the spread of biological and chemical weapons, though their watershed moment dates back to the First World War [290], leading to the CWC in 1997. Unlike nuclear weapons, biological and chemical ones are relatively inexpensive and easy to produce. A large proportion of the required materials for chemical weapons can be used for legitimate (dual-use) purposes and deployed via basic delivery mechanisms – a similarity to SaaW, despite an otherwise very different nature. Certain biological agents, such as Anthrax, can be grown in less than a week and are even cheaper to produce.

If these different weapons are sought solely to further security, then the effort to reduce the proliferation of one type may cause greater interest in another, a theory suggested previously [25]. Some work has been done to document the NBC weapon spread comprehensively, for example including the delivery systems and technical capability requirements [64]. It has also been suggested that protection against insurgents may be a motivator [290, p.96] or to spread terror [344].
The research focussing on the interrelationship of the three types is limited, with one notable exception concluding that the underlying demand was indeed correlated [151]. Furthermore, the weapons appear to function as complements during the pursuit stage, but become substitutes upon acquiring nuclear weapons, thus the coining of the term 'poor man's atomic bomb'. In this sense, SaaW may become an even poorer man's weapon of choice; however, the destructive element and shock-factor is currently not near that of NBC weapons.
Comparisons to the cyber domain have been made, for example by exploring the applicability of the CWC [115] as discussed in Chapter 2. Another approach has been a side-by-side analysis of common characteristics between biological and cyber weapons [173], which in summary are: problems of attribution; multi-use nature of the underlying technologies; attractiveness to weaker powers and non-state actors due to asymmetry; force multiplication; potential for collateral damage; importance of secrecy and surprise; and, lastly, questionable deterrent value.

4.5 SALW

Gaining traction in the late 1990s, the discourse on SALW has grown exponentially since, however often with a stronger policy than academic focus. The early work centred on SALW's uniqueness, or the question thereof, not unlike the recent work on all things cyber. Whilst annual yearbooks on the topic exist, such as the 'Small Arms Survey', a very limited number address an overarching framework: application of diffusion to the topic [172], layer addition to the model of conventional weapon spread allowing for more grey and black market areas [176], [177], and the

creation of a new, three-level framework aiming to better uncover the dynamics and structure of SALW [37].

More specifically, Bourne suggests three images of weapon spread: the availability and acquisition structures of technology, which includes know-how at a global level, as well as the respective infrastructure; the availability and acquisition structures of weapons, split into the trade threshold, supplier base, recipient base and non-state threshold; and lastly, the policy and normative structures affecting acquisition [37, pp.17-34]. He also questions the prevailing concepts of a vast, global SALW stock combined with buzzing illicit trade and shadowy brokers (ibid., p.34). This depiction appears to be also rather on point regarding many aspects of SaaW.

4.6 From Theories to Determinants

As discussed in sections above, there are a great number of theories and explanations on what drives actors to pursue weapon technologies, ranging from external to internal, sociological to technological or economic factors. One possibility would be to delve deeply into a specific school of thought and then build upon that; another approach would be to analyse strengths and weaknesses. However, with the aim of creating a balanced and comprehensive basis for discussion, this chapter draws up a set of variables capturing as many factors as possible, with similar work having been done in the nuclear realm [71], [110], [147].

The numerous factors have been extracted from various theories discussed above and broken into three sub-sets, before having been applied to SaaW, focussing on the source of the impetus: motivation, capability, and restraint – all of which are interlinked. The weighting of these factors depends on the actors in question and their goals: for example, is the actor a large nation state that intends to create a Stuxnet-like weapon, or a small group seeking to run low-to-medium level interference?

While the model presented below allows modularity, future work seeks to increase node granularity and apply additional levels of fine-tuning, as well as create a

wider case study pool. Each node can be explored further and split into sub-elements. It should also be noted that in reality almost all nodes are interconnected, and that these connections are often bi-directional, i.e. A affecting B but also B affecting A. However, in order to create a model that can be implemented as a BN discussed in Chapter 5, they have been made acyclic depicting only the most influential connections and nodes.

4.6.1 Restraints

Restraints are, essentially, reasons for an actor not to develop or proliferate SaaW and they could also form part of the motivation group, albeit in a negative manner. Against a backdrop of increasing and intensifying conflict arising in the cyber domain, it is particularly important to explore and analyse what elements can limit, or even control, further escalation. These restraints have been split into four main elements:

- Potential Collateral
- Domestic Safeguards
- International Agreements
- Fear of Retaliation

These in turn have one or two sub-levels of factors, namely:

- Human Collateral
- Environmental Collateral
- Opponent's Military Power
- Opponent's Alliances
- Public Perception
- Security

4.6.1.1 Potential Collateral

The inclusion of this node serves a dual purpose, firstly to connect to conventional warfare methodology, and secondly, to allow for future developments that will likely lead to even greater interconnectivity, such as the IoT.

Standard methodology exists in conventional warfare to identify and assess collateral damage, as "definitions are clear and the harms described are tangible because they relate to persons and property" [251, p.11]. For example, NATO standard AJP3.9, pertaining to Allied Joint Doctrine For Joint Targeting, has Section VII devoted to collateral damage considerations. For the purpose of Collateral Damage

Estimation (CDE), collateral damage is here defined as "the unintentional or incidental physical damage to non-combatants, non-military objects or environment arising from engagement of a legitimate military target" [219, p.25]. Furthermore, "CDE provides a probability, but not a certainty, of collateral damage for a specific weapon system" with a "methodology [that] recognises levels of collateral damage as estimated by certified analysts" [219, p.26]. This is a process that can be planned and managed for, but becomes problematic in the cyber domain given the highly interconnected, and often nebulous, nature.

SaaW are considered to pose a much smaller risk, particularly risk to human life because of their non-kinetic nature [200]. At the same time, many publicly known attacks have been indiscriminate causing collateral damage, despite examples that they can be highly discriminate, with comparisons having been made to precision-guided munitions [2]. With yet growing interconnectivity and IoT ubiquity on the horizon, including autonomous vehicles and biomedical devices, the potential for collateral damage will likely increase.

This node includes the sub-nodes *Human* and *Environmental Collateral* that could occur as part of SaaW use, such as unintended consequences. Three other sub-nodes have been considered: an economic node to denote financial ripple effects, a node to represent CIA and related damage to IT and digital systems, and one node to represent reputation. Whilst they are not currently part of the model, they have, amongst others, also been suggested by respondents in the context of an instantiation of the model (see Section 6.2).

4.6.1.2 Domestic Safeguards

Safeguards often come in international and domestic forms and are considered to be (voluntary) control measures. Examples of the former in the nuclear domain include unrestricted access, inspections or monitoring systems supporting the claim of a peaceful nuclear program. However, this is not the case for SaaW given their nature and a lack of formal legal rules pertaining specifically to questions of cyber

security or warfare, as discussed in Section 2.2. International safeguards are limited to the *International Agreements* discussed below.

The domestic form encompasses specific steps taken by an actor to conform to obligations. Taking the nuclear domain as an example again, this would include the Office for Nuclear Regulation in the UK, which supports and intervenes with duty-holders, including Euratom and IAEA. Cyber equivalents however are hard to come by. Instead, this node currently relies on *Polity IV Data* and *Public Perception*. Whilst the influence of these two nodes can be disputed, accounting for them gives increased flexibility for future work.

4.6.1.3 International Agreements

As discussed previously, there are no agreements binding actors, although there has been some scholarly work on the topic [116], or the study on how international law might apply to cyber warfare [266]. Nonetheless, aspects of the Wassenaar Arrangement pertaining to malware and its related technology apply [330], as do other 'traditional' agreements.

4.6.1.4 Fear of Retaliation

The pursuit of new weapon technology usually threatens to upset the balance of power, whether at a local or global level. Particularly state actors may fear reprisal but may gain a bargaining chip to (re)negotiate trade deals, aid, or status. It could be argued that retaliation is to a large extent bound to attribution, which presents a problem in the cyber domain, given the problems thereof, as well as lacking international agreements or explicit norms. Nonetheless, offensive cyber ambitions made official could fuel other actors' ambitions or invite cyber attacks. Retaliation via non-cyber means is possible in the near future, with arguable only two known examples thus far including the 2015 US drone strike that killed the then-head of Daesh's hacker groups, Junaid Hussain, as well as the 2019 Israeli Defense Forces launch of an air-strike in retaliation to a cyber-attack by Hamas' cyber operations amid a periodic flare-up in tensions and hostility [98]. For non-state

actors, external threats may intensify if they are known to seek out additional weapons, particularly NBC ones.
One additional aspect to consider is the US Active Cyber Defense Certainty Act introduced in 2017 [126], extending the powers of victims of cyber attacks beyond the limits imposed by the Computer Fraud and Abuse Act (CFAA) of 1986. It allows authorised individuals and companies to legally explore systems outside their own networks in order to:

- Establish attribution (source, nature, cause)
- Disrupt attackers without damaging systems
- Retrieve and destroy any files stolen during the course of the attack
- Monitor attackers' behaviour
- Use beaconing/dye pack technology

However, this is not agreed upon internationally, and such cross-border activities could thus lead to an incident, in turn affecting the probability of retaliation.
The *Fear of Retaliation* has currently been split into *Opponent's Military Power* and *Opponent's Alliances*. Additional nodes could include economic factors or the probability of attribution. It would also be feasible to essentially include a 'copy' of the model but from the opponent's perspective, or to create several versions based on not only the envisaged SaaW but also target and attack type.

4.6.1.5 Sub-Levels

Human Collateral For many actors, human collateral damage is to be avoided where possible and is illegal under many definitions. For example Article 51(5)(b) of the 1977 Additional Protocol I to the Geneva Convention prohibits "an attack which may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which would be excessive in relation to the concrete and direct military advantage anticipated" [83]. Aside from the moral and ethical aspects, it also damages an actor's reputation. There are however some exceptions, such as specific terrorist groups who do not ascribe to international norms and instead thrive on collateral.

Cyber attacks (currently) rarely reach the required level of harm, which from the armed forces perspective is considered to be an operation that causes injury or death to persons, or damage or destruction to objects [267]. This does not include inconvenience, fear or stress. Indirect effects are more likely, for example in the form of inadvertently affecting components used in personal devices, medical institutions, or traffic control systems, as have been seen in attacks such as WannaCry.

Environmental Collateral Damage to the environment has been associated with NBC weapons that have a lasting effect through pollution or contamination. Here, the meaning has been expanded to include the cyber environment, thus encompassing for example the loss or destruction of data. In future versions the cyber component will become a separate node.

Public Perception History has shown how public perception can affect policy, whether via debates, protests, or elections, whilst the previous chapter highlighted the differences in opinions on SaaW. Overall, it appears that SaaW is considered a 'clean weapon', particularly in comparison to traditional alternatives, which may greatly reduce resistance to their use.

Opponent's Military Power Military strength combined with the will and ability to project it are often considered to act as restraining elements. Whilst this can be assessed in regards to traditional weapons and domains, it is complicated by the cyber element.

Firstly, an opponent may decide to retaliate via conventional means, though this is unlikely in light of a 'mere' cyber attack as it would not be seen as proportionate. However, it may be the proverbial tip of the iceberg, particularly during times of rising tensions or conflict. Secondly, assessing relative military power within the cyber domain is problematic given, *inter alia*, the anonymity and limited-use nature (see Chapter 2.2).

Opponent's Alliances Alliances play an important role when an actor may be considered 'minor' or 'weak' but has strength through their partnerships, particularly military ones, such as NATO. Whether a cyber attack alone would lead to invocation of consultation articles or even collective defence ones (such as Articles 4 and 5, respectively, in NATO's case), is under discussion [8], [137].

4.6.2 Motivations

Clearly, without any motivation to develop and proliferate the best capabilities will not be utilised and are rendered irrelevant. Nonetheless, it should also be noted that negative motivation can act as a restraint, and *vice versa*.
Motivation itself is not a single entity but a combination of various interlinked and sometimes competing factors, which can be further split into five main factors:

- Fear of Retaliation
- Foreign Policy
- Domestic Policy
- Ideological Factors
- Economic Factors

Similarly to *Restraints*, these elements have several interconnected sub-levels:

- Public Perception
- Security
- Industry Interests
- Historic Factors
- Religious Factors

4.6.2.1 Fear of Retaliation

The lack of fear could be motivation itself, stemming from the opponent's inability or willingness to retaliate. The former may be due to the opponent's lack of power or alliances, or the attribution problem with its plausible deniability; the latter could stem from normative or policy constraints. In short, knowing reprisal is unlikely can thus act as an incentive.

4.6.2.2 Foreign Policy

Concerns about security often manifest themselves in an actor's foreign policy, or a non-state actor's equivalent, possibly as a response to threats to sovereignty.

This is in turn influenced by *Public Perception*, *Security*, *Industry Interests*, *Historic*, and *Religious Factors*.

4.6.2.3 Domestic Policy

This node represents the various internal pressures experienced by the actor, with the weighting dependent on the scenario. For example, public opinion might have a strong influence on one actor yet be irrelevant for another, whilst in others this could be the case for religious aspects. It reflects the actor's narrative and stance towards weapons in general, and will include defence against criminal activity, particularly for a state actor.

Non-state actors may not have domestic policy to manage; however, their public image (within the group and externally) is often of great importance, which can be seen in groups as different as Anonymous and Daesh.

Domestic Policy shares *Public Perception*, *Security*, *Industry Interests* and *Historic* factors with *Foreign Policy*, alongside of two additional ones, *Polity Data* and *Religious Factors*.

4.6.2.4 Ideological Factors

In this model, an actor's ideology is comprised of *Historic* and *Religious Factors*, currently not accounting for aspects such as ethics or morality.

4.6.2.5 Economic Factors

Modern society relies heavily on the cyber domain, so whilst for example stable state actors may not profit in the form of the traditional arms industry, they do so by providing a secure technological environment that can attract and defend businesses and investments.

Unstable or rogue states on the other hand may actively use malware for financial gain, similar to criminal organisations, terrorist groups or other nefarious non-state actors, as well as individuals. Alternatively, they may be involved in grey and black-market activities, from selling (or buying) botnets to zero-days. It should

be noted that legitimate forms of income utilising vulnerabilities also exist, for example in the form of bug bounty programs or penetration testing.

4.6.2.6 Sub-Levels

Public Perception If the narrative can present SaaW as a solution without collateral and relative low cost in comparison to traditional alternatives, opposition is likely to be limited.

Security This element has been a strong motivator for arms acquisition, and this is not changed by the cyber element, particularly in the case of a state actor. Threatening conditions are highly dependent on the actor but also the regional and global security situation, as has been shown for example in the nuclear domain. The threat, however, is no longer limited to 'weapon reach', security guarantees are less effective, nor is there a traditional period of mobilisation. Indeed, the potential adversary may not even be fully known. Another aspect is the fact that cyber infrastructure has become a core element of modern society, thus creating a myriad of additional attack vectors, and making it difficult to balance offensive and defensive elements. An analyst may choose to include comparative indicators, such as conflict history or enduring rivalries. This node has four components, *Global Security*, *Regional Security*, *National Security* and *International Prestige*. It should be noted that in this model, the first three of those components have an inverse state relationship with the *Security* node. For example, if the state of *Global Security* is set to 'low', the idea is that it is a 'high' motivator.

Global Security This node represents the level of global security as perceived by the actor, and its relevance greatly depends on the actor. Data is sourced from Militarized Interstate Dispute [233] and Correlates of War [260].

Regional Security Regional disputes, whether territorial, religious, or economic, have been a cause of instability and conflict for millennia. Whilst technology has transcended regional constraints, locality still plays an important

role, as can for example be seen in the relations between India and Pakistan, or the Ukraine and Russia. This node uses the same source data as *Global Security*.

National Security Whilst non-state actors do not have national security *per se*, they do have group security interests and often face numerous external threats. The nature of cyber weapons and their reach may allow a group to strike from the relative safety of another place, act as a force-multiplier or source of income, particularly if they lack capabilities for, or inclination towards, traditional weaponry.

International Prestige Important throughout history, from times of war to peaceful competition, there is an added complexity regarding cyber weapons: it is often of interest to an actor (particularly a state actor) to keep capabilities a secret and avoid overt showcasing, and to ensure plausible deniability after an attack. Attribution will still occur (and thus prestige will be gained or lost), but at a much higher level of uncertainty.
Non-state actors on the other hand may wish to advertise their capabilities and successful exploits, challenging their perceived status – for altruistic or nefarious reasons. It may also affect recruitment or economic gain, whilst individuals may seek 'bragging rights' or use prestige gained as an entry to specific groups. This node is also influenced by historic factors.

Industry Interests This aspect is vital in respect to cyber weapons, particularly given how interconnected and reliant modern society is upon this technology. State actors may derive scientific impetus from dual-use and civil applications, underlying economic pressures or the fear of being 'left behind'. Sheer curiosity is more likely to apply to individuals or organisations that are constrained by different accountabilities.
The traditional arms industry is a lucrative business for many actors, particularly from export weapons and weapon systems. The cyber domain does not have an equivalent, although security products could be considered part of this. However,

actors can profit by providing a technological environment that can attract (and defend) businesses and investments, or via service industries and start-ups. These also affect foreign and domestic policy, for example through trade agreements. Unstable or rogue states on the other hand may actively use malware for financial gain, similar to criminal organisations, terrorist groups or other nefarious non-state actors, as well as individuals. Alternatively, they may be involved in grey and black-market activity, from selling (or buying) botnets to zero-days [1], [30], [128], [296]. It should be noted that legitimate forms of income utilising vulnerabilities also exist, for example in the form of bug bounty programs or penetration testing. For non-state actors, the focus may be more one of having to stay ahead or to overcome barriers; individuals on the other hand may consider this as a challenge or a way to make themselves known. At the same time, security researchers may stumble across vulnerabilities and find attack vectors in the lab, hopefully adhering to responsible disclosure procedures and preventing attacks in the wild.

Historic Factors An actor's history may affect motivation to develop and proliferate weapons, whether as part of a conscious decision or cultural bias. For example, it provides some explanatory power as to why Germany does not engage in nuclear weapon technology, despite technical and economic capabilities. An actor's history directly affects foreign and domestic policy, but also contributes to an overall ideology. It should be noted that an actor's history may affect motivations differently when comparing cyber and traditional domains because of the nature of SaaW.

Religious Factors This source impetus is likely to be negligible for a large proportion of state actors, however, it can be a vital narrative for a small subset, as well as for non-state actors, and therefore warrants consideration. Initially set as a sub-element of domestic policy, the changing distribution of power and emerging actors, as well as the asymmetric nature of the cyber domain that is heavily weighted to the offensive [44], has propelled this factor to become a separate node with direct motivational effect.

4.6.3 Capabilities

An actor's capability is also of essence: without that, the only options to gain SaaW are to steal or buy. The capabilities can be split into six elements:

- Economic Factors
- Strategic
- Tactical
- Research
- Development
- Reconnaissance

4.6.3.1 Economic Factors

Cyber weapons are orders of magnitude cheaper than most traditional weapons, however they are often single use, have a limited lifespan and, if highly specialised, are only useful against a very specific system – particularly if collateral is to be avoided. Elements can of course be reused, for example Stuxnet and its cousins [24], and it can potentially be fully re-used until discovery. Thus, whilst cheaper, there is no single blueprint or mass-production ability – at least not for the 'top range'. All actors will therefore need to constantly invest in research and development, intelligence gathering, test beds/testing, hardware purchases, infrastructure, and human resources amongst other things. It should however also be noted that since Stuxnet, the cost for that 'weapon' would have decreased substantially due to evolving technology. Those actors not concerned about collateral or with less specialised targets are even less constrained financially, being able to adapt known malware, hire bots [195], [240] or hackers [18].

4.6.3.2 Strategic Factors

Scholarship, including at professional military education institutions, often breaks war into three levels: tactics, operations, and strategy [95]. Strategy is considered to be an idea of how to employ national instruments of power in a coherent manner to achieve certain objectives, whilst tactics refer to the order and use of various forces in relation to one another, and operations forms the 'glue' between both [95, p.17]. For state actors, there must be a strategy for using weapons in the cyber domain, whether a generic long-term plan or for specific aims, or even to allow for the possibility thereof. It should act as a guide of how various departments/teams will

interact in the future, without which cyber weapons are a loose end in military operations [238].

4.6.3.3 Tactical Factors

This element pertains to how strategic goals will be carried out 'on the ground'. For weapons in the cyber domain, this could for example entail the delivery or dissemination via cleverly placed USB sticks, or the management of a specific attack. State actors will need a coherent framework for action and conduct, neither of which limit non-state actors. However, non-state actors may lack the resources, from infrastructure to sufficiently trained personnel. The 'operational' element is currently not modelled.

It is partially dependent on an actor's *Military Spending* (or equivalent for non-state actors), as well as other factors such as the skills of the personnel and equipment, which are currently not represented separately. In turn, the tactical capability directly affects the delivery mechanism, particularly if it involves remote locations and requires inside access or other forms of direct physical access. This link is again more important for state actors, but even small groups or individuals need to consider this element.

4.6.3.4 Research

Previously forming a single node together with *Development*, it now encompasses all related research, whether academic, industrial or government based. For example, a specific vulnerability could be found via academic research in an institution that follows responsible disclosure procedures and has no skills in the utilisation and weaponisation thereof. It directly influences technical aspects (*Test Bed*, *Payload Package*, and *Delivery Mechanism*) but also informs the actor of possible capabilities (offensive and defensive). It should also be noted that this area is advancing at a much greater pace, similar to the overall domain. Furthermore, a large proportion of research is published or even open source, therefore allowing for the application and use by others.

4.6.3.5 Development

Similar to the economic aspects above, the constraints are based on what result is envisaged by the actor. Custom code requires extremely high expertise, whilst basic SQL injections or similar attacks are rudimentary: the difference is akin to bespoke architecture versus mass produced large panel system-building.
It is formed by the technical aspects of *Test Bed*, *Payload Package*, and *Delivery Mechanism* but it is likely to involve wider considerations, such as measures to prevent blow-back or collateral damage. In the case of Stuxnet for example, great care was taken to avoid collateral damage, unlike later indiscriminate attacks based on parts of the same code. It could be argued that state actors may be held to a higher level of responsibility, however, the question is whether this would be upheld in times of war.
A previous iteration of this model focussed heavily on the technical aspects, with various sub- and sub-sub-elements, for example breaking down delivery mechanisms into direct/indirect types, with insiders, externals, bot-based, drive-by-downloads, social engineering and so forth. Similarly, the payload element was broken down into various aspects such as code development, with its different attack vectors and obfuscation techniques. However, without a very specific scenario, technical experts and access to classified data it is not possible to gain more information than from only using the three sub-nodes *Test Bed*, *Payload Package*, and *Delivery Mechanism* (for examples, see Appendix B).
On a side note, expertise in this area has far fewer barriers to entry and is not limited to the traditional academic route: every day digital devices have increased greatly in processing power, have dropped in price and there is an endless number of resources available for free online.

Test Bed Experiments are used to test weapons by confirming their usability and reliability, to measure potency and to improve the technology. New weapons are tried out in laboratories, in designated areas, ranges or underground, often targeting mock-up buildings. The majority of weapons nowadays are created

as part of production lines, becoming more and more correct by design, limited to testing to random samples.
In the context of cyber, this however is not the case. Whilst a large proportion of kinetic weapons can be used on a variety of targets, SaaW usually require specific vulnerabilities to have the desired effect. This is particularly the case when attacking a highly specific system, necessitating a *Test Bed* acting as mock-up of the target. The importance and sophistication of the *Test Bed* varies according to the type and goal of the envisaged SaaW.

Payload Package Whilst splitting the payload and delivery of SaaW in this manner is a gross oversimplification from a technical perspective, it does serve the conceptual purpose of this model.
Here, the *Payload Package* refers to the computer code that causes harm in the target system. This also includes components for propagation, detection evasion and sustained communication with the attacker or deletion in cases this is applicable. Future work should include an expansion of the technical intricacies, such as high-level suggestions shown in Appendix B, which include bot(net) and C&C components, or more detailed build, install and delivery mechanisms.

Delivery Mechanism Similarly to the node above, this node does not accurately reflect the division of labour but is sufficient for this discussion. Here, this node refers to means of transporting the *Payload Package* to the target system. This could be achieved through a variety of means, ranging from exploits, drive-bys and man-in-the-middle attacks to phishing or well-placed USB sticks.

4.6.3.6 Intelligence (Intel) & Reconnaissance (Recce)

Also known as *recon*, this node refers to intelligence gathering, such as information about the target systems' hard- and software components, the location, as well as the human element. It is to some extent limited by the financing available, which for many states comes from a combination of military assets and *Military Spending*, as well as other intelligence services. A further source of information

may be bought or stolen data, which may be more commonly used by non-state actors lacking resources for their own operations. This outsourcing option seems likely to play a far greater role in the cyber domain, particularly due to increased anonymity and plausible deniability.

Information was originally a separate node, with sub-elements including media (national and foreign), intelligence gathering, covert sources, diplomatic channels, and information sharing. However, it was collapsed into this node following discussions, mainly because it would be very difficult to obtain data and assign values.

It directly influences technical aspects (*Test Bed*, *Payload Package*, and *Delivery Mechanism*), as it provides the information for the target system and access points.

4.6.3.7 Sub-Levels

Military Spending This sub-node is influenced by the *Economic Factors*, as military spending is often considered a function of economic capabilities (although several counterexamples exist).

GDP The economy of an actor, particularly a state, can be measured in various ways, with the Gross Domestic Product (GDP) being one of the most common ones. It represents the market value of all the goods and services produced in a set period of time, often used to denote the economic performance and to make international comparisons. However, this does not reflect differences between states, such as inflation or cost of living, thus GDP *per capita* at Purchasing Power Parity (PPP) is used in studies to compare living standards. The data-sources are *Angus Maddison* [314] and the *Penn World Table* [315].

4.7 Conclusion

This chapter centred on the factors that lead an actor to develop and proliferate in the cyber domain, exploring a different facet of SaaW. It drew on the previous chapter which examined what it means for software to be a weapon and expanded

the literature review from Chapter 2 with a focus on weapons and proliferation theories that were then applied to SaaW. This allowed for the creation of a conceptual model that can act as a unified starting point for a variety of stakeholders. It does this by dividing SaaW pursuit into three interconnected pillars, namely *restraints*, *motivations*, and *capabilities*, with numerous elements and sub-elements.

Restraints reflect the reasons for an actor not to develop or proliferate SaaW, which can be used to analyse what elements can limit, or even control, further escalation. This pillar has four main elements: *Potential Collateral; Domestic Safeguards; International Agreements; and Fear of Retaliation. Motivations* on the other hand account for reasons why an actor may pursue SaaW and has five main elements: *Fear of Retaliation; Foreign Policy; Domestic Policy; Ideological Factors; and Economic Factors.* The last pillar refers to *Capabilities*, without which options of pursuit are limited to theft or purchase, and has six main elements: *Economic; Strategic; Tactical; Research; Development; and Reconnaissance.*

This chapter discussed each in turn, examining interrelationships and sub-elements, both competing and complementing, laying the foundation for the operational model presented in the next chapter. There, the foundations of BN will be presented, including probability and reasoning about uncertainty, before variables, connections and data acquisition are discussed. This provides the background for the three case studies in Chapter 6.

5
Operational Model

Contents

While Chapter 3 explored what it means for software to be a weapon, the previous chapter created a conceptual model of what factors influence an actor

to develop and proliferate in the cyber domain. Aside from the literature review, it drew on weapon and proliferation theories to create a group of interconnected determinants that act as variables, split into three main pillars representing *restraints*, *motivations* and *capabilities*. This fifth chapter now builds upon this using BN to create an operational model, which is then used for the three case studies presented in Chapter 6. It helps answer instances of the third research question: *What is the probability that a given actor is pursuing SaaW?*

It begins with an argument for the use of BN, including previous work on the application to weapon proliferation, followed by a foundational discussion, including uncertainty, reasoning, and dependency. An introduction of the underlying principle of conditional probability is next, followed by a presentation of data and knowledge acquisition. Next, this chapter turns more specifically towards SaaW, shedding light on potential alternative implementations, prior to discussing how the conceptual model from Chapter 4 has been operationalised. This background on BN and its application provides the foundation for the three case studies in Chapter 6.

5.1 Bayesian Use Cases

BNs were chosen to implement the conceptual model from Chapter 4 as they aid reasoning about knowledge that has levels of uncertainty, allow for the combination of quantitative and qualitative data and because they provide an intuitive visual form that is supported by a formal mechanism.

They also have a long history of being applied in variety of situations, ranging from medical diagnosis [209], [273], device troubleshooting [40], [170], to digital forensics [179]–[181]. They have also been applied to the question of proliferation, albeit within the context of nuclear weapons, with three works discussed below [71], [110], [147]. The first of these was submitted as a thesis in 2008 by Holcombe [147], who developed a BN to assess the probability of nuclear proliferation and applied it to two case studies, India and Iran. His work centres on describing the various causes and capabilities using conditional probabilities based on political theories and expert

opinions. He first discusses the different (political) theories and models of nuclear proliferation, divided into three main areas: external sources, domestic pressures, and individual influences. These are then used to create a set of determinants, leading to four groups of factors: intentions, including security threats, prestige, scientific agendas and domestic policy; capabilities, such as economic, tactical and technical; restraints created by international agreements and safeguards; and lastly, actions, encompassing intelligence reports and media/news coverage. His initial states for factors are divided into 'not contributing (N)', 'possibly contributing (P)' and 'definitely contributing (D)'. An analyst splits the overall probability and the values are initially set to the least threatening ones. The relative weights between the factors are set equal until applied to a specific scenario. The case study on India was run with events for each calendar year between 1946 and 2000 in order to determine the overall annual proliferation risk and to validate the model. It was run on an earlier model and therefore did not include a factor group (actors) added later, and the values of the relative weights between the factor groups were kept constant for the first iteration, then adjusted for the second. His analysis of Iran is described in more detail and follows the same methodology. In short, he showed that his immediate goal of creating an analytical system integrating disparate information to provide unbiased probability of proliferation risk was successful. Furthermore, he acknowledges that access to expert opinion, as well as classified information could greatly improve accuracy.

Another thesis submitted in the same year by Freeman [110] focussed on the acquisition aspect of nuclear proliferation, implementing a BN for analysis. Whereas traditionally nation-states that are susceptible to ideas of deterrence took centre-stage in the proliferation discourse, Freeman includes the threat posed by other actors, such as terrorist entities. His aim was to gain an understanding of the various pathways that may lead to the procurement of nuclear weapons based on existing evidence. He developed a methodology using BNs, evaluating key resources, as well as motivations, to form prior probabilities for various pathways of a given actor. His overall BN is divided into several sections: intentions and

resources input, including deliverability, yield and capabilities; obtaining Highly Enriched Uranium, including enrichment capabilities and uranium feed portions; reprocessing, including equipment and spent/reprocessed fuel; Special Nuclear Materials, the acquisition of which is a pre-requisite for weaponisation, either by indigenous development, purchase or theft; the weaponisation, including the weapon package, as well as pusher/tamper/reflector components; the completed nuclear device, including purchase or theft pathways. Unlike Holcombe, Freeman did not use specific case studies to apply his BN but created four (fictional) scenarios: a small fanatical religious group; a semi-developed rogue nation-state; an international extremist group with nation-state sponsorship; and lastly, a developed nation-state. His results show that in the case of nuclear proliferation, an actor's motivations influenced the pathways more than the resources available, but the resources had a greater bias on the overall success chance. Aside from constraints of input affecting the results, limitations of the model include it being unable to account for parallel weapon programs or deception.

The third is a paper by Coles et al. [71], who illustrated how a BN model of social factors can contribute to nuclear proliferation assessments of a generic model state. Alongside factors such as political, economic, and security elements, Coles' model also includes psychological aspects. That model used datasets [159], [286], with the appendix providing a categorisation and detailed source attribution but the model is still under development [334].

The following section will provide a brief overview of Bayesian theory and principles before later parts apply it to the pursuit of SaaW.

5.2 Probability & Bayesian Networks

The introduction of prior knowledge is not permitted as part of classical inferential models, due to concerns of adding information that may skew experimental results. Yet in reality, there are many instances where prior knowledge contributes to making informed decision – a strength of BNs.

The first person to create an equation that allowed for new evidence to be taken into account and update beliefs was Reverend Thomas Bayes, whose work was published posthumously in 1763 [20] including amendments by Richard Price. His work included details on prior probability alongside of conditional probability theorems that have become the foundation of Bayes' Theorem. Laplace developed this further [185], adding a probability-based approach to inductive reasoning.

BNs have been applied successfully for several decades across domains, providing a convenient framework for encoding knowledge pertaining to uncertainty and the reasoning thereof. From modelling medical diagnosis [209], [273] to commercial trouble-shooting [40], [170] or financial models. They elegantly combine qualitative and quantitative aspects, depicted as an intuitive graphical interface that is based on a formal mathematical framework. Furthermore, they can be constructed from very differing sources, combining human expert knowledge with data.

The majority of BNs constructed centre on discrete variables, as continuous ones pose a far greater challenge in terms of representation and inference. Whereas probabilistic interactions of the former can be enumerated exhaustively in CPTs, this is not adequate for the latter, nor are there suitable alternatives. Furthermore, no universal inference algorithm exists for continuous variables, thus limiting applicability to very specific cases and restrictive scenarios [188]. Hereafter, the term variable is used to denote discrete variables, unless otherwise stated.

5.2.1 Uncertainty

Unlike computing processes that are deterministic in nature, a large aspect of human life involves various levels of uncertainty, with examples ranging from weather patterns and traffic forecasts to financial markets. Planning involves weighing up what is likely and what not, based on experiences and data, whether casually or in a formalised sense of a *degree of belief in a proposition.*

In the context of analysing SaaW development and proliferation, there are a large number of uncertainties pertaining to data and the analyst: the former might be inaccurate or lacking; the latter will be biased by experiences and motivations. Thus,

uncertainty becomes a key component, and challenge, for arriving at conclusions via inferences. Whilst there will always be uncertainty, the levels of ambiguity can be reduced [156], with probability being one of the most prevalent approaches.
Similar to digital forensic evidence, where *probabilistic dependence* exists between the forensic hypothesis and the supporting evidence [180], that relationship also exists in the context of SaaW between the various factors and the overall hypothesis forming conditional probability.
However, it should be noted that probability is not without problems: firstly, quantitative data (or a number) is not necessarily available; secondly, it means that comparisons are possible – whilst this is usually good, it raises the issue of incomparable likelihoods [131]. Alternative measures, including their advantages and disadvantages, such as *Dempster-Shafer belief functions*, *possibility measures*, and *ranking functions* are not within scope.

5.2.2 Observation & Reasoning

Phases of observation and analysis vary from one domain and subject matter to another. Nonetheless, common elements include a form of *identification*, *observation*, *analysis* and *presentation*, which is very similar to the digital forensic process [205]. Whilst weapon development and proliferation analysis is rarely created for courts of law, it is used by analysts, the military and policy-makers.
In this type of analysis, there are two main types of uncertainty: the first is based upon the unreliable observation of data, which includes unreliable or missing data, as well as the selection thereof; the second type rests upon the analysts' human nature, their interpretations based on, *inter alia*, bias, motivation, experience, skill-set, and mood.
Given the context of SaaW, where data (current and historic) is lacking, the weight of the interpretation is often heavily skewed towards the analyst. Kwan [180, p.14] stated that "since it is a common phenomenon that evidence is always incomplete, rarely conclusive, and imprecise, it is necessary to establish a reasoning model so that

the reliability of the forensic conclusion can be upheld". Although he said this in the context of digital forensics, it is equally true for SaaW development and proliferation. Given these uncertainties in observation and reasoning, probabilistic inference is inherently involved. Linking data with a hypothesis 'tells the story' but an analyst's reasoning has to be justified and hold up under scrutiny. Thus, a model is required that can not only handle uncertainty and limit bias, but one that can also be examined and evaluated.

5.2.3 Dependency

The probability of the final conclusion is dependent on the various factors, or evidences, contributing to it, known as *probabilistic dependency* or *conditional probability*. This relationship is represented by inductive as well as deductive elements: the former by the *belief of evidence given the hypothesis*, often referred to as *likelihood*; the latter is the reverse, it is the *belief of the hypothesis given the evidence*, known as *inference*.

When multiple causes contribute to a common event independently, these events are *casually independent* to each other. If the overall hypothesis is dependent on several events, it is said to be based on the *joint probability* of these. Furthermore, events or evidences exist in a form limbo, known as *conditional interdependence*: when nothing is known about the hypothesis' state, the events are all dependent; yet, if the hypothesis' state is certain, the events are independent.

5.3 Bayesian Foundations

The Bayesian approach is a common way of modelling and analysing inference and likelihood, particularly given cases of dealing with conditional (in)dependencies of hypothesis and evidence, as well as computational complexity of the Joint Probability Distribution (JPD) [180, p.25].

BNs are a type of casual belief network depicted as a DAG, encoding a factorisation of a JPD, and representing the dependency model of the interest area, either explicitly or implicitly. The graph structure is represented by directional arcs that

show casual influences between the linked stochastic variables depicted as nodes. The probability distributions over individual variables are conditional on their parents, representing individual factorisation components.

For BNs where all nodes are discrete, the conditional probability distributions are stored in CPTs, indexing all possible combinations of states. Defining these quickly becomes problematic, or even impossible from a practical point of view, due to the exponential nature [342].

5.3.1 Conditional Probability

The basic underlying principle of BNs is *conditional probability*, representing the probability of an event A given event B, which is mathematically denoted and defined by:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \tag{5.1}$$

Combined with the *fundamental rule* for probabilities, where $P(A, B)$ is the probability of the joint event $(A \cap B)$:

$$\begin{aligned} P(A \mid B)P(B) &= P(A, B) \\ P(B \mid A)P(A) &= P(A, B) \end{aligned} \tag{5.2}$$

Following on from this, *Bayes' Theorem* is denoted as:

$$P(A \mid B) = \frac{P(B \mid A)\, P(A)}{P(B)} \tag{5.3}$$

The Bayesian perspective represents the degree of belief in a hypothesis as the probability thereof, for example $P(A)$, being true. When new evidence is gained in the form of B, it affects the degree of belief in the hypothesis, thus becoming $P(A \mid B)$ – either strengthening or weakening the initial hypothesis.

In other terms, $P(B \mid A)$ is the posterior probability, whilst $P(B)$ denotes the prior probability of the B before any information on A is known or observed; $P(A)$ is the prior probability of A, also known as the normalising constant. In words, this gives the relationship of:

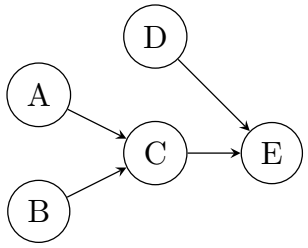


Figure 5.1: Example DAG

		B			
		b_1	b_2	b_3	b_4
A	a_1	0.7	0.6	0.5	0.2
	a_2	0.1	0.2	0.3	0.5
	a_3	0.2	0.2	0.2	0.3

Note: For each b_j,$\sum_A P(A \mid b_j) = 1$

Figure 5.2: Tabulated view of variable state combinations

$$likelihood = \frac{posterior\ probability\ \text{x}\ normalising\ constant}{prior\ probability} \tag{5.4}$$

and it can be rewritten as:

$$posterior\ probability\ \propto\ likelihood\ \text{x}\ prior\ probability \tag{5.5}$$

5.3.2 Variables

Variables can have more than one state, with the probability distribution summing to 1. For variable A with states $a_1, ..., a_n$, $P(A \mid B) = (x_1, ..., x_n)$ gives the probability distribution $\sum_{i=1}^{n} x_i = 1$. The combination of states for each variable is an n x m table and for variables with several states (5.2) becomes:

$$P(a_i \mid b_j)P(b_j) = P(a_i, b_j) \tag{5.6}$$

which can also be represented in a tabulated manner as shown in Table 5.2. Similarly, the fundamental rule can be applied to variables $P(A \mid B)P(B) = P(A, B)$ and also be depicted in a tabular manner (see Figure 5.2).

The number of combinations grows exponentially, and thus poses a great challenge if data has to be sourced from experts and cannot be learnt or imported from databases. Of the five nodes depicted in Figure 5.1, three are parent-less: A, B, and D. If they each have four states (such as *high* – *h*, *medium* – *m*, *low* – *l*, *unknown* – *u*), they have an independent fixed prior probability distribution across their states recorded in a *Marginal Probability Table* (MPT), with each summing to 1 (see Table 5.1).

Table 5.1: Example MPTs

Var	State	Prob
A	H	0.65
	M	0.1
	L	0.2
	U	0.05

Var	State	Prob
B	H	0.3
	M	0.4
	L	0.1
	U	0.2

Var	State	Prob
D	H	0.3
	M	0.25
	L	0.2
	U	0.25

Two have parents and thus are formed by CPTs, with C being dependent on A and B, whilst E is dependent on C and D.

5.3.3 Connections

The qualitative part is depicted as a DAG, with the nodes representing the casual relationship between the variables and the hierarchy creating the logical structure of the model. This is then combined with a quantitative counterpart in the form of conditional probability functions, which define the relationships between nodes. These connections between nodes take three basic forms of *serial*, *converging* and *diverging*, as seen in Figure 5.3 and the equations in 5.7.

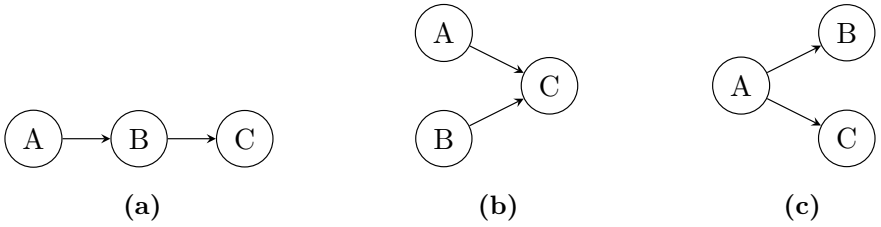


Figure 5.3: Connection types: (a) serial, (b) converging, and (c) diverging

$$\begin{aligned} \text{For serial:} \quad & P(A,B,C) = P(C \mid A)P(B \mid A)P(A) \\ \text{For converging:} \quad & P(A,B,C) = P(C \mid A,B)P(A)P(B) \\ \text{For diverging:} \quad & P(A,B,C) = P(C \mid A)P(B \mid A)P(A) \end{aligned} \tag{5.7}$$

5.3.4 D-Separation & Markov Property

Two variables are considered to be *D-Separated* if the information between two nodes is essentially 'blocked' by some evidence about the nodes in between, informing (and

limiting) how information is passed on. For example, taking Figure 5.3(a), A and C have B to act as a block in a serial connection; for converging ones 5.3(b), A and B require C to be unknown; the last option, 5.3(c), covers diverging connections, where B and C need A to be known. A further element that dictates the conditional independence of a variable or node is the casual *Markov Property*. The *Markov Blanket* covers a node, its parents, its children, as well as their parents, and makes that node conditionally independent of any others. A benefit thereof is that it can reduce the resources required for computation of BN inference by reducing the parameters needed for the calculation of the JPD. Under the probability chain rule, also known as the general product rule, the joint probabilities of $P(x_1, ..., x_n)$ are

$$\prod_{i=1}^{n} P(x_i \mid x_1, ..., x_{i-1}) \tag{5.8}$$

This makes the number of parameters required infeasible given the exponential nature. But given the Markov Property, this can be reduced greatly by factorising the JPD. Thus,

$$P(x_1, ..., x_n) = \prod_{i=1}^{n} P(x_i \mid parent(x_1)) \tag{5.9}$$

5.4 Data & Knowledge Acquisition

Knowledge acquisition can often pose a substantial challenge for BNs, particularly if there are no datasets to learn from, thus requiring expert opinions to populate CPTs that grow exponentially with the number of associated parent-nodes.

5.4.1 Data Sets

Data sets can either provide information for a specific node, a collection thereof or even inform the creation of the whole network. In data-rich environments, Learning Bayesian Networks (LBNs) can be utilised, either *constraint* or *score-based* ones. The former relies on testing conditional independences between variables, whilst the latter scores the network's data fit [75], [138], [234], [291]. However, this is not the

case in the context of SaaW and the focus is therefore on specific (quantitative) data sets that can contribute to specific nodes, combined knowledge gained from experts.

5.4.2 Expert Opinion

As discussed previously, there are many real-world applications that rely on expert elicitation, although a current trend is to slowly phase this out in favour of big-data coupled with learning algorithms. Nonetheless, some research suggests that incorporating expert opinions can result in much better models [74], [244], [346]. However, one key challenge is to ensure accuracy when eliciting knowledge, particularly if the expert does not have mathematical or related understanding [215]. Alternatives have thus been created, from language used [246], [319] or iterative processes [96] to a reduction of overall questions and data required [85].
Another difficulty stems from the inherent bias in expert opinions beyond that inherent in any human judgement. This can be due to the limited understanding of probability and statistics, which can threaten both validity and reliability, or the unfamiliarity of the expert on this particular aspect of the topic. There are several ways to mitigate this, including standardised scripts, examples, training exercises or feedback and revision processes [160], [163], [222].
A complete process to manage elicitation based on the experiences of the NUREG 1150 study was proposed in the early 90s [165]. After shortcomings were identified in the initial draft, the elicitation process was formalised to include seven stages [165, p.197]:

1. Identification and selection of the issues
2. Identification and selection of the experts
3. Discussion and refinement of the issues
4. Training for elicitation
5. Elicitation
6. Analysis, aggregation, and resolution of disagreements
7. Documentation and communication

Furthermore, the prior distributions and CPTs are set out openly, allowing other domain experts to challenge or refute these.

5.4.3 Das' Weighted Sum

In order to mitigate BNs inherent exponential problem of populating CPTs, particularly when using experts, Das [85] devised an algorithm by weighting the relative parent-node's influence strength and using CPCs to create a linearly growing set of probability distributions. Both aspects are elicited from domain experts, and CPTs are then populated by computing appropriate weighted sums. Das uses information geometry to show this accurately captures experts' judgemental strategies. He first introduces an example of a small business, where the aim is to assess levels of 'efficiency' (E) based on 'personal morale' (PM), 'personnel training' (PT) and 'managerial expertise' (ME). Converted into a basic BN, E becomes a child-node (E) with three parents (PM, PT, ME), each of which has five states (very low – vl, low – l, average – a, high – h, very high – vh). Denoting a typical *parental configuration* as Π consisting of three elements, the probability distributions for the CPT have the form of:

$$\begin{aligned}\{p(E = vl|\Pi), p(E = l|\Pi), p(E = a|\Pi),\\ p(E = h|\Pi), p(E = vh|\Pi)\}\end{aligned} \tag{5.10}$$

Despite this small scenario, the challenge of gaining expert opinions can already be seen, as it would already require 5^3 parental distributions, and thus questions asked. Das therefore proposes the use of CPCs, which refer to parent state combinations that *make sense* to an expert. Using these, particularly if there is a one-to-one mapping of parent states, the number of questions that need to be asked drop significantly, resulting in a linear instead of exponential growth based on parent numbers. Continuing his example above, an expert would subjectively interpret the compatible parental configurations as:

$$\begin{aligned}\{Comp(PM = s)\} \equiv \{PM = s, PT = s, ME = s\},\\ for \quad s = vl, l, a, h, vh\end{aligned} \tag{5.11}$$

Taking compatibility as equivalence:

$$\begin{aligned} \{Comp(PM = s)\} &\equiv \{Comp(PT = s)\} \\ &\equiv \{Comp(ME = s)\} \\ &\equiv \{PM = s, PT = s, ME = s\}, \\ &\quad for \quad s = vl, l, a, h, vh \end{aligned} \tag{5.12}$$

In turn, this leads to the following probability distribution for E:

$$\begin{aligned} &\{p(E = e|\{Comp(PM = s)\} \\ &= p(E = e|\{Comp(PT = s)\} \\ &= p(E = e|\{Comp(ME = s)\}, \\ &\quad for \quad e, s = vl, l, a, h, vh \end{aligned} \tag{5.13}$$

Thus, if the expert were to provide the distributions corresponding to the parental configurations, all 5x3 CPCs are obtained. However, this still leaves the non-compatible states which are required to populate the CPT. To fill these in, Das applied a weighted sum algorithm that only requires the relative weighting $(w_1, ..., w_n)$ and the $k_1, ..., k_n$ probability distributions CPCs:

$$\begin{aligned} &p(x^l|y_{1^{s_1}}, y_{2^{s_2}}, ..., y_{n^{s_n}}) \\ &= \sum_{j=l}^{n} w_j p(x^1|\{Comp(Y_j = y_{j^{s_j}})\}), \\ with \quad l = 0, 1, ..., m \quad and \quad s_j = 1, 2, ..., k_j \end{aligned} \tag{5.14}$$

For incompatible distributions, weights assigned by experts are used:

$$\begin{aligned} \{p(E = e|PM = vh, PT = vl, MT = vl) = \\ w_1 p(E = e|\{Comp(PM = vh)\} \\ +w_2 p(E = e|\{Comp(PT = vl)\} \\ +w_3 p(E = e|\{Comp(ME = vl)\} \\ for \quad e = vl, l, a, h, vh \end{aligned} \tag{5.15}$$

5.5 Evaluation

A problem of BN is that the model does not have an inbuilt method to check for extremely sensitive or even incorrect posterior inference. However, reliability, as well as accuracy, of the posterior hypothesis can be compared against prior probabilities and likelihood in the form of a sensitivity analysis [229], [230].
This type of evaluation examines the changes it outputs based on iterative input variations [258]. When examining robustness, the structure and priors are the most important aspects to consider [118], with the focus normally limited to the quantitative aspect of the BN [77].
The simplest form of conducting a sensitivity analysis involves changing all the variables and their possible combinations – one parameter at a time. This quickly becomes infeasible [186] and leads to a first-order approximation technique, which, however, can breakdown when larger variations are considered [171]. To mitigate this, a method was presented for "computing the coefficients in the functions for all possible parameters, using just one propagation in a junction tree" [171, p.318].

5.6 SaaW Application

As seen in the literature review at the beginning of this chapter, applying BN to weapon development and proliferation is not new. However, the application to the cyber domain is, particularly given the different nature of SaaW. Alternative approaches are briefly discussed before applying the related discourse on proliferation from Chapter 4 and expanding [277].

5.6.1 Alternative Approaches

Whilst Bayesian approaches have been considered appropriate to interpret the degree of belief of scenarios given prior probabilities, it would be remiss not to mention potential other approaches.
Game Theory would essentially require a conflict or contest between at least two actors, interacting in a game governed by a set of rules. Information can be either

perfect or not, with the latter posing a challenge as there is no clear strategy to winning. Neumann proved that all two-player zero-sum games have optimal strategies, which are a random mix of individual ones, assuming the game is played many times over. However, games can have multiple players, leading to coalitions against others, and they are not limited to being zero-sum, leading to a generalised optimal – the equilibrium solution. This solution is a combination of strategies for each player set in such a way that there is no reason to deviate, unless other players do. One drawback is that there can often be more than one best-response equilibrium, or even an infinite number of them.

In 1960, Schelling was the first to establish the study of bargaining and strategic behaviour in his seminal work [263], shining a new light on game theory and its application. Since then, game theory has been applied to numerous aspects, for example super-powers [38] or conflict resolution [242]. However, it does not really solve the problem for the question at hand, but it could offer a different perspective. For example, given only one actor is currently examined at a time, it could be used to pit *Restraints* and *Motivations* of one actor against each other. Furthermore, it might allow for a better depiction of the reciprocal nature of war, which this current envisaged BN does not, unless two (or more) competing ones are run at the same time. Yet, the game theoretic approach would also require a very clear and detailed rule set and win-conditions, yet rules in the cyber domain are anything but clear. Furthermore, some information is often hidden, adding another stumbling block for a game theoretic approach. There would also be the question of how to map the varying factors contributing: would they be actors or amalgamated into strategies? Another approach would have been an implementation using MN, which, unlike BNs are an undirected graphical model. Whilst it offers the advantage of allowing for cyclic relations, the lack of directionality does not suit modelling causal relationships, which exist in the context of SaaW pursuit.

5.6.2 Contributing Factors: Nodes

Whilst a brief discussion of factors follows, the model is most easily accessible in its graphical form, with the variables depicted as nodes and their dependencies as arcs (Figure 5.4).

5.6.2.1 Interrelationships: Arcs

Whilst the various factors will in reality have bi-directional interactions and contain feedback loops, BN constraints enforce directional. Initial nodes and arcs were selected based on literature reviews and informal discussions, which were then fine-tuned resulting in the current model.

The CPTs for some nodes in this model are relatively large due to nodes with up to six parents with four states each. Yet, more should be added to reflect reality more accurately. They were created using Das' approach discussed in Section 5.4.3.

5.6.2.2 Leaf Nodes: Priors

A number of nodes essentially act as initiators by selecting one state from the prior probabilities. These probabilities are drawn from quantitative data or elicited from experts (see 5.6.3).

The elicited leaf nodes in this model are (left to right in Figure 5.4): *Human* and *Environmental Collateral*; *Opponent's Military Power* and *Opponent's Alliances*; *National Security*; *Historic Factors*; *Religious Factors*; *Strategic Factors*; and lastly, *Tactical Factors*.

5.6.2.3 Internal Node States

Many BN use binary states for nodes denoting 'yes' and 'no', and are for example applied in digital forensics to denote the presence or absence of an evidential trace [179], with some allowing for a small fraction of uncertainty via a third state. For this topic and model at hand, the binary approach appears to be too limiting, leading therefore to four states: *high*, *medium*, *low* (with 'good', 'average', 'poor'

or equivalent substitutes where the former do not make sense conceptually) and *unsure*, with the last one explicitly accounting for uncertainty.

However, this poses a challenge as these categories attempt to capture fluid and abstract concepts: one person's 'high' may be another's 'medium'. In order to mitigate this alongside other respondents' biases, all the responses per scenario were merged and averaged.

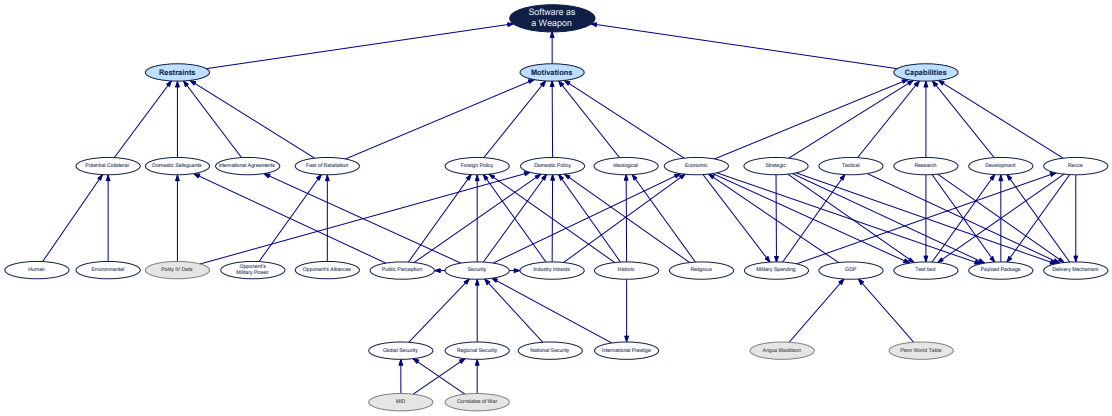


Figure 5.4: Overview of Whole Model. *Note: Grey nodes refer to datasets*

5.6.2.4 Restraints

The restraints are based on the conceptual model from Chapter 4, with a detailed discussion of elements presented in Section 4.6.1. The initial model had four elements, consisting of *International Agreements*, *External Threats*, *Domestic Barriers* and *Environmental Risks*. Here, similar ones are proposed consisting of *International Agreements*, the *Fear of Retaliation*, *Domestic Safeguards*, and *Potential Collateral Damage*. The first three are the same, however two have been relabelled to be more specific. Furthermore, the *Fear of Retaliation* has been split into *Opponent's Military Power* and *Opponent's Alliances*. *Environmental Risks* has become a parent node to *Potential Collateral Damage*, along with *Human Risk*.

5.6.2.5 Motivations

Similarly, the motivations are based on the conceptual model and details can be found in Section 4.6.2. However, for the operational model, these have been re-

arranged, with the main difference being *External Threats/National Security* and *Prestige* now being a parent node to *Foreign* and *Domestic Policies* via a *Security* node. Similarly, *Public Opinion* also affects these as opposed to *Motivations* directly.

5.6.2.6 Capabilities

Akin to the two other pillars, *Capabilities* is based on the previous chapter (see Section 4.6.3). Early versions did not include *Intelligence/Reconnaissance* (Intel/Recce) and *Technical/Scientific* was one node instead of being split into *Research* and *Development*. Whilst the latter two are often considered integrated aspects, including in company departments being designated as 'R&D', it is argued that in the context of cyber, they do not require that close a relationship and can come from unrelated sources.

5.6.3 Knowledge Acquisition

As discussed in the previous section, knowledge acquisition can often pose a substantial challenge for BN. This section addresses the limited datasets that are applicable and information on the interviews conducted implementing Das' approach (see Section 5.4.3). It is very likely that a number of proprietary or classified datasets exist, however, they are not readily available.

5.6.3.1 Existing Datasets

Unlike the nuclear realm, datasets are very sparse and often not publicly available. The ones that have been used are *Polity V* [198], contributing to *Domestic Safeguards* and *Domestic Policy*; *Militarized Interstate Dispute* [233] and *Correlates of War* [260] into *Global* and *Regional Security*; and lastly, two sets of data measuring GDP being Angus Maddison [33] and the Penn World Table [104]. It should be noted that these data sets currently 'lag' behind several years, however, they are sufficient for the example scenarios.

Two further datasets that have been considered are *World Religion Data* [196] and *Formal Alliances* [119]. The former may contribute to the *Religion* node, however, the adherence to religion within an area does not necessarily affect an actor's

religious impetus. The latter would affect *International Agreements* and *Opponent's Alliances* and will likely be included in future iterations of the model.

5.6.3.2 Weighted Sum Algorithm

Das' methodology is implemented for the three scenarios discussed below, as otherwise several nodes would need a large number of probability distributions to be elicited. *Domestic Policy*, for example, would require 4^6 (4096) values alone. CPCs have been pre-defined, as otherwise each respondent would potentially have a separate set in mind.

5.6.3.3 Interviews

Ethical approval for the interviews was gained under CUREC. Following an initial pilot study, 30 interviews were conducted, each lasting approximately one hour. After a few questions regarding topic familiarity and expertise, the model and task were explained. Each node (variable) and connects were discussed, giving the interviewee the chance to ask questions and comment.

Once the interviewee was comfortable with the model, the three scenarios (detailed below) were explained and then iterated upon, providing initial priors for nodes without parents, relative weighting between parents, as well as compatible configurations for each one. Comments of a qualitative nature were also recorded. The process followed the steps suggested by [165], albeit in a far more time-constrained manner. The data from the interviews for each scenario was combined, creating a single set of values per scenario.

5.6.3.4 Alternative Data Sources

A solution to large CPTs has been the application of parametric conditional distributions, of which the most common ones are the binary noisy-OR [235] and the multi-valued noisy-MAX models [90], [139]. These models offer a logarithmic reduction of the number of parameters required to complete CPTs in respect of growing parent numbers. However, they do this by assuming independence, which works well in scenarios such as the *Burglar – Earthquake – Alarm* model but differs

from the relationship between variables in the model presented here. However, a robust comparison akin to Zagorecki & Druzdzel [343] has not been completed. Another solution, albeit out of scope of the current paper, would be a re-design of the model itself, attempting to avoid nodes with a large number of parents, particularly when further granularity is added. This could either be done 'manually' when creating the network, or it could be applied by utilising CfPD [228], [250], [322].

5.6.4 Scenarios

Three different scenarios were used for this case study and they are discussed in detail in Chapter 6. They are defined by different actor types: a *Generic*, a *State*, and a *Terrorist* one. The aim of the first scenario was to familiarise the interviewee with the task further, and to establish a base line. The second one used the United Kingdom as a basis, with the software weapon sharing similarity to Stuxnet but being more advanced. The last scenario was a terrorist non-state actor with strong religious impetus seeking to create disruption and destruction, as opposed to a sophisticated and targeted weapon.

5.6.5 BN Software

While Python was first used to implement the operational model, the latter and final versions were created and examined using the academic version of GeNIe [21] by BayesFusion LLC, as it provided a better user experience. This software allows for the creation and evaluation of various types of nodes, their connections, and associated priors as well as CPTs. The model and section overviews are screen-shots from the system.

5.7 Conclusion

This chapter introduced BNs, beginning with underlying concepts of uncertainty and conditional probability. It familiarised the reader with how variables are managed in terms of internal states, as well as external interconnections, before broaching the topic of data acquisition and alternative approaches.

A BN was then applied to SaaW, using the conceptual model from Chapter 4 as a basis, which examined what factors influence an actor to develop and proliferate in the cyber domain. The qualitative part is depicted as a DAG, with the nodes representing the casual relationship between the variables and the hierarchy creating the logical structure of the model. This is then combined with a quantitative counterpart in the form of conditional probability functions, which define the relationships between nodes. The resulting operational model rests on the three main pillars discussed in the previous chapter, namely *restraints*, *motivations* and *capabilities*. It features an intuitive graphical model with a robust underlying structure that models an actor's probability of pursuing SaaW. One of the core challenges was data acquisition, as in cases where data sets are limited or lacking, experts are required to populate CPTs that grow exponentially with the number of associated parent-nodes. To mitigate this, Das' algorithm, which weights the relative parent-node's influence strength and combines it with CPCs, was applied to create a linearly growing set of probability distributions.
This operational model provides the foundation for the three case studies in the next chapter and contributes to answering the third research question: *What is the probability that a given actor is pursuing SaaW?*

6
Case Studies

Contents

The previous two chapters presented a conceptual and an operational model, which is used for three case studies to answer the last research question: *'What is the probability that a given actor is pursuing SaaW?'*. Case studies serve a dual purpose: they validate the method and they assess its utility, but they are not without

limitations [292]. They can range from completely fictional scenarios to mirrors of reality, with the latter being more useful but also difficult to achieve due to the lack of information – or the access to it. Utilising mixed methods, these case studies also serve to generate new data, which is still rare in this young field.

The BN discussed in the previous chapter is applied to three actors, representing terrorists, state powers, and generic attackers. Alongside five datasets, 30 interviews provide data to populate the network nodes with prior probability nodes and relative weightings of dependencies. Whilst the same BN structure was used for all three cases, the label, or conceptional meaning, of nodes changed, adjusting to the actor in question. It does not, for example, make sense to talk about 'national security' when referring to a non-state actor. 'Group security', however, does make sense. The dynamics are clearly different, but they are sufficiently related for the current iteration of the model. Similarly, certain nodes or connections will carry limited or no weight for one actor but be vital to another. Keeping the structure intact allows for easier comparison and analysis. Additional models or versions could however be tailored to very specific actors or types of SaaW and targets in questions, for which this current version could provide a starting template.

The remainder of this chapter is structured as follows: first, the three case studies are discussed, followed by overall qualitative results, as well as priors and weights. Next, each actor is iterated along the three pillars of *restraints*, *motivations*, and *capabilities*.

6.1 Case Studies

Originally, the scenarios were to include a time element, representing a specific actor in 2008, 2013 and 2018. However, it was discovered during the pilot study that the participants struggled to remember and differentiate these years, both from a technical and geo-political point of view. Furthermore, the combination of time commitment and repetitiveness of the task time was too inhibitive. All scenarios are therefore set in 2018. Ethical concern was also raised picking specific, real, actors, thus all scenarios are set in a 'mirror reality'. The interviewees were

given an overview and explanations of the network, followed by a series of questions pertaining to it, such as the various elements and connections. This was then followed by an iteration of the three scenarios.

6.1.1 Generic Actor

The main purpose of this scenario was to allow participants to become familiar with the model and elicitation process. During initial discussions and test runs, the level of freedom to be given to participants in the scenario design was also discussed. On the one hand, increased freedom would provide interesting data, such as what actors and capabilities the respondents have in mind, but it would also likely result in as many different scenarios as participants. It could, however, be used as an additional study or future work.

Here, the scenario for the *Generic Actor* was set as follows: the actor was to be a western state with average-to-high military power and living standards, such as France. As member of an alliance system, the actor is involved in peace-keeping missions but no other conflicts. The capabilities sought are conceptually akin to Stuxnet: the ability to degrade specific uranium enrichment operations by attacking industrial PLCs of another state actor, whilst limiting the potential for blow-back or collateral – emphasis here was placed to err on the side of caution.

6.1.2 State Actor

This scenario used the same setting of capabilities pursued as above, albeit the hypothetical state actor being based on the UK. In summary, this actor is a parliamentary democracy with a population of roughly 66million, an annual GDP growth of 1.5% and GDP/capita of around USD 42k. It is a member of an alliance system akin to NATO and an intelligence alliance similar to *Five Eyes*. There are some sophisticated cyber operation capabilities but the previous focus has been on defensive elements.

6.1.3 Terrorist Actor

This scenario diverges from the other two, by having the actor be a religiously driven group instead of a state actor, that by many is considered a terrorist grouping. It was used to explore to what extent the same model could be applied to non-state actors, despite the highly varying internal and external dynamics. The group in this scenario has taken over territories in several states and seeks to expand widely, claiming religious, political, and military authority, mimicking Daesh/ISIS. However, it is considered to be centrally led, without networked, or self-acting, structures. The capabilities sought belong to the family of encryption ransomware, for example propagating via infected e-mail attachments and/or the EternalBlue exploit, seeking payment in a cryptocurrency. Unlike NotPetya, there is no targeting of a specific industry sector or state-actor. Instead, the spread is to be as far ubiquitous as possible and contains an automatic transport mechanism marking it as a worm, similar to WannaCry.

Several participants were also willing to complete an additional, fourth scenario, this time the *Terrorist Actor* being the same as above, however, the capabilities pursued instead being based upon Stuxnet in-line with the two state-actor scenarios above.

6.2 Qualitative Results

These results centre on participants' comments regarding the model in general, its nodes and connections. Unsurprisingly, many interviewees had diverging opinions, and whilst the majority of suggestions and comments were only mentioned once, several others were repeated and are presented below.

Potential Collateral This node currently has two parents, accounting for *Human* and *Environmental* aspects. 14 participants suggested to add another parent to represent 'digital collateral', explicitly accounting for any damage done to computers, data and related technologies including aspects of CIA. Three proposed the addition of an 'economic' parent node, to reflect commercial/industrial fallout. Several asked

if these nodes included the idea of 'blow-back', whether it was synonymous with 'unintended consequences' and how concepts and definitions of 'harm' would fit. Accounting for economic and digital damage, or harm, makes sense in light of attacks such as NotPetya and WannaCry. Similarly, data breaches and data scandals, for example Under Armour's MyFitnessPal and Facebook-Cambridge Analytica in 2018, would also support adding these additional sub-nodes. On the other hand, this could be counter-argued with highly targeted cases, such as Stuxnet, that despite its extensive spread focussed on not causing collateral damage.

Adding additional nodes would shift *Potential Collateral* closer to Agrafiotis et al.'s CHM [4], with its six elements across four levels discussed previously in Chapter 2. However, unlike their model, psychological/emotional and cultural elements were not raised, and levels remained centred on property/infrastructure and national aspects. The idea of blow-back is also very interesting, particularly given increased publicly known usage of cyber-capabilities, for example by the US [293]. Here, blow-back could be understood in two different forms, either more generally, such as a cyber-attack leading to a response in turn, or more specifically, the same capability backfiring, damaging the initiating actor. In this model, the former aspect should however be placed under *Fear of Retaliation*, possible as its own node, whilst the latter would be added under *Potential Collateral*. This would require detailed knowledge about the actor's own susceptibility to the proposed capability, as well as a measure to define what level is considered acceptable in what scenario. For implementing this sub-node, one solution would be to mirror *Potential Collateral*: one for the target, one for the actor themselves. Yet, this would also increase the complexity greatly, particularly the suggestions above to add more sub-nodes is followed. An alternative would be to have a single sub-node, which encompasses the potential for blow-back more generally.

Domestic Safeguards Three participants suggested that *Ideology* should be a parent to this node, whilst another four suggested a similar parent labelled

'morality'. Future work could also explore various domestic safeguard regimes for specific weapon systems to improve this model.

Fear of Retaliation One of its parents is *Opponents Military Power*, which 19 participants believed should be split into cyber and non-cyber components. Furthermore, two remarked that capability should be split into offensive and defensive. Three brought up 'economic sanctions', which currently have not been addressed. Eleven suggested that *Fear of Retaliation* should be two disconnected or different nodes, one each for *Restraints* and *Motivations* as the current structure caused confusion.
Events in 2019, particularly such as the US launching cyber-attacks on Iranian rockets and missiles [14], [162] or other alleged retaliatory cyber-attacks on Iran [7], would further support splitting into cyber and non-cyber components, as would the idea of blow-back discussed under potential collateral.

Attribution Linked to retaliation, nearly every participant asked about 'attribution', with two specifically stating that it is a burden of proof. The question of feasibility thereof remains. It was suggested that an 'attribution' node should be parent to both *Restraints* and *Motivations*, or to *Fear of Retaliation*. Similarly, connections to *International Prestige* could be made, however, they swing both ways: diminish it for being caught red-handed, or bolster because of it.

Cost as a Restraint Three proposed the inclusion of an 'opportunity cost' as a restraint, whilst six suggested to connect the *Economic* node to *Restraints* as a parent. This would need to be investigated further, including connections to other concepts mentioned, such as the idea of blow-back.

Security This node posed a slight challenge as a security motivation often stems from insecurity, and the two data sets (MID and CoW) for *Global* and *Regional Security* measure this in the form of conflict. Simplified, this means that 'low' values

in those two nodes are more likely compatible with a 'high' security motivation, leading to several suggestions of re-labelling.

Education & Skills The topic of a capable workforce was often mentioned, however only five suggested an explicit 'skills' node to contribute to *Capabilities*. Their suggestions for the connections of this new node diverged, from being a parent to *Research* and *Development* to being a direct *Capabilities* parent.

Development Seven believed that the *Development* node would be sufficient by itself, essentially merging *Test Bed*, *Payload* and *Delivery*. Others however believed they should be connected more and expanded upon. The respondents with a more technical background in particular found these nodes lacking, wanting more clarity.

6.3 Priors and Weights

The results and comparisons of the three scenarios can be seen in Figures 6.1 and 6.2). Whilst the state actor has the most balanced weighting, *Potential Collateral* and *Fear of Retaliation* carry slightly more importance. The generic and terrorist scenario similarly show repercussions as the strongest restraint, however, in the latter case, it outweighs the other three factors combined. *International Agreements* become, unsurprisingly, almost negligible in that scenario.

The *Human Collateral* distribution for the state actor scenario was surprising, expecting a larger value for 'high'. At the same time, it is equally surprising to see the terrorist group actor having such a large value for 'high'. This could be due to easy misinterpretation of the question, by struggling to reconcile that the terrorist group may want (or is motivated by) human collateral, however, in this instance, the question is to what extent it is a restraining factor.

It should be noted that in the *Motivations* context, the *Fear of Retaliation* is reversed: the idea is that a lack of fear can act as an encouraging factor. In terms of *Ideology*, the relative weighting is reversed for the state and terrorist actors,

Human Coll.	G	S	T
H	0.381	0.544	0.236
M	0.234	0.200	0.153
L	0.242	0.153	0.493
U	0.143	0.103	0.118

Enviro. Coll.	G	S	T
H	0.207	0.314	0.159
M	0.341	0.371	0.167
L	0.324	0.207	0.536
U	0.127	0.107	0.139

Opp. Mil. Power	G	S	T
H	0.350	0.419	0.267
M	0.206	0.257	0.216
L	0.246	0.144	0.370
U	0.198	0.180	0.147

Opp. Alliances	G	S	T
H	0.356	0.341	0.187
M	0.239	0.279	0.269
L	0.213	0.204	0.407
U	0.193	0.176	0.137

Historic	G	S	T
H	0.267	0.325	0.302
M	0.256	0.229	0.339
L	0.333	0.316	0.206
U	0.144	0.130	0.153

Religious	G	S	T
H	0.269	0.229	0.663
M	0.186	0.253	0.166
L	0.413	0.386	0.084
U	0.133	0.133	0.087

Global Sec	G	S	T
H	0.314	0.331	0.217
M	0.276	0.294	0.183
L	0.276	0.254	0.409
U	0.134	0.120	0.191

Regional Sec	G	S	T
H	0.361	0.435	0.363
M	0.299	0.282	0.217
L	0.203	0.163	0.296
U	0.137	0.120	0.124

National Sec	G	S	T
H	0.495	0.586	0.383
M	0.196	0.137	0.211
L	0.134	0.117	0.223
U	0.174	0.160	0.183

Strategic	G	S	T
H	0.309	0.390	0.287
M	0.269	0.247	0.287
L	0.231	0.190	0.230
U	0.191	0.173	0.196

Research	G	S	T
H	0.391	0.430	0.280
M	0.251	0.216	0.210
L	0.183	0.186	0.293
U	0.174	0.169	0.217

Figure 6.1: Node priors across the *Generic* (G), the *State* (S) and *Terrorist* (T) scenarios.

while it is almost balanced for the generic scenario. *Historic* factors will likely carry a higher weight for any secular entity.

Within *Security*, the 'national' or 'group' aspect is the most strongly weighted contributing factor for all scenarios. However, *International Prestige* has an almost equal standing for the terrorist group, which seems counter-intuitive, unless it is taken to mean notoriety in this context.

6.4 Generic Actor

An initiated BN for the generic scenario (and evidence selected) can be seen in Figure 6.3, and enlarged in three sections in Figures 6.4-6.6. The leaf nodes were set to the highest prior value. Overall, the states 'high', 'medium', and 'low' were almost equally split for SaaW proliferation, which given the unspecific scenario nature are not surprising.

SaaW	G	S	T
Restraints	0.262	0.289	0.158
Motivations	0.334	0.335	0.465
Capabilities	0.404	0.376	0.377

Restraints	G	S	T
Pot. Collateral	0.258	0.273	0.150
Dom. Safeg.	0.215	0.244	0.112
Int. Agr.	0.213	0.250	0.073
Retaliation	0.314	0.233	0.665

Motivations	G	S	T
Retaliation	0.162	0.200	0.134
For. Policy	0.230	0.254	0.122
Dom. Policy	0.216	0.207	0.099
Ideological	0.157	0.142	0.405
Economic	0.234	0.197	0.239

Capabilities	G	S	T
Economic	0.212	0.192	0.233
Strategic	0.140	0.147	0.125
Tactical	0.147	0.150	0.154
Research	0.168	0.170	0.128
Development	0.184	0.178	0.184
Intel/Recce	0.148	0.162	0.176

Pot. Collateral	G	S	T
Human	0.683	0.664	0.677
Enviro.	0.317	0.336	0.323

Fear of Retaliation	G	S	T
Military	0.607	0.607	0.589
Alliances	0.393	0.393	0.411

Foreign Policy	G	S	T
Pub. Percep.	0.185	0.211	0.318
Security	0.388	0.371	0.276
Industry Int.	0.254	0.239	0.114
Historic	0.174	0.179	0.292

Development	G	S	T
Test Bed	0.404	0.378	0.255
Payload	0.300	0.311	0.432
Delivery	0.295	0.311	0.313

Domestic Policy	G	S	T
Polity IV	0.161	0.186	0.052
Pub. Percep.	0.174	0.168	0.206
Security	0.246	0.271	0.152
Industry Int.	0.184	0.161	0.072
Historic	0.116	0.126	0.181
Religious	0.118	0.088	0.337

Mil. Spending	G	S	T
Economic	0.572	0.535	0.572
Strategic	0.428	0.465	0.428

Test Bed	G	S	T
Economic	0.271	0.250	0.312
Strategic	0.176	0.191	0.134
Research	0.292	0.282	0.241
Intel/Recce	0.262	0.277	0.313

Ideology	G	S	T
Historic	0.514	0.603	0.355
Religious	0.486	0.397	0.645

Payload	G	S	T
Economic	0.220	0.236	0.289
Strategic	0.203	0.201	0.158
Research	0.300	0.263	0.230
Intel/Recce	0.277	0.300	0.323

Economic	G	S	T
Security	0.344	0.382	0.448
Industry Int.	0.356	0.304	0.205
GDP	0.300	0.315	0.347

Delivery	G	S	T
Strategic	0.213	0.204	0.238
Tactical	0.218	0.221	0.198
Research	0.244	0.251	0.229
Intel/Recce	0.326	0.324	0.335

Security	G	S	T
Global Sec.	0.229	0.226	0.165
Reg. Sec.	0.244	0.249	0.246
Nat. Sec.	0.334	0.330	0.311
Int. Prestige	0.193	0.196	0.277

Figure 6.2: Relative node weightings across the *Generic* (G), the *State* (S) and *Terrorist* (T) actor scenarios.

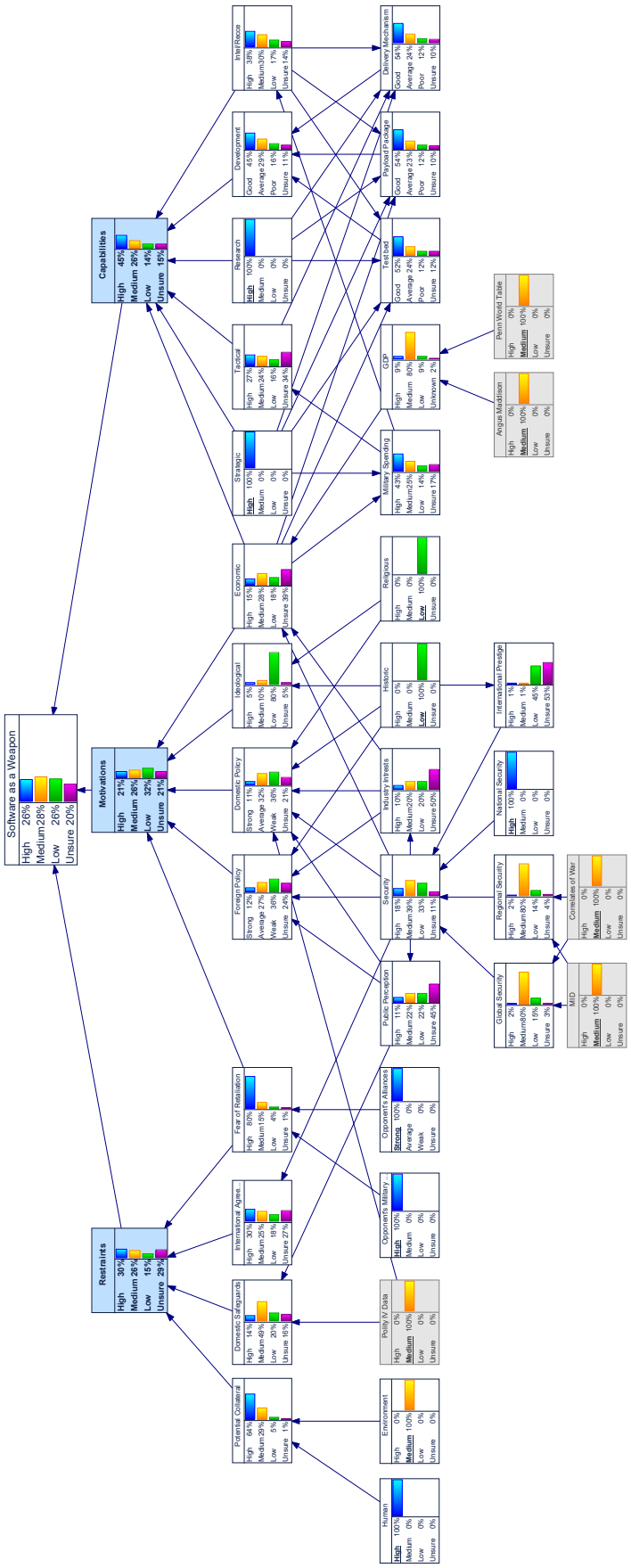


Figure 6.3: Results: Initiated BN for the Generic Scenario

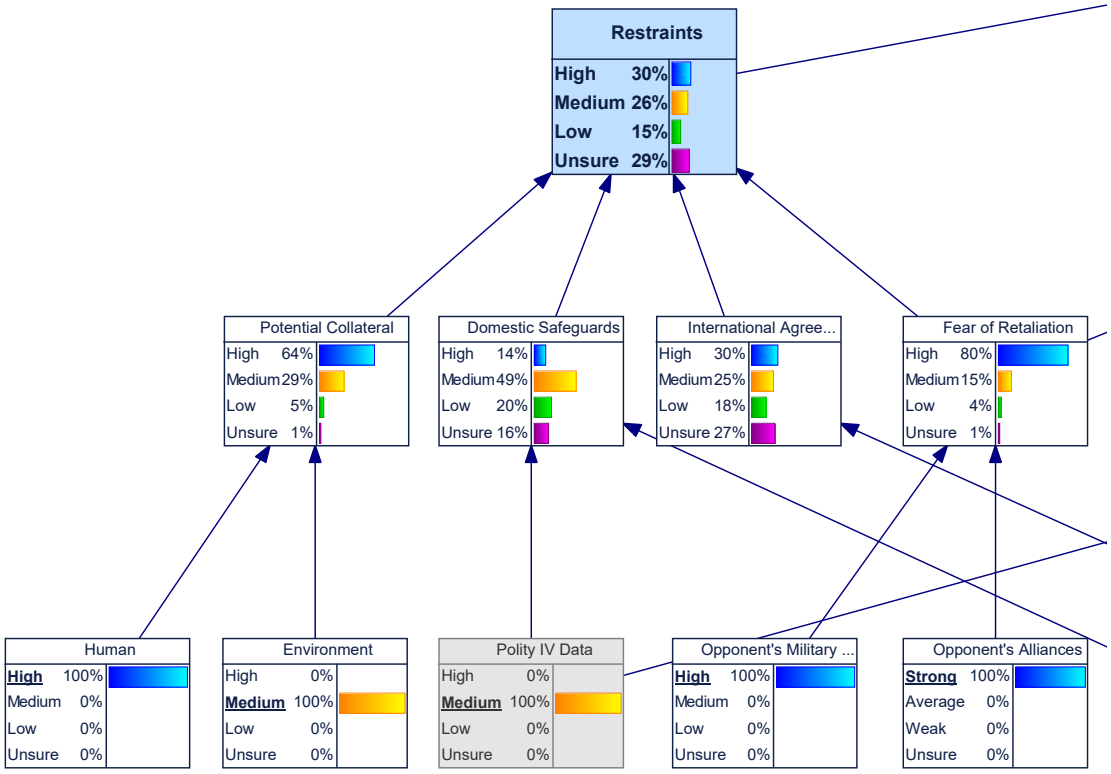


Figure 6.4: Enlarged view of Generic Actor BN centring on the *Restraints*

6.4.1 Restraints

Overall, this node results in a 'high' state, with 30%, followed by another 26% 'medium', 'unsure' is almost equal to 'high' with 29%. However, given the percentages of 'high' states of some the parents, it is somewhat surprising that this is not equally reflected. This difference is due to the weights given, combined with the ambiguous *International Agreements* parent and 'medium' *Domestic Safeguards*.

Potential Collateral arrives at 64% 'high', driven by the 'high' state of its *Human* parent node paired with 'medium' *Environmental*, with a weighting of 0.68 to 0.32, respectively. The relative importance of the two sub-nodes changes minimally across scenarios, however the priors vary much more greatly. In this case, the *Polity IV Data* parent node was set to 100% 'medium', whilst *Public Perception* was predominantly 'unsure', resulting in 49% 'medium' and an almost even split across the other states for *Domestic Safeguards*.

International Agreements has *Security* as a parent, combined with levels of

uncertainty, which likely stem from the topic at hand combined with it being a generic scenario with less tangibility. This is reflected by an overall ambiguous split. *Fear of Retaliation* is 'high' given both parent nodes (*Opponent's Military Power* and *Alliances*) are also 'high', with a little room for uncertainty.

6.4.2 Motivations

In the generic scenario, the overall level is 'low' at just over 30%, driven by similar values of domestic and foreign policy, as well as low historic and religious impetus. For comparison, the terrorist actor scenario has almost reversed values.

Foreign Policy provides little motivation, with 36% 'weak' ('low'), followed by 27% 'average' or 'medium'. It is strongly driven by similar security concerns combined with uncertainty stemming from *Industry Interests* and *Public Perception*. These values are not surprising given the scenario setting. Similarly, *Domestic Policy* is 'weak' ('low'), with additional 'low' parental influences from *Historic* and *Religious* factors. Furthermore, *Ideological* elements are overwhelmingly 'low' due to both parents' settings, which also suggests that the respondents had actors in mind where these factors have very little to no influence. The *Economic* node appears mainly 'unsure' at 39%, followed by 'medium' (28%) levels. High levels of ambiguity surrounding *Industry Interests* (50% 'unsure') are evident, combined with the 'medium' *GDP*.

6.4.3 Capabilities

Overall, the *Capabilities* are envisaged as 'high' (45%), with strong *Strategic* and *Research* contributions, both of which have no parents, as well as *Development*. Without parents, the *Strategic & Research* nodes have been set to their highest prior value. The *Tactical* node appears split, with 34% 'unsure', followed by 27% high. The *Development* node on the other hand is predominately 'high' with 45% gained from all three technical nodes. This is a little bit surprising, as a higher level of uncertainty was expected given a non-specific actor. Similarly, the 38% 'high', followed by 30% 'medium' of the *Intel/Recce* node is unexpected, as to some extent,

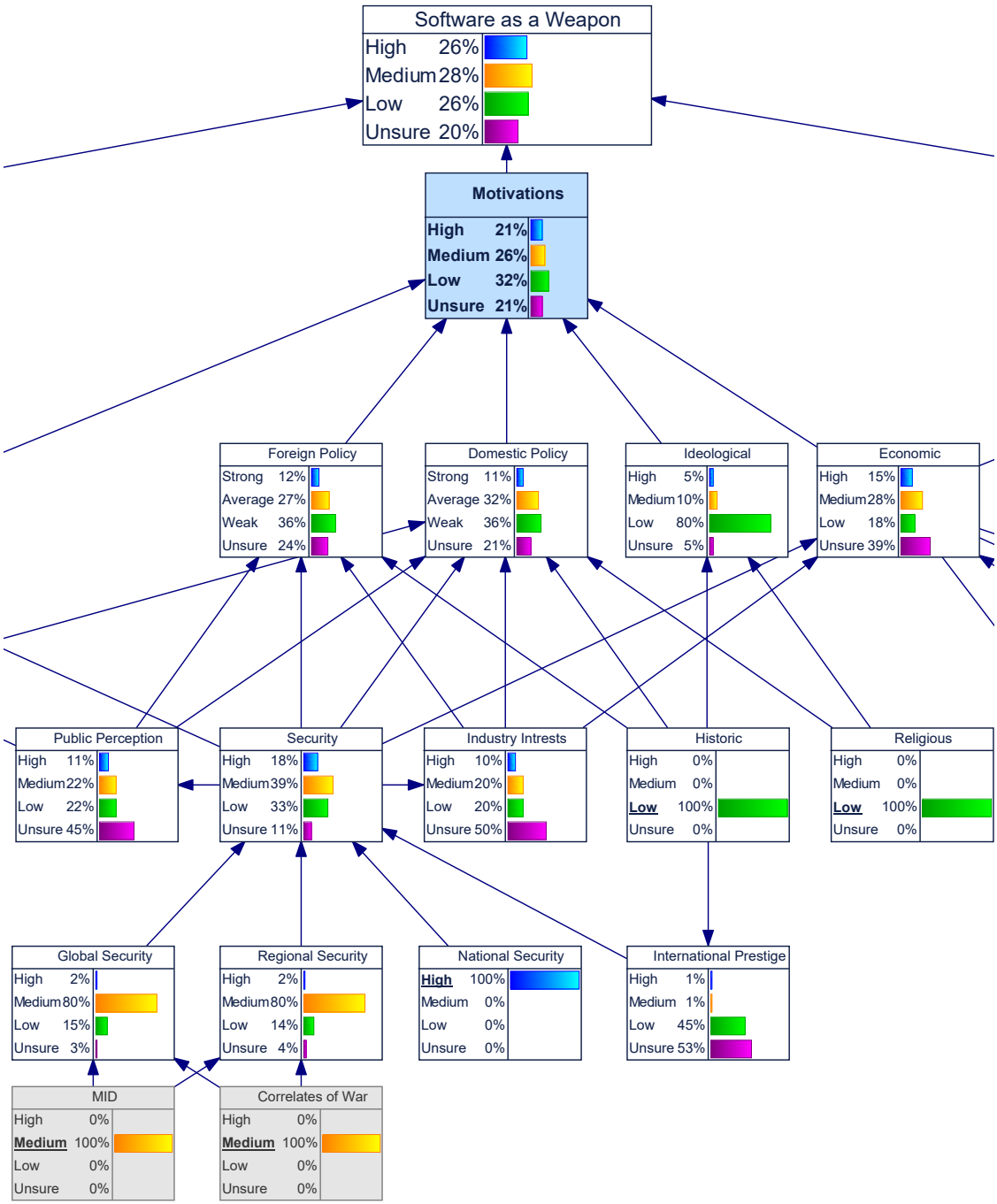


Figure 6.5: Enlarged view of Generic Actor BN centring on the *Motivations*

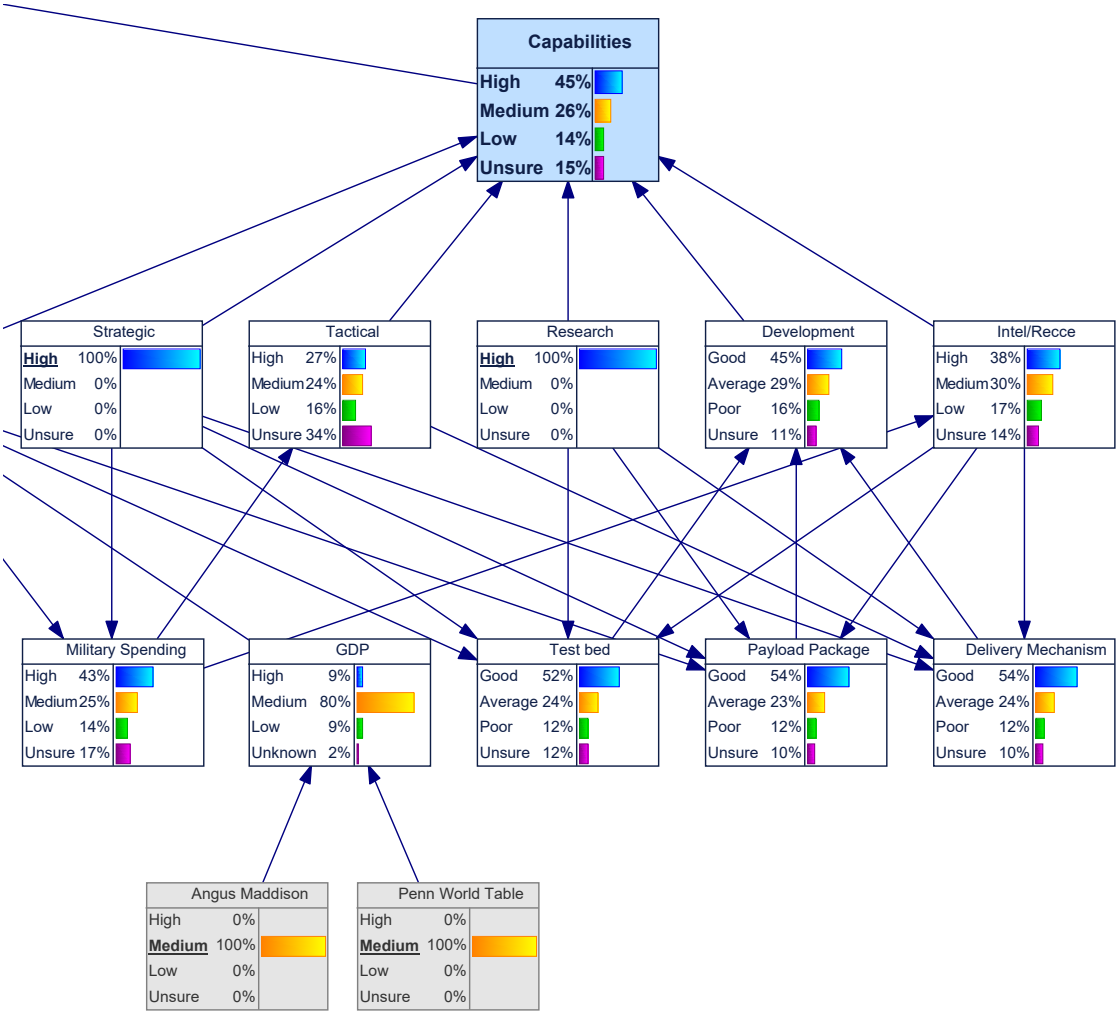


Figure 6.6: Enlarged view of Generic Actor BN centring on the *Capabilities*

it is one of the most obscure nodes by default. It appears that the participants were either very sure, or that mistakes were made in the weighting or the structure.

6.5 State Actor

This actor has three states competing, with 'high', 'medium' and 'low' being at 28%, 26% and 31%, respectively, indicating the potential for internal struggle. On the one hand, *Restraints* are 'high', yet so are *Capabilities*, whilst *Motivations* are almost 40% 'low', but also just over 40% split between 'medium' and 'high'. This means

that while for now the pursuit is considered unlikely, it could easily shift, particularly if the motivation was to increase, whether via internal factors or geopolitical events. This change would then require additional disincentives, such as an increase of currently existing restraining elements or the addition of new ones. Alternatively, it could also be explored how the actor's capabilities could be reduced to achieve this. The sensitivity for each state of this node is shown in Figures 6.7 - 6.8, comparing the relative importance of factors. The sensitive variable is shown as an uncertain value, whilst all others are kept stable.

6.5.1 Restraints

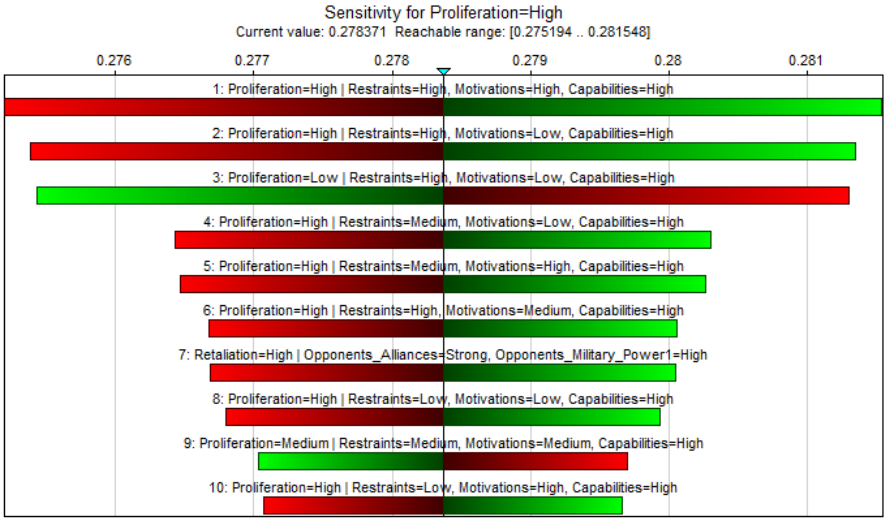
In this scenario, restraints are 'high' state at 47%, clearly indicating that this actor has very low incentives. Of the parent nodes, uncertainty is mainly driven by *Domestic Safeguards* and *International Agreements*.

The fear of doing accidental harm depicted as *Potential Collateral* is a strong restraining factor (67%) in this scenario, with 'high' and 'medium' states from its parent nodes *Human* and *Environmental*, respectively. The *Domestic Safeguards* node is also 'high' at 41%, with the other states split almost equally. In this scenario, the *Polity IV Data* parent node was set to 100% 'high', whilst *Public Perception* was predominantly uncertain (48% 'unsure'). *International Agreements* has *Security* as a parent, which, combined with levels of uncertainty, results in a 'high' state. In this scenario, the opponent is considered to be very strong in terms of *Military Power* and *Alliances*), resulting in a 'high' *Fear of Retaliation* combined with some uncertainty.

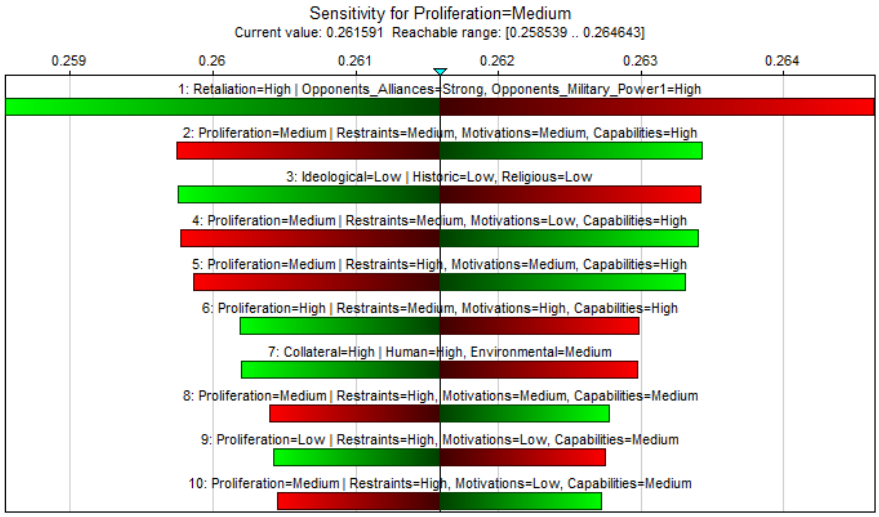
6.5.2 Motivations

In the generic scenario, the overall level is 'low' at 39%, driven by a relative stability domestically and abroad, as well as low historic and religious impetus. It is only marginally different from the generic actor scenario, with a higher tendency towards the 'low' state, despite larger variance in parent nodes.

Foreign Policy provides little motivation, with 52% 'weak' ('low'), followed by 25% 'uncertain' – much lower than in the generic actor scenario. It is similarly driven by



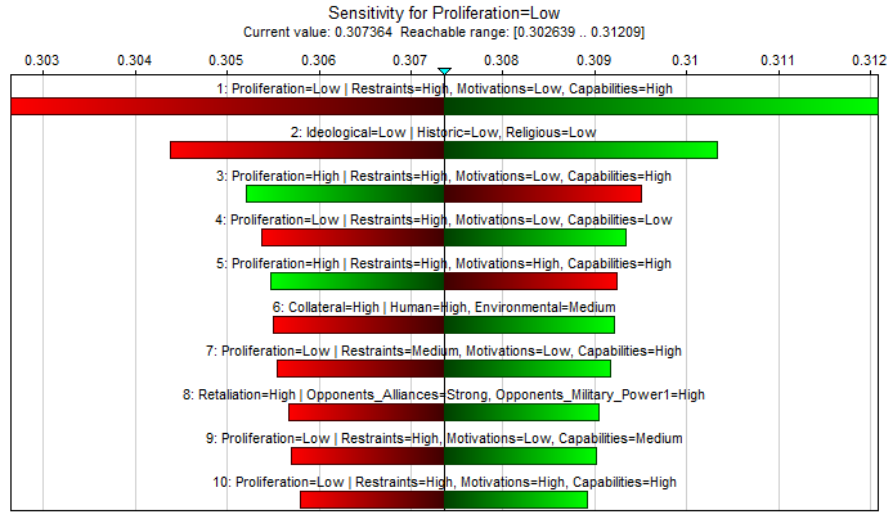
(a)



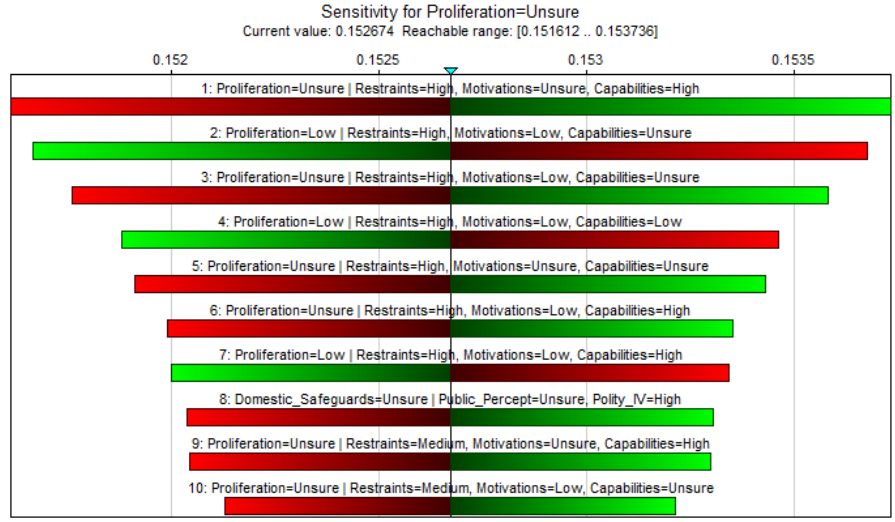
(b)

Figure 6.7: Tornado Plot Comparison: State Actor's SaaW Node States (a) High and (b) Medium

'low' security concern combined with uncertainty stemming from *Industry Interests* and *Public Perception* (53% and 48%, respectively). The state distribution for the two latter nodes is surprising given that the scenario is no longer ambiguous, but it could be explained by either the ambivalent nature of the factors or the respondents' unfamiliarity with them. Similarly to its foreign counterpart, *Domestic Policy* is in its 'low' state at 48%, stemming from the *Historic* and *Religious* leaf nodes set to 'low', combined with a 'low' *Security* impetus. The *Economic* node shows a



(a)



(b)

Figure 6.8: Tornado Plot Comparison: State Actor's SaaW Node States (a) Low and (b) Unsure

shift towards higher levels in comparison to the generic actor scenario, with 26% versus 15% 'high' but all four states are nearly evenly split.

6.5.3 Capabilities

This actor is considered to be considerably strong in this area, with overall *Capabilities* at 51% 'high', followed by 19% 'medium'. This is not surprising given the leaf nodes' settings (all at 'high'), with some ambivalence in *Tactical* and *Intel/Recce*

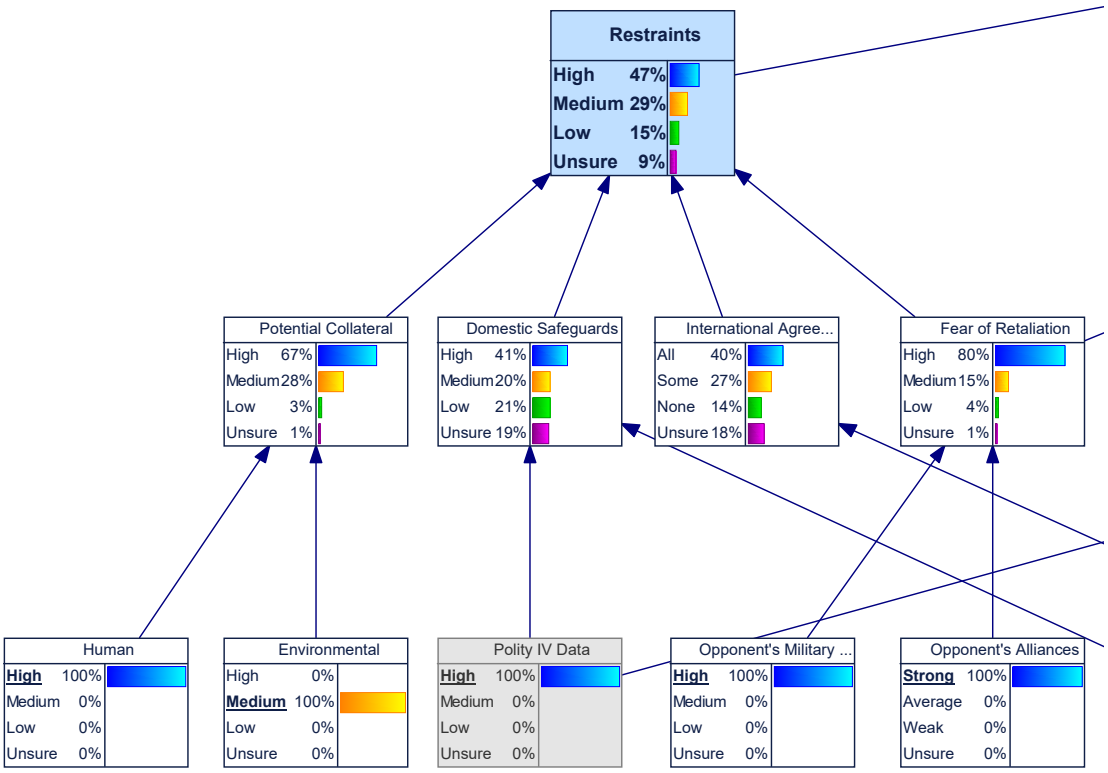


Figure 6.9: Enlarged view of State Actor BN centring on the *Restraints*

aspects. The former is not unexpected given the CPT results, however, the latter is. Given the scenario, it is surprising to see *Intel/Recce* at 31% 'unsure', with the other states split almost equally. However, this can be possibly explained by only one parent node contributing. The *Development* node is predominately 'high' with 47% gained from all three technical nodes, all of which are between 50% and 60% 'high'.

6.6 Terrorist Actor

This actor has overall 'medium' to 'low' probability of proliferating SaaW, despite the 'low' *Restraints*, which was somewhat surprising given the scenario. This is driven by the lack of *Capabilities* and limited *Motivations*, which means that based on these results no additional measures would need to be taken to discourage the actor. Unlike the *State Actor*, however, the *Restraints* are'low', thus any shift in *Motivations* or *Capabilities* will be harder to counterbalance.

The sensitivity for each state of this node is shown in Figures 6.12 - 6.13, comparing

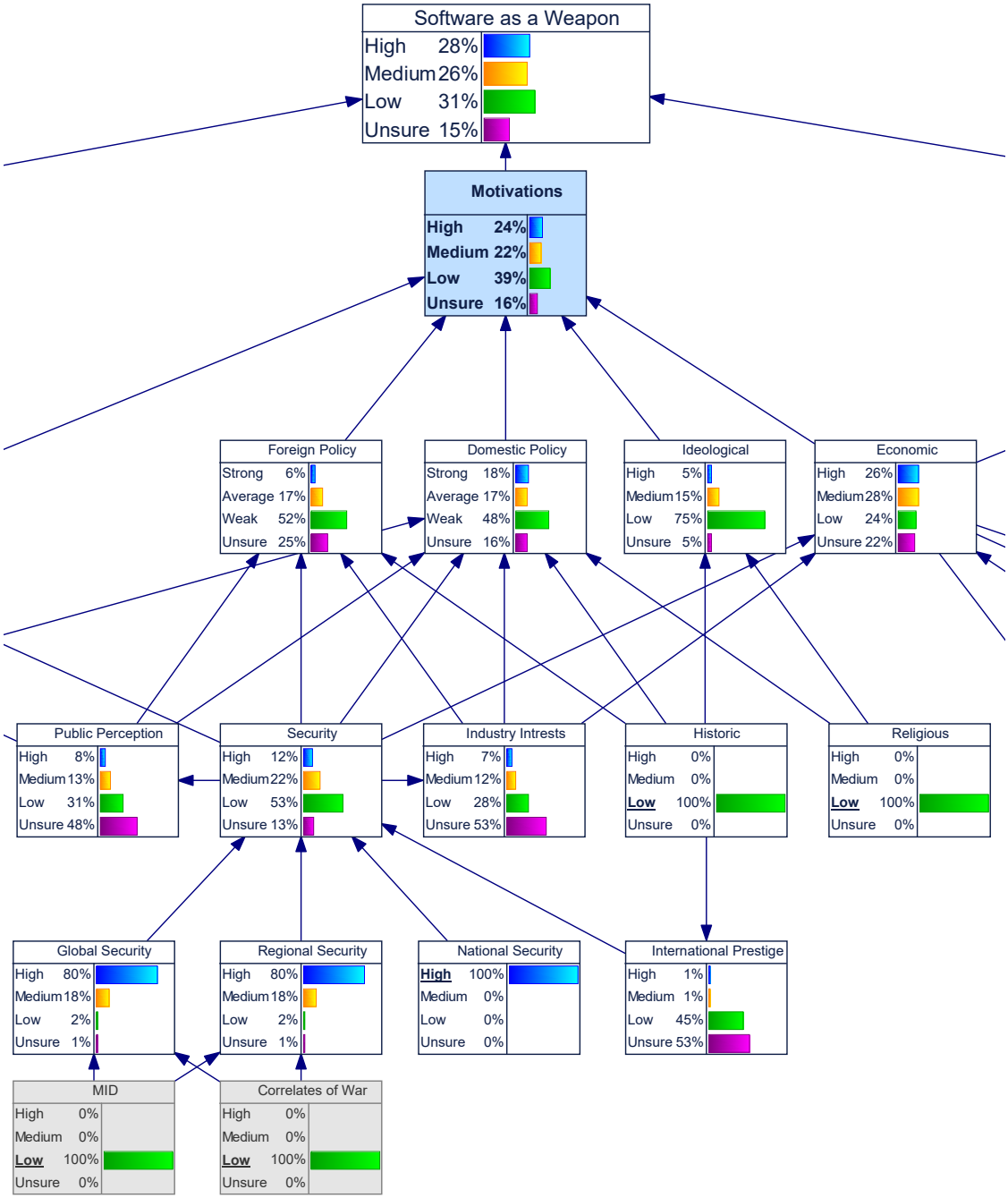


Figure 6.10: Enlarged view of State Actor BN centring on the *Motivations*

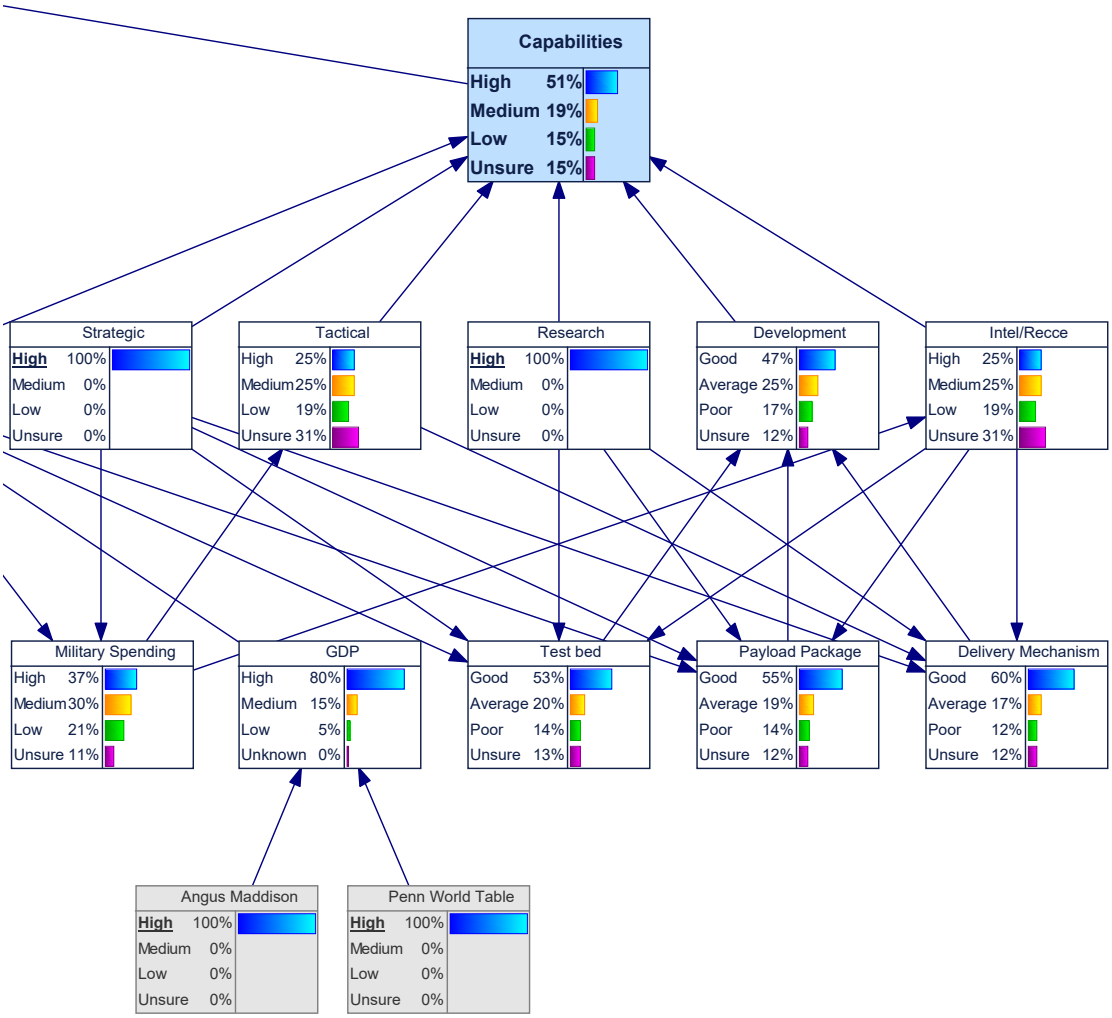


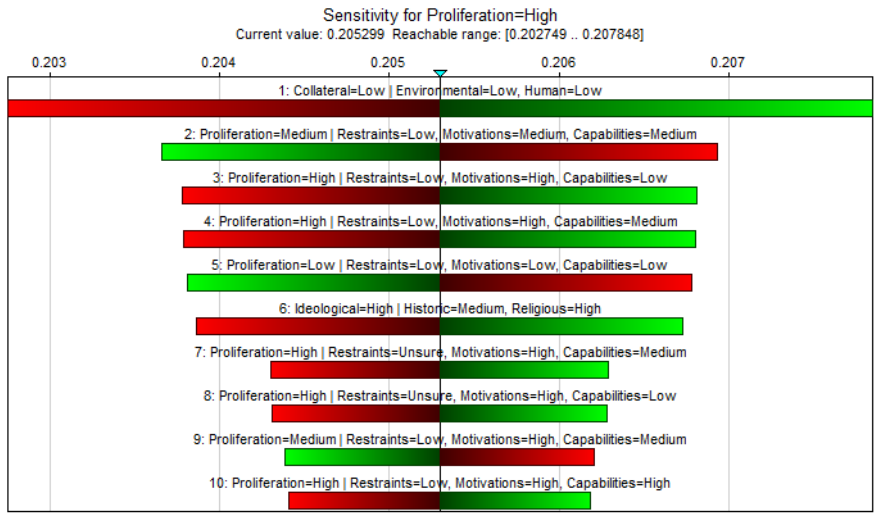
Figure 6.11: Enlarged view of State Actor BN centring on the *Capabilities*

the relative importance of factors. The sensitive variable is shown as an uncertain value, whilst all others are kept stable.

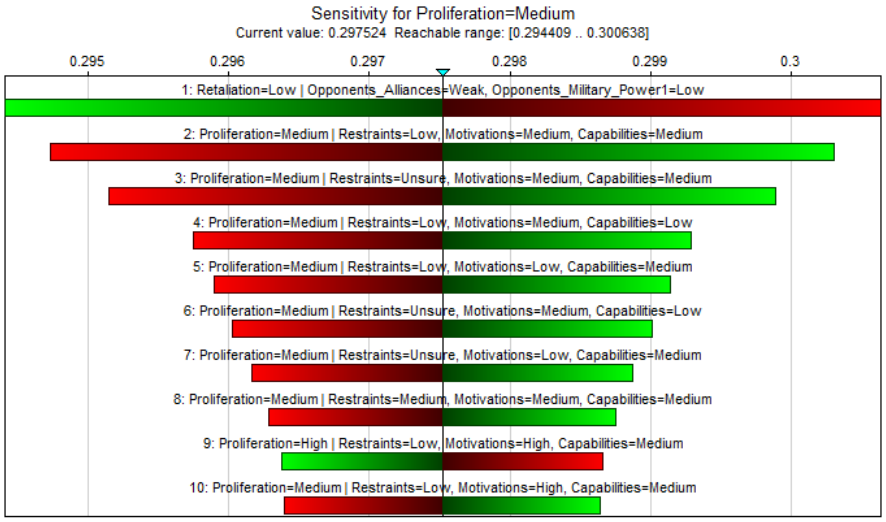
6.6.1 Restraints

Overall, this node is 'low' with 39% but it also has a propensity for 'unsure' at 34%. The former is driven by the disregard for collateral and retaliation, combined with a lack of *Domestic Safeguards*. However, there seems to be more reluctance of the experts when completing the CPTs, favouring more uncertainty. Based on discussions with participants, this is due to the unpredictable nature of this actor as well as the difficulty of taking the actor's perspective.

Unsurprisingly, the node *Potential Collateral* is overwhelmingly 'low', which



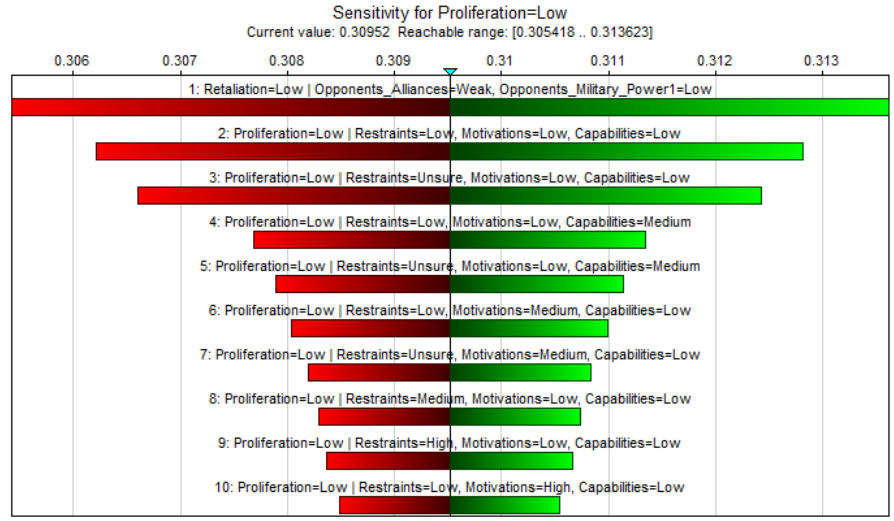
(a)



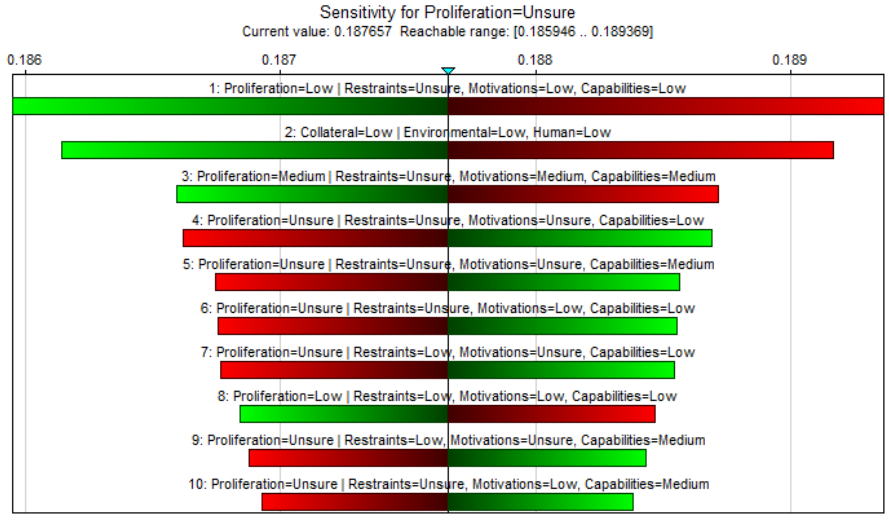
(b)

Figure 6.12: Tornado Plot Comparison: Terror Actor's SaaW Node States (a) High and (b) Medium

corresponds to the limited historic data available pertaining to non-cyber events. Similarly, the node on *Domestic Safeguards* is also 'low', driven by the *Polity IV Data*. There was argument to place this node into 'unsure' given the actor's status. This would result in a switching of 'low' and 'unsure' states in this node: whilst currently the four states are split into 13%, 21%, 52% and 14%, respectively, they would become 11%, 18%, 22% and 48% with that adjustment. In turn, this would affect the *Restraints* child node marginally, resulting in 'low' and 'unsure' both at 37%.



(a)



(b)

Figure 6.13: Tornado Plot Comparison: Terror Actor's SaaW Node States (a) Low and (b) Unsure

Fear of Retaliation has two parent nodes (*Opponent's Military Power* and *Alliances*), both of which are 'low' restraining factors, resulting in this node also being overwhelmingly 'low' with 80%. This is followed by 'medium' 14%, with the remainder split across the other two states.

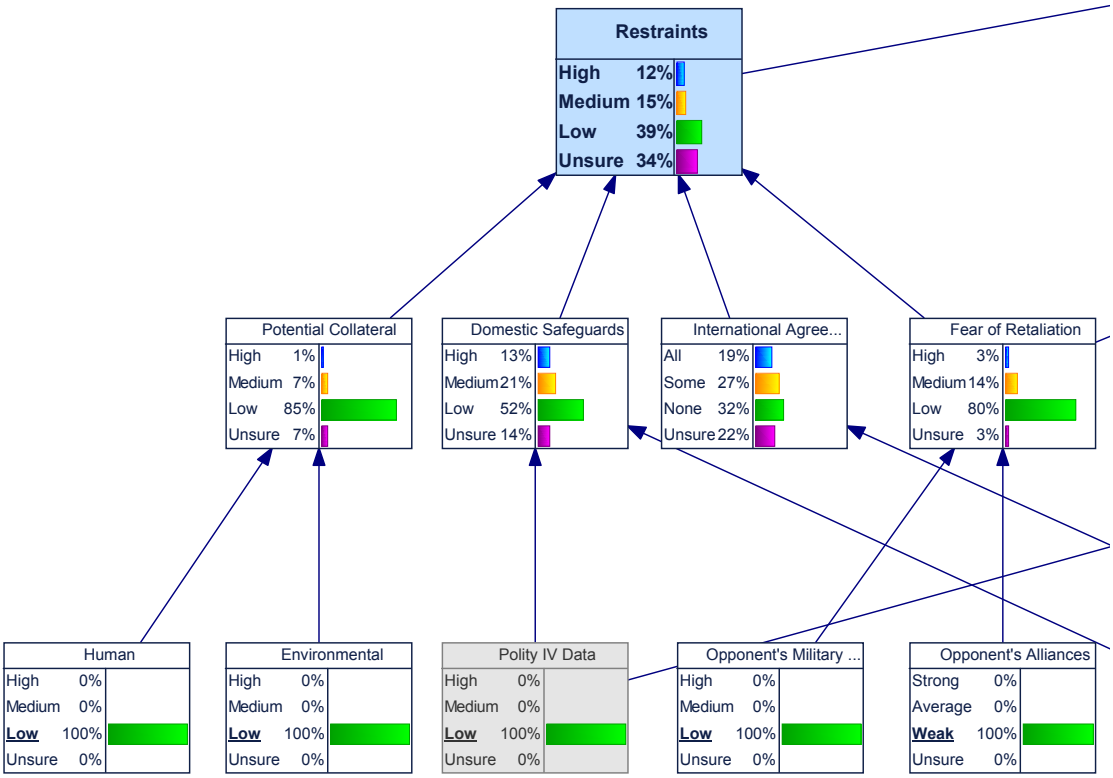


Figure 6.14: Enlarged view of Terrorist Actor BN centring on the *Restraints*

6.6.2 Motivations

For the terrorist scenario, the overall motivations are split, peaking at 'medium' and 'low' almost equal and accounting for 63%, whilst 'high' is 21%. The greater *Security* and *Ideological* impetus appear to be balanced by the overall 'medium' *Foreign* and *Domestic Policies*, as well as *Economic* factors.

Although a non-state actor may not have a political structure akin to a state actor, they do often have politics of dealing with group-external actors. However, it should be noted that they may not necessarily be cohesive or stable. In this scenario, *Foreign Policy* provides average motivation, mixing 'high' *Security* concerns with great uncertainty in *Public Perception* and *Industry Interests* (47% and 51%, respectively). Further to *Foreign Policy*, the concept of *Domestic Policy* needs to be adjusted for non-state actors: many will not be in control of a territory or have state-like entities. Nonetheless, they will have group internal policies and plans. That is particularly true in this scenario, which is based on Daesh, who see themselves as a state and even

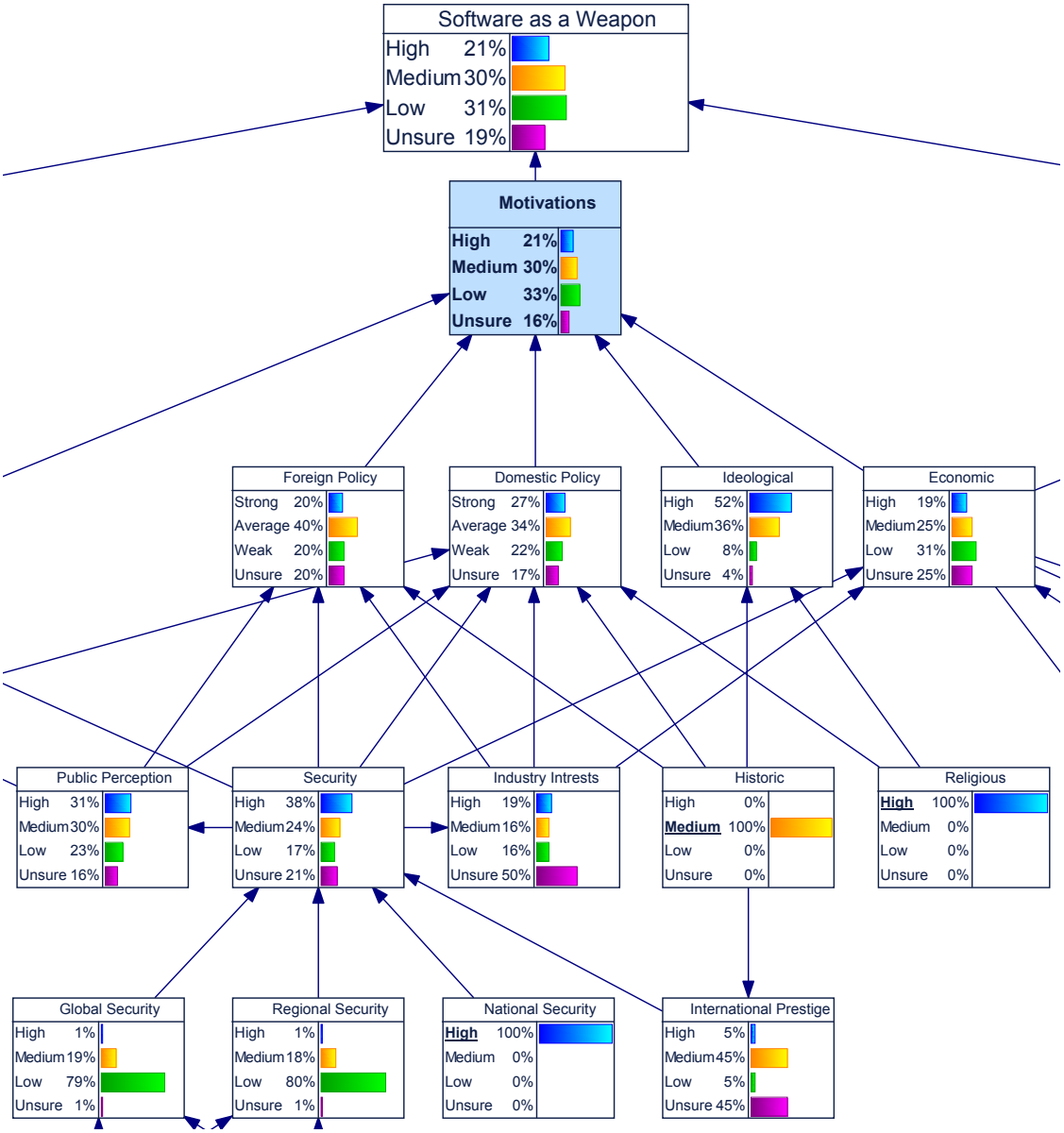


Figure 6.15: Enlarged view of Terrorist Actor BN centring on the *Motivations*

control some territory. Here, *Domestic Policy* is averagely motivating ('medium'). *Ideological* impetus on the other hand is seen as 'high' (52%), followed by 'medium' (36%) due to both parents' settings, with the *Religious* one being more prominent. The *Economic* element is overall 'low' at 31%, with 25% for 'medium' and 'unsure', which stems predominantly from the *GDP* measure that had both data sets leaf nodes set to 'low'. If these nodes were set to 'unsure', it would result in *Economic* also peaking in that state, as well as 1% more uncertainty in *Motivations* and 2% in *Capabilities* overall. This result highlights a potential problem, particularly in

presentation, as a lack of economic capability could be a major motivating factor.

6.6.3 Capabilities

The scenario considered a predominantly disruptive type of SaaW, with overall *Capabilities* resulting in 'low' at 38% and 35% 'medium'. Using a destructive version for comparison, 'low' peaked with 57%, followed by 16% for 'medium' and 'unsure'.

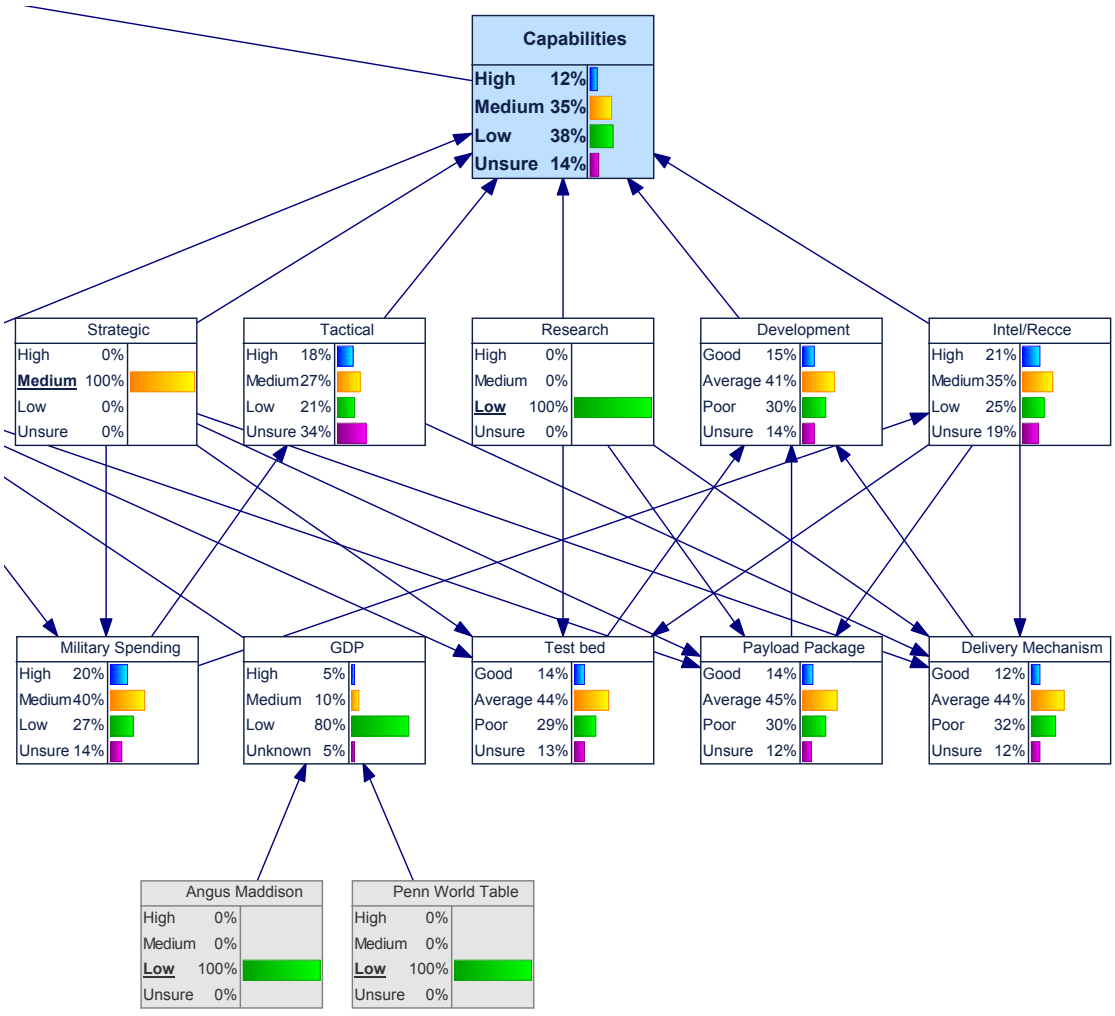


Figure 6.16: Enlarged view of Terrorist BN centring on the *Capabilities*

Given the limited knowledge available regarding the *Tactical* node, 34% 'unsure' is lower than expected, and it is followed by 27% 'medium'. The results for *Development* echo all three technical parent nodes leading to a predominately 'average' ('medium') with 41%. As the scenario calls for a SaaW that is not greatly targeted, it is surprising to see that *Intel/Recce* gained influence on *Capabilities*

overall in comparison to the State and Generic scenarios. Nonetheless, it arrives at a 'medium' state with 35%, followed by 25% 'low'.

6.7 Conclusion

This chapter presented a BN of factors contributing to the proliferation of SaaW across three actor scenarios, a generic, a state and a terrorist actor. Alongside five datasets, 30 interviews were conducted to populate the nodes with prior probability nodes and relative weightings of dependencies shown. The BN implemented Das' approach of weighting relative parent-node's influence strength, combined with compatible parental distributions, to create a linearly growing set of probability distributions as discussed in the previous chapter. The model was initiated for three scenarios: a *Generic Actor*, a *State Actor*, and a *Terrorist Actor*. This generated new data and provided qualitative as well as quantitative results, answering the third research question.

Several suggestions raised by the participants need to be taken into account regarding the network structure. This includes additional nodes within *Restraints*, for example to explicitly denote 'digital' or 'economic collateral', as well as opponent's cyber capabilities. Recent cyber attacks, such as NotPetya and WannaCry, as well as cases of data breaches and scandals would support the addition of such sub-nodes. This would also allow for a better differentiation, separating highly targeted cases such as Stuxnet, which despite its extensive spread sought to limit causing damage to unrelated systems. Furthermore, these additional sub-nodes would better connect to other work on harm done, such as the CHM [4].

Blow-back was another interesting aspect that was raised, which could be understood in two different ways. The first would be in a general form, for example a cyber attack leading to a response. The other one would refer to the same capability damaging the initiating actor. Keeping the current model as it is, the former could be placed under *Fear of Retaliation*, possible as its own *Cyber Retaliation* sub-node. The latter could be added under *Potential Collateral*, for example as *Cyber Blow-back*.

Furthermore, questions of economic cost and attribution were raised, as were proposals to re-label or re-define *Security* related nodes within *Motivations*, which future work should explore. The *Development* and its sub-nodes were found lacking particularly by those with technical background, seeking more details, whilst others countered that there was already too much detail and that one node would suffice. Based on this, future work should explore the possibility of different levels of detail within the model. For example, those with highly technical knowledge would be given a detailed breakdown, whilst those without would only see a top-level node. This might be reversed or further tailored for other parts of the model, for example *Security* elements.

This chapter also introduced quantitative results, with the scenario-based state actor having an ambivalent proliferation probability, with 'high', 'medium' and 'low' states being very close to each other. On the one hand, *Restraints* are 'high', yet so are *Capabilities*, whilst *Motivations* are almost 40% 'low', but also just over 40% split between 'medium' and 'high'. This means that for the current scenario, the efforts to curtail the pursuit are sufficient for the time being. However, the various factors that affect motivation should be closely monitored, as a shift could upset the balance. In that case, new approaches and concepts to prevent pursuit will need to be explored, for example in the form of stronger discouragements or disincentives. The scenario-based terrorist actor on the other hand has an overall 'medium' to 'low' probability, driven by the lack of *Capabilities* and limited *Motivations*, despite the 'low' *Restraints*, which was somewhat surprising. This would imply that nothing further needs to be done to discourage the actor in this case. However, it also means that there is no counterbalance should motivational aspects or capabilities shift, thus continued monitoring and exploration of options to curtail the actor would be recommended. Furthermore, future work should also asses the economic aspect further, particularly how financial gain may affect motivational aspects and offset economic limitations.

7 Conclusion

Contents

Cyber security is still a young and rapidly-evolving field, not only in regards to technological developments but also its interaction with national and international security. This is particularly true when exploring the various uses of software as a weapon, ranging from mere nuisance to severe attacks.

The most closely related concept is that of cyber weapons, often also referred to as digital or virtual weapons – yet, there is no agreement on concepts or a definition. Similarly, there is disagreement on the topic of cyber war, with some refuting the possibility whilst others do not. The vast majority of work in this area takes a technical approach, centring on vulnerabilities or various forms of malicious software, including ways to either secure systems or highlight the flaws in doing so.

This thesis contributed to the topic by addressing SaaW taking a mixed-methods approach bringing together elements of CS, IR, and Strategic Studies. Across seven chapters, it first set out the motivation, including a presentation of three interrelated questions, which will be summarised in turn below.

7.1 Software as a Weapon

Chapter 2 presented a literature review that contextualised weapons and their pursuit as a strategic tool in relation to war and warfare, before addressing cyber weapons and related constructs more specifically. This provided the groundwork for examining the first question '*What does it mean for software to be a weapon?*' discussed in Chapter 3.

The high-level results showed that several aspects matter when discussing SaaW, particularly the interaction of design, use, and damage caused. Relying solely on designed or even intended use is not sufficient, as it would exclude potential accidents or mistakes. Yet at the same time, relying only on damage done or harm caused is also not adequate. This led to three categories, the first of which spans a spectrum of software specifically designed or used to cause damage or harm, with no other primary use. While malware can be a weapon, where the exact line between is to be drawn in remains contested and in flux, with the main disagreement being between CIA and cyber-physical damage requirements. The second category of SaaW centres on dual-use software, which for example is vital for security research or reviews on the one hand but can equally be used as a first step for, or as part of, cyber attacks on the other. The last and final category encompasses software that for example causes accidental and unintentional damage.

A questionnaire was designed to explore opinions on software, malware, and weapons, sourcing public attitudes towards across three respondent groups: *Academics*, *Military Officers* and *General Public (Others)*, contributing to the findings. Aside from sourcing attitudes towards SaaW, there was a contention that experience or work sector, as well as knowledge of computer science or international relations might lead to different responses. *Military* officers were chosen because they will have been trained extensively in the use of weapons regardless of their specialism, as well as the responsibility and consequences of their use. *Academic* respondents on the other hand are often considered to have access to, and influence on, decision makers, thus also giving them an important role in forming opinions.

A total of 46 questions including consent and demographics were asked, addressing three aspects: weapons, their nature and constitution; software, malware, and understandings thereof; as well as SaaW in the context of international security, including capabilities and proliferation. Most questions used the common five-point Likert rating scale, allowing for a 'neutral' response as well as open-ended free-text, leading to analysis using MWU and KWH tests, qualitative analysis, text-mined comparisons, as well as a structural examination of the questionnaire itself via a PCA. One aspect of the questionnaire centred on whether or not software/malware capabilities lead to more insecurity than security in general, for a state actor, or the international system. Whilst opinions were split, most respondents remained neutral or tended to agree, regardless of groups. This could be seen to support the argument of expanding cyber commands (or equivalent institutions) and related strategies to combat further insecurity, yet it could also be used to argue that capabilities in the cyber domain act as a disrupting force fuelling the security dilemma. Similarly, only about one fifth disagreed that software/malware capabilities provide a deterrent, adding to the ongoing academic debate on the topic.

Unsurprisingly, views on the meaning of term 'weapon' varied greatly. However, there was also some agreement, with about two-thirds considering it to be an object that is designed or can be used to cause harm or damage, and is offensively driven. Furthermore, a majority of respondents believe that there is a difference between an object being a weapon and being used as one, regardless of age, background, or expertise. Whilst this concurs with initial thoughts on the topic, harm or damage done by defensive measures should also be considered. This not only includes the increase of active countermeasures but also cyber strategies being adopted, such as the US' approach of persistent engagement/forward defence and the implications thereof [288].

Whether malware is a weapon or not split opinions, as did when asking whether it depends on a threshold. On the one hand, *Academics* are more likely to disagree that software should be treated like any physical object, *Military* respondents are more likely to agree and *Others* are indecisive. When suggesting a three-tiered

classification of objects and software/malware (tool/dual-use/weapon), there is agreement across the groups.

Examining the analogies used to describe concepts, the respondents appear to think of 'software' as having a broad function, whilst 'malware' is characterised by unique expressions forming a bridging construct to a 'weapon'. It appears as if malware is a stepping stone to software becoming a weapon, yet something 'more' is required. Yet this 'more' remains elusive and provides a starting point for future work. There is also disagreement as to what effect causes it to become a weapon, ranging from requiring physical damage to a living being, structure or system, or merely an attack on CIA.

Thus, an effects-based definition is not sufficient, yet neither is one solely reliant on intent, as for example Rid & McBurney's. When asked whether damage mattered regardless of intent, most respondents disagreed. Yet at the same time, there was no agreement that intent was the decisive factor. Therefore, SaaW should also account for unintended consequences, such as collateral damage or economic effects. Particularly the former is a criterion in most state actors' military engagements, being illegal under many definitions, such as Article 51(5)(b) of the 1977 Additional Protocol I to the Geneva Convention.

Future work should first narrow down the questionnaire based on the results of the PCA, reducing it to around 15-20 questions. This would make it more efficient yet analytically viable, whilst at the same time making higher response rates more likely. Furthermore, this would provide a basis for longitudinal data, allowing for an assessment of the changing nature and attitudes. Another aspect that should be considered is exploring the 'more' that links concepts of malware to those of a weapon, as that could provide a basis for frameworks in the cyber domain. Another aspect that could be addressed is the examination of the sources and effects of bias, which could, for example, be performed through a psychometric evaluation.

7.2 Factors Contributing to SaaW Pursuit

Chapter 4 combined aspects of the literature review and the questionnaire to explore what factors contribute to the pursuit of SaaW. It further drew upon weapon and proliferation theories, with the purpose of creating a group of determinants that act as variables, which are then interconnected to create a conceptual model. This model is naturally simplified but nonetheless introduces objectivity to the discourse and can provide a variety of stakeholders with a unified starting point.

The literature to what motivates actors to pursue nuclear weapons has shifted since the end of the Cold War, in line with more practical definitions of security [46]. Most proliferation literature has been devoted to WMD – NBC weapons – often grouped together for convenience to distinguish them from conventional weapons technologies. Although these weapon technologies are in many ways different, particularly in their destructive capability, the decades of research on the motivations behind the spread of other weaponry cannot be ignored.

Having discussed the various theories, the numerous factors were extracted, applied to the cyber domain, and broken into three interlinked sub-sets: restraint, motivation, and capability. Restraints are, essentially, reasons for an actor not to develop or proliferate and are split into four main elements with subcategories:

- Potential Collateral
- Domestic Safeguards
- International Agreements
- Fear of Retaliation

Motivation on the other hand are the counterpart to restraint, and are formed through five main interlinked and sometimes competing factors with subcategories:

- Fear of Retaliation
- Foreign Policy
- Domestic Policy
- Ideological Factors
- Economic Factors

Lastly, an actor's capability consists of six core factors, again with subcategories.

- Economic
- Strategic
- Tactical
- Research
- Development
- Reconnaissance

Each element was discussed in turn, examining interrelationships and sub-elements, both competing and complementing, creating a conceptual model. It is used as a basis for the operational BN model created in Chapter 5 and discussed below in the next section.

One of the main shortcomings of this model is that an opponent is only considered via limited direct input, whilst in reality most interactions, particularly conflict and war are reciprocal. Furthermore, in reality, there are often numerous actors at play. This could be mitigated by having several models that interact or are connected. A question then arises is how much of the 'other' an actor can see and with what 'lag', which might necessitate adding an additional layer or barrier between the actors. Furthermore, future versions could be broken down further into actor templates that are more tailored, for example to state and non-state actors, or even various types of state actors.

7.3 Probability of SaaW Pursuit

The sixth chapter applied the conceptual and operational model to three scenarios, engaging with the last research question: *'What is the probability that a given actor is pursuing SaaW?'*. It did this by using three actors, representing terrorists, state powers, and generic attackers. Data was based on expert opinions solicited via 30 interviews, as well as five datasets, populating nodes with prior probability and relative weightings of dependencies shown. Das' approach of weighting relative parent-node's influence strength was implemented, creating a linearly growing set of probability distributions and Tornado plots were used to show relative node influence strength. The main advantage of using a BN is its ability to model and account for uncertainty, which is particularly prevalent in the cyber domain.

Overall, the *State Actor* was shown to have ambivalent proliferation probability, with 'high', 'medium' and 'low' conditions being very close to each other. On the one hand, *Restraints* are 'high', yet so are *Capabilities*, whilst *Motivations* are almost 40% 'low', but also just over 40% split between 'medium' and 'high'. This result implies that pursuit is currently unlikely, however, an increase in motivation may

shift the balance. This could, for example, be caused by geopolitical or domestic factors and events. Given the level of restraining elements, additional disincentives could be sought here, for example by seeking to increase them or by adding new ones. Alternatively, the motivational factors or capabilities would need to be explored further, to see how the actor can be best discouraged.
The *Terrorist Actor* on the other hand has an overall 'medium' to 'low' probability, driven by the lack of *Capabilities* and limited *Motivations*, despite the 'low' *Restraints*. This would, similarly to the state-actor above, mean that currently nothing would need to be done to discourage the actor further. However, unlike the state-actor, the restraining element is 'low', thus any shift in the *Motivations* or *Capabilities* would be harder to counter-balance. Therefore, it would be recommended to continue monitoring developments and explore further options to curtail this actor.

Here, there are several starting points for future work. Firstly, this concerns suggestions raised by the participants regarding the network structure. This includes additional nodes within *Restraints* to explicitly denote 'digital' or 'economic collateral', as well as opponent's cyber capabilities. The concept of blow-back should also be explored further, caused on the one hand by a cyber attack leading to a response, and by the same capability damaging the initiating actor on the other. The former could be implemented as its own *Cyber Retaliation* sub-node under *Fear of Retaliation*, whilst the latter could sit as *Cyber Blow-back* under *Potential Collateral*. Furthermore, questions were raised on economic cost and attribution, which are not reflected in the current model but future work should take into account. Other proposals included the re-labelling or re-defining of *Security* related nodes within *Motivations*, which should also be explored.
Secondly, the same BN structure was used for all three cases but the label, or conceptional meaning, of nodes changed, adjusting to the actor in question. It does not, for example, make sense to talk about 'national security' when referring to a non-state actor. 'Group security', however, does. The dynamics are clearly different but

for this iteration of the model, they were considered to be sufficiently related. The advantage of keeping the structure intact allows for easier comparison and analysis. However, it reduced the informative and explanatory power. Future versions could be shaped around a specific actor type or even subtype incorporating greater detail of the interrelationships and relative importance of factors. These would then act as starting points that could then be applied to, and modified for, specific actors and/or scenarios, particularly as certain nodes or connections carry limited or no weight for one actor but are vital for another. Similarly, certain nodes or sub-nodes should be explored further, particularly including an expansion of the technical intricacies, such as high-level suggestions shown in Appendix B, which include bot(net) and C&C components, or more detailed build, install and delivery mechanisms.

This also ties in with a third aspect, namely the challenge of improving data acquisition and datasets. Whilst many real-world applications that rely on expert elicitation and effort was made to follow an accepted process [165], this work would benefit greatly from improved data access, possibly big-data coupled with learning algorithms. If reliance on elicitation remains, a greater focus should be placed on dividing the model up into 'subject-matter' groupings for experts in those areas to focus on. This would also address concerns raised for example regarding the *Development* node and its sub-nodes. Those with a more technical background wanted greater granularity, whilst others considered the current level of detail too great, instead suggesting the use of only one node. An alternative solution to dividing-up the focus areas could be a 'zoom' solution that allows the responder to decide on the level of detail.

Lastly, a time component or comparisons to conventional or nuclear weapon scenarios could be run, for example by examining the applicability of CfPD and related methods, which could span both aspects. The SaaW BN is currently a snapshot in time. However, as actors and weapons change over time, so should the model, adjusting the values of the variables and interconnections. A Dynamic Bayesian Network (DBN) would capture this process by essentially creating a time-series, representing multiple states of variables across time in given steps. It consists of a set

of variables at a given time in combination to a transition model, linking time t with for example $t-1$. In order to keep track of the current state, all past observations have to be taken into account. More formally, a DBN is a generalisation of Kalman Filters (KF) and Hidden Markov Models (HMM). Whilst HMM is very explicit and each node represents a system state, the DBN uses the nodes to represent the system dimension. This means that a DBN may have exponentially fewer parameters than its corresponding HMM and that inference maybe equally faster. However, whilst DBNs account for value changes over time, they do not allow structure or parameter changes, so other alternatives will also have to be considered.

Appendices

A

Summary of Questions and Responses

Table A.1: Summary of Questions and Responses: Q8-31

#	Question	Str. Dis.	Dis.	Nth.	Agr.	Str. Agr.
8	What is a weapon?	free-text				
9	Please list three different types of weapons	free-text				
10	An everyday object or tool (e.g. a cup, a chair, a shirt) can be a weapon	8	11	10	29	38
11	An everyday object or tool (e.g. a cup, a chair, a shirt) can be used as a weapon	3	0	2	33	58
12	There is no difference between an object being a weapon and an object being used as a weapon	25	37	12	9	13
13	What turns an object into a weapon/when does an object become a weapon?	free-text				
14	Weapons are only used offensively	34	47	3	10	2
15	There are three types of objects: 1) those created for the sole purpose of being used as a weapon (offensively or defensively); 2) those created for dual use, either as a tool or a weapon, depending on the situation; 3) those created to be used as a tool but that can situationally become/be used as a weapon. To what extent do you agree with the categorisation above?	2	9	5	51	29
16	If an object with an un-precedented destructive potential is created, but nobody is aware of it, it is still a weapon.	7	14	24	31	20
17	What is software?	free-text				
18	Software should be treated like any physical object or tool	4	22	24	34	12
19	What is malicious software (malware)?	free-text				
20	All malware is equally dangerous.	25	52	9	5	5
21	On what basis should malware be classified?	free-text				
22	Software and/or malware can cause physical damage	3	4	6	46	37
23	All software has the potential to become malware	5	23	31	29	8
24	Malware is a weapon	2	13	20	40	21
25	Malware can be used as a weapon	1	2	6	48	39
26	Malware can sometimes be/be used as a weapon	6	9	13	47	21
27	Malware needs to cross a threshold to be considered a weapon	12	34	15	34	1
27a	If you agree, what would you base the threshold on?	free-text				
27b	Should that type of weapon be given a new name to reflect its potency, e.g. cyber weapon?	3	11	29	41	5
28	It makes sense to approach software/malware attacks with (traditional) weapon terminology (e.g. ‘warhead’, ‘trigger’, ‘payload’)	7	24	24	38	3
28a	And/or measures (e.g. ‘calibre’, ‘yield’, ‘range’)	10	21	25	39	1
29	Software/malware causing physical damage to a living being, a structure or system (a ‘cyber-physical attack’) is a weapon; software/malware causing damage to data integrity, accessibility and confidentiality is not a weapon	28	45	8	15	0
30	Software/malware causing damage to a living being, a structure or system (a ‘cyber-physical attack’) is a weapon, regardless of the type or severity of damage	2	7	7	52	28
31	The damage caused determines whether the software/malware is a weapon or not, regardless of the intentions of the attacker	18	44	12	18	4

Table A.2: Summary of Questions and Responses: Q32-46

#	Question	Str. Dis.	Dis.	Nth.	Agr.	Str. Agr.
32	The intention of the attacker determines whether the software/malware is a weapon or not, regardless of the damage caused	5	27	13	36	15
33	Software/malware can be separated into three types: 1) created for the sole purpose of being used as a weapon (offensively or defensively) 2) created for dual use, either as a tool or a weapon, depending on the situation 3) created to be used as a tool but that can situationally become a weapon	4	14	23	45	10
34	Most attacks by software/malware fall only into three categories: 1) Mere nuisance 2) Summary offences / misdemeanour / petty crime 3) Indictable offence / felony	3	8	25	48	12
35	Defensive software capabilities are vital to a state's security	3	2	3	38	50
35a	Any additional comments?	free-text				
36	Offensive software capabilities are vital to a state's security	2	16	25	30	23
36a	Any additional comments?	free-text				
37	Software/malware capabilities lead to more insecurity than security in general	3	15	37	33	8
37a	For a state:	7	14	37	29	9
37b	For the international system:	7	14	37	29	9
37c	Any additional comments?	free-text				
38	Software/malware capabilities provide a deterrent in general	3	21	29	39	4
38a	When visible/previously demonstrated:	5	15	26	39	11
38b	When merely suggested/implied:	4	20	37	33	2
38c	Any additional comments?	free-text				
39	It is possible to showcase software/malware capabilities without losing a technological advantage	4	20	30	38	4
39a	Any additional comments?	free-text				
40	Software capabilities should be regulated globally, similar to the production and/or proliferation of other weapons	9	16	27	39	5
40a	Any additional comments?	free-text				
41	Software/malware capabilities have propelled a vast array of new actors into the security field, nationally and internationally	0	2	16	52	26
41a	Any additional comments?	free-text				
42	Software/malware capabilities have rendered state-centric security models obsolete	5	42	34	12	3
42a	Any additional comments?	free-text				
43	Software/malware is a vital component of modern warfare	0	3	14	41	38
43a	Any additional comments?	free-text				
44	Software/malware capabilities can be clearly differentiated between civil and military applications	12	38	22	22	2
44a	Any additional comments?	free-text				
45	Software/malware exemplify asymmetric warfare	1	9	29	42	15
45a	Please explain you answer:	free-text				
46	An attack by software/malware can be considered an act of war	0	5	20	52	19
46a	Any additional comments?	free-text				

B

Initial Technical Capability Components

Figure B.1: Bot(net) and C&C components

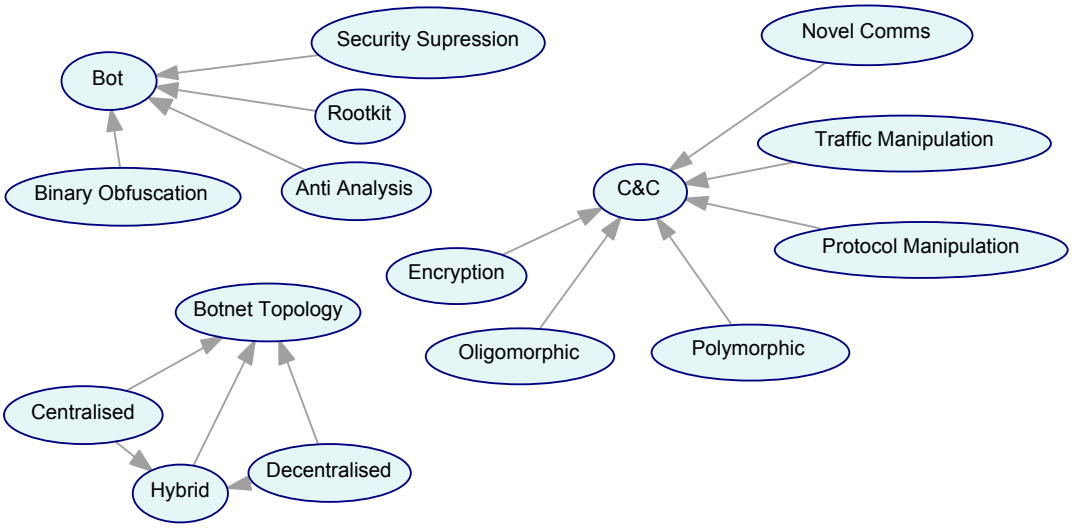
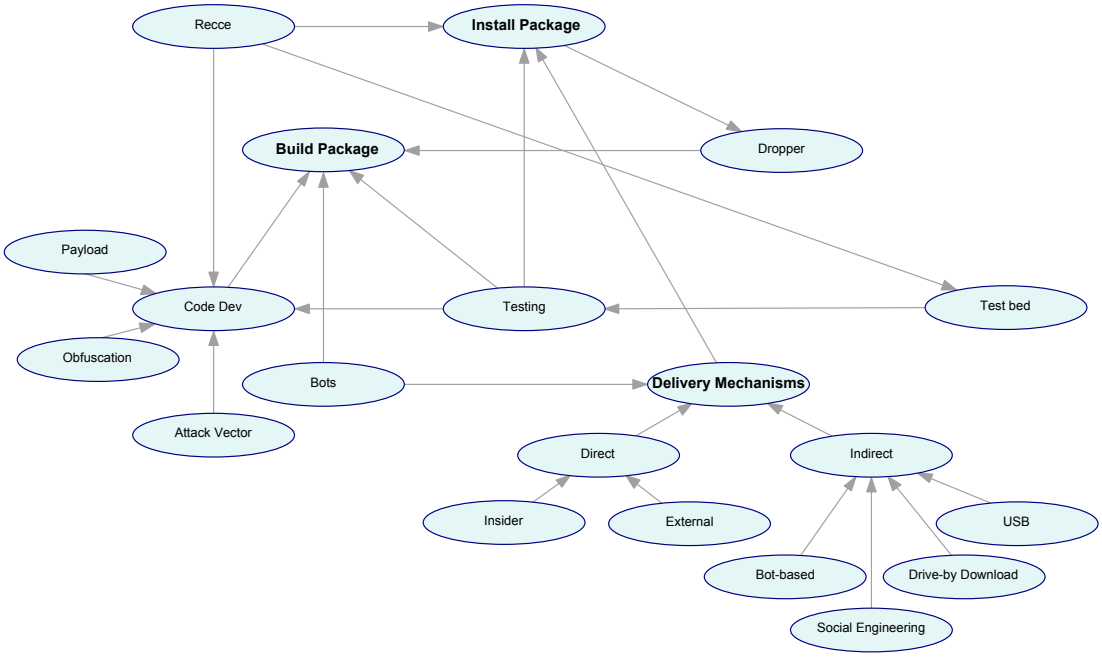


Figure B.2: Build, Install and Delivery Mechanisms



Bibliography

[1] L. Ablon and A. Bogart, *Zero Days , Thousands of Nights The Life and Times of Zero-Day Vulnerabilities and Their Exploits*. 2017, pp. 65–71. [Online]. Available: http://www.rand.org/pubs/research_reports/RR1751.html.

[2] J. Acton, "Cyber Weapons and Precision-Guided Munitions", in *Understanding Cyber Conflict: Fourteen Analogies*, G. Perkovich and A. E. Levite, Eds., Georgetown University Press, 2016, ch. 3, pp. 45–60.

[3] E. Adler, "The emergence of cooperation: national epistemic communities and the international evolution of the idea of nuclear arms control", *International Organization*, vol. 46, no. 01, p. 101, 1992. [Online]. Available: http://www.journals.cambridge.org/abstract_S0020818300001466.

[4] I. Agrafiotis, M. Bada, P. Cornish, S. Creese, M. Goldsmith, E. Ignatuschtschenko, T. Roberts, and D. M. Upton, "Cyber Harm: Concepts, Taxonomy and Measurement", 2016.

[5] I. Agrafiotis, J. R. C. Nurse, O. Buckley, P. A. Legg, S. Creese, and M. Goldsmith, "Identifying attack patterns for insider threat detection", *Computer Fraud & Security*, vol. 2015, no. 7, pp. 9–17, 2015. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S136137231530066X.

[6] I. Agrafiotis, J. R. C. Nurse, M. Goldsmith, S. Creese, and D. Upton, "A taxonomy of cyber-harms: Defining the impacts of cyber-attacks and understanding how they propagate", *Journal of Cybersecurity*, pp. 1–15, 2018.

[7] I. Ali and P. Stewart, *Exclusive: U.S. carried out secret cyber strike on Iran in wake of Saudi oil attack: officials*, 2019. [Online]. Available: https://www.reuters.com/article/us-usa-iran-military-cyber-exclusive/exclusive-u-s-carried-out-secret-cyber-strike-on-iran-in-wake-of-saudi-oil-attack-officials-idUSKBN1WV0EK (visited on 12/10/2019).

[8] G. Allison, *NATO working on cyber attack trigger for Article 5*, 2018. [Online]. Available: https://ukdefencejournal.org.uk/nato-working-on-cyber-attack-trigger-for-article-5/ (visited on 12/03/2018).

[9] D. Alperovitch, "Towards establishment of cyberspace deterrence strategy", *2011 3rd International Conference on Cyber Conflict*, pp. 1–8, 2011.

[10] J. Altmann, W. Liebert, G. Neuneck, and J. Scheffran, "Preventive Arms Control as a Prerequisite for Conversion of Military-Related R&D", in *Conversion of Military R&D*, J Reppy, J Rotblat, J Holdren, and V Avduyevsky, Eds., Macmillan, 1998, pp. 255–271.

[11] C. Alwardt and J. Polle, "Internationale Rüstungskontrollbemühungen zu autonomen Waffensystemen: Definitionen, Technik und sicherheitspolitische Implikationen", *Sicherheit und Frieden*, vol. 36, no. 3, pp. 133–139, 2018.

[12] C. Anderson, "Considerations on Wassenaar Arrangement Control List Additions for Surveillance Technologies", p. 18, 2015. [Online]. Available: https://www.accessnow.org/cms/assets/uploads/archive/AccessWassenaarSurveillanceExportControls2015.pdf.

[13] L. Arimatsu, "A treaty for governing cyber-weapons: Potential benefits and practical limitations", *2012 4th International Conference on Cyber Conflict, CYCON 2012 - Proceedings*, no. September 2011, pp. 91–109, 2012.

[14] Associated Press, *US launched cyber attack on Iranian rockets and missiles*, 2019. [Online]. Available: https://www.theguardian.com/world/2019/jun/23/us-launched-cyber-attack-on-iranian-rockets-and-missiles-reports (visited on 12/10/2019).

[15] A. Backstrom and I. Henderson, *New capabilities in warfare: An overview of contemporary technological developments and the associated legal and engineering issues in Article 36 weapons reviews*, 886. 2013, vol. 94, pp. 483–514.

[16] M. R. Baer, "Toward Criteria for International Cyber Weapons", in *IEEE World Cyberspace Cooperation Summit IV (WCC4)*, 2013.

[17] C. Bagot and R. Rush, *British-American Diplomacy Exchange of Notes Relative to Naval Forces on the American Lakes*, 1818. [Online]. Available: https://avalon.law.yale.edu/19th_century/conv1817.asp.

[18] B. Barth, *Snack attack: A crimeware-as-a-service menu for wannabe hackers*, 2016. [Online]. Available: https://www.scmagazine.com/snack-attack-a-crimeware-as-a-service-menu-for-wannabe-hackers/article/527865/ (visited on 01/08/2018).

[19] I. Barzashka, "Are Cyber-Weapons Effective?", *The RUSI Journal*, vol. 158, no. 2, pp. 48–56, 2013.

[20] M. Bayes and M. Price, "An Essay towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, F. R. S. Communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S.", *Philosophical Transactions of the Royal Society of London*, vol. 53, pp. 370–418, 1763. [Online]. Available: http://rstl.royalsocietypublishing.org/cgi/doi/10.1098/rstl.1763.0053.

[21] BayesFusion, *GeNIe Academic*, 2018. [Online]. Available: http://www.bayesfusion.com/.

[22] BBC Technology, *Ransomware cyber-attack threat escalating - Europol - BBC News*, 2017. [Online]. Available: http://www.bbc.co.uk/news/technology-39913630 (visited on 05/14/2017).

[23] A. T. Beck, C. H. Ward, M Mendelson, J Mock, and J Erbaugh, "An Inventory for Measuring Depression", *Archives of General Psychiatry*, vol. 4, pp. 561–571, 1961. arXiv: 1011.1669v3.

[24] B. Bencsáth, G. Pék, L. Buttyán, and M. Félegyházi, "The Cousins of Stuxnet: Duqu, Flame, and Gauss", *Future Internet*, vol. 4, no. 4, pp. 971–1003, 2012.

[25] R. K. Betts, "Paranoids, Pygmies, Pariahs and Nonproliferation Revisited", *Security Studies*, vol. 2, no. 3-4, pp. 100–124, 1993.

[26] D. Betz, *Carnage and Connectivity*. London: Oxford University Press, 2015.

[27] ——, "'cyberwar' is not coming", *Infinity Journal*, vol. 1, no. 3, pp. 21–24, 2011.

[28] D. J. Betz and T. Stevens, "Cyberspace and the State: Toward a Strategy for Cyber-power", *Adelphi Series*, vol. 51, no. 424, 2011.

[29] A. Bezverkhyi, *Petya.A / NotPetya is an AI-powered cyber weapon, TTPs lead to Sandworm APT group*, 2017. [Online]. Available: `https://socprime.com/blog/petya-a-notpetya-is-an-ai-powered-cyber-weapon-ttps-lead-to-sandworm-apt-group` (visited on 01/28/2020).

[30] L. Bilge and T. Dumitras, "Before We Knew It: an Empirical Study of Zero-Day Attacks in the Real World", *Proceedings of the 2012 ACM Conference on Computer and Communications Security – CCS'12*, pp. 833–844, 2012.

[31] D. Blake and J. S. Imburgia, "Bloodless weapons?", *Air Force Law Review*, vol. 66, pp. 157–204, 2010. [Online]. Available: `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1850831`.

[32] W. Boebert, "A Survey of Challenges in Attribution", *Proceedings of a workshop on Deterring CyberAttacks*, pp. 41–52, 2010. [Online]. Available: `https://cs.brown.edu/courses/csci1800/sources/lec12/Boebert.pdf`.

[33] J. Bolt, R. Inklaar, H. De Jong, and J. L. Van Zanden, "Rebasing 'Maddison': new income comparisons and the shape of long-run economic development", 2018, [Online]. Available: `https://www.rug.nl/ggdc/historicaldevelopment/maddison`.

[34] W. H. Boothby, *Weapons and the Law of Armed Conflict*. Oxford University Press, 2016.

[35] E. D. Borghard and S. W. Lonergan, "Confidence Building Measures for the Cyber Domain", *Strategic Studies Quarterly*, vol. 12, no. 3, pp. 10–49, 2018.

[36] V. Boulanin, "Implementing Article 36 Weapon Reviews in the Light of Increasing Autonomy in Weapon Systems", *SIPRI Insights on Peace and Security no. 1*, pp. 1–28, 2015.

[37] M. Bourne, *Arming Conflict: The Proliferation of Small Arms*. Palgrave Macmillan, 2007.

[38] S. J. Brams, *Superpower games: applying game theory to superpower conflict*. Yale University Press, 1985.

[39] S. Bratus, D. J. Capelis, M. Locasto, and A. Shubina, "Why Wassenaar Arrangement's Definitions of Intrusion Software and Controlled Items Put Security Research and Defense At Risk — And How To Fix It", Tech. Rep., 2014, pp. 1–13.

[40] J. S. Breese and D. Heckerman, "Decision-theoretic troubleshooting: A framework for repair and experiment", in *Proceedings of the 12th Annual Conference on Uncertainty in Artificial Intelligence (UAI-96)*, San Francisco, CA: Morgan Kaufmann, 1996, pp. 124–132. [Online]. Available: `http://dl.acm.org/citation.cfm?id=2074299`.

[41] S. W. Brenner, "'at light speed': Attribution and response to cybercrime/terrorism/warfare", *The Journal of Criminal Law and Criminology*, vol. 97, no. 2, pp. 379–476, 2007.

[42] C. S. Brown and D. Friedman, "A Cyber Warfare Convention? Lessons from the Conventions on Chemical and Biological Weapons", *Arms Control and National Security: New Horizons*, pp. 45–63, 2014.

[43] G Brown and A Metcalf, "Easier Said Than Done: Legal Reviews of Cyber Weapons", *Journal of National Security Law & Policy*, vol. 7, pp. 115–138, 2014.

[44] B. Buchanan, *The Cybersecurity Dilemma: Hacking, Trust and Fear Between Nations*. Oxford University Press, 2017.

[45] T. Burgers and D. R. S. Robinson, "Keep Dreaming: Cyber Arms Control is Not a Viable Policy Option", *Sicherheit & Frieden*, vol. 36, no. 3, pp. 140–145, 2018.

[46] B. Buzan, "New Patterns of Global Security in the Twenty-First Century", *International Affairs (Royal Institute of International Affairs 1944-)*, vol. 67, no. 3, pp. 431–451, 1991.

[47] ——, "Rethinking security after the Cold War", *Cooperation and conflict*, vol. 32, no. 1, pp. 5–28, 1997.

[48] B. Buzan, C. A. Jones, and R. Little, *The logic of anarchy: neorealism to structural realism*. Columbia University Press, 1993.

[49] S. Caltagirone, A. Pendergast, and C. Betz, "The Diamond Model of Intrusion Analysis", vol. 298, no. 0704, pp. 1–61, 2013.

[50] J. C. Campbell, "Danger Assessment", Johns Hopkins University, School of Nursing, Tech. Rep., 2001.

[51] Carnegie Endowment for International Peace, *Cyber Norms Index*. [Online]. Available: `https://carnegieendowment.org/publications/interactive/cybernorms` (visited on 09/01/2019).

[52] E. H. Carr, *The Twenty Years' Crisis*. New York: Harper & Row, 1939.

[53] J. Cater, "National Security Directive 18: U.S. National Security Strategy", Washington, DC, Tech. Rep., 1977, p. 7.

[54] CBSNews, *'Great Cannon' is China's latest cyber weapon*, 2015. [Online]. Available: `https://www.cbsnews.com/video/great-cannon-is-chinas-latest-cyber-weapon/` (visited on 09/17/2019).

[55] G. Chafetz, "The End of the Cold War and the Future of Nuclear Proliferation: An Alternative to the Neorealist Perspective", *Security Studies*, vol. 2, no. 3-4, pp. 125–158, 1993.

[56] A. Cherepanov and R. Lipovsky, *GreyEnergy: Updated arsenal of one of the most dangerous threat actors*, 2018. [Online]. Available: `https://www.welivesecurity.com/2018/10/17/greyenergy-updated-arsenal-dangerous-threat-actors/`.

[57] B. C. K. Choi, A. W. P. Pak, and W. P. Anita, "A Catalog of Biases in Questionnaires", *Preventing Chronic Disease*, vol. 2, no. 1, pp. 1–13, 2005.

[58] S. J. Cimbala, "On Nuclear War : Deterrence, Escalation, and Control", *Military and Strategic Affairs*, vol. 4, no. 3, pp. 25–43, 2012.

[59] C. Cimpanu, *China resurrects Great Cannon for DDoS attacks on Hong Kong forum*, 2019. [Online]. Available: `https://www.zdnet.com/article/china-resurrects-great-cannon-for-ddos-attacks-on-hong-kong-forum/` (visited on 01/20/2020).

[60] ——, *Surprise! NotPetya Is a Cyber-Weapon. It's Not Ransomware*, 2017. [Online]. Available: `https://www.bleepingcomputer.com/news/security/surprise-notpetya-is-a-cyber-weapon-its-not-ransomware/` (visited on 01/20/2020).

[61] P. Cirenza, "An Evaluation of the Analogy Between Nuclear and Cyber Deterrence", PhD thesis, Stanford University, 2015.

[62] ——, *The flawed analogy between nuclear and cyber deterrence*, 2016. [Online]. Available: `https://thebulletin.org/2016/02/the-flawed-analogy-between-nuclear-and-cyber-deterrence/` (visited on 12/12/2010).

[63] J. Cirincione, *Bomb Scare: The History and Future of Nuclear Weapons*. New York: Columbia University Press, 2007.

[64] J. Cirincione, J. B. Wolfsthal, and M. Rajkumar, *Deadly Arsenals: Nuclear, Biological, and Chemical Threats*. 2005.

[65] R. Clarke and K. Robert, *Cyber war: The next threat to national security and what to do about it*. New York: HarperCollins, 2010.

[66] C. von Clausewitz, "Meine Vorlesungen über den kleinen Krieg", in *Schriften-Aufsätze-Studien-Briefe*, W Hahlweg, Ed., Göttingen, 1966, pp. 208–598.

[67] ——, "Meine Vorlesungen über den Kleinen Krieg", in *Lehrmeister des Kleinen Krieges von Clausewitz bis Mao Tse-Tun und Che Guevara*, Darmstadt, 1968.

[68] ——, *On War*, M. E. Howard and P Paret, Eds. Princeton University Press, 1989.

[69] R. Clayton, "Anonymity and traceability in cyberspace", Tech. Rep. 653, 2005, p. 189. [Online]. Available: `http://www.cl.cam.ac.uk/TechReports/`.

[70] D. Coats, *Russia's INF Treaty Violation*, 2018. [Online]. Available: `https://www.dni.gov/index.php/newsroom/speeches-interviews/speeches-interviews-2018/item/1923-director-of-national-intelligence-daniel-coats-on-russia-s-inf-treaty-violation` (visited on 01/20/2020).

[71] G. A. Coles, A. J. Brothers, J. Olson, and P. D. Whitney, "Assessing State Nuclear Weapons Proliferation: Using Bayesian Network Analysis of Social Factors", in *Proceedings of the Pacific Northwest International Conference on Global Nuclear Security*, Pacific Northwest National Lab, Richland: OSTI, 2010.

[72] T. Z. Collina, *Russia Breaches INF Treaty, U.S. Says*, 2014. [Online]. Available: `https://www.armscontrol.org/act/2014-09/news/russia-breaches-inf-treaty-us-says` (visited on 01/20/2020).

[73] S. Collins and S. McCombie, "Stuxnet: the emergence of a new cyber weapon and its implications", en, *Journal of Policing, Intelligence and Counter Terrorism*, vol. 7, no. 1, pp. 80–91, 2012.

[74] A. C. Constantinou, N. Fenton, and M. Neil, "Integrating expert knowledge with data in Bayesian networks: Preserving data-driven expectations when the expert variables remain unobserved", *Expert Systems with Applications*, vol. 56, pp. 197–208, 2016.

[75] G. F. Cooper and E Herskovits, "A Bayesian Method for the Induction of Probabilistic Networks From Data", *Machine Learning*, vol. 9, no. 4, pp. 309–347, 1992.

[76] P. T. Costa and R. R. McCrae, "Age differences in personality structure: a cluster analytic approach", *Journal of Gerontology*, vol. 31, no. 5, pp. 564–570, 1976.

[77] V. M. H. Coupe and C Linda, "Properties of Sensitivity Analysis of Bayesian Belief Networks", *Annals of Mathematics and Artificial Intelligence*, vol. 36, no. 4, pp. 323–356, 2002.

[78] A. Craig and B. Valeriano, "Conceptualising cyber arms races", in *2016 8th International Conference on Cyber Conflict (CyCon)*, vol. 2016-Augus, IEEE, 2016, pp. 141–158.

[79] M. van Creveld, *The transformation of war*. Free Press, 1991.

[80] M. Croon, *Methods for correlational research: Factor analysis, path analysis, and structural equation modeling*. Pearson, 2008.

[81] R. Crootof, "Regulating New Weapons Technology", in *The Impact of Emerging Technologies on the Law of Armed Conflict*, E. T. Jensen and R. T. Alcala, Eds., Oxford University Press, 2019, pp. 3–26.

[82] ——, "The Killer Robots are Here: Legal and Policy Implications", *Cardozo Law Review*, no. January, p. 80, 2015. [Online]. Available: `https://www.researchgate.net/profile/Rebecca_Crootof/publication/288825550_The_Killer_Robots_Are_Here_Legal_and_Policy_Implications/links/568432a508ae051f9af040d5/The-Killer-Robots-Are-Here-Legal-and-Policy-Implications.pdf`.

[83] *Customary IHL - Practice Relating to Rule 14. Proportionality in Attack*. [Online]. Available: `https://ihl-databases.icrc.org/customary-ihl/eng/docs/v2_rul_rule14` (visited on 01/08/2019).

[84] C. Daase, "Clausewitz and Small Wars", in *Clausewitz in the Twenty-First Century*, H. Strachan and A. Herberg-Rothe, Eds., Oxford University Press, 2007, pp. 182–195.

[85] B. Das, "Generating Conditional Probabilities for Bayesian Networks: Easing the Knowledge Acquisition Problem", *arXiv*, 2008. arXiv: `0411034v2 [cs]`. [Online]. Available: `http://arxiv.org/abs/cs/0411034v2`.

[86] K. Davis, "The demographic foundations of national power", *Freedom and Control in Modern Society*, pp. 206–243, 1954.

[87] Z. S. Davis, "The Realist Nuclear Regime", *Security Studies*, vol. 2, no. 3-4, pp. 79–99, 1993.

[88] R. S. Dewar, "The "Triptych of Cyber Security": A Classifi cation of Active Cyber Defence", in *Cyber Conflict (CYCON), 2014 6th International Conference on*, 2014, pp. 7–21.

[89] B. Dharamkar and R. R. Singh, "A Review of Cyber Attack Classification Technique Based on Data Mining and Neural Network Approach", *International Journal of Computer Trends and Technology.*, vol. 7, no. 2, pp. 100–105, 2014.

[90] F. J. Diez, "Parameter adjustement in Bayes networks. The generalized noisy {OR}-gate", *Proceedings of the 9th Conference on Uncertainty in Artificial Intelligence*, pp. 99–105, 1993.

[91] Dimensions, *Cyber Security Publications Overview - Dimensions*, 2019. [Online]. Available: `https://app.dimensions.ai/analytics/publication/viz/overview-publications?search_text=cybersecurity%2Ccybersecurity&search_type=kws&search_field=full_search` (visited on 01/10/2019).

[92] Y. Dinstein, *The Conduct of Hostilities under the Law of International Armed Conflict*. Cambridge University Press, 2016.

[93] B. Doherty, *John Perry Barlow 2.0*, 2004. [Online]. Available: `http://reason.com/archives/2004/08/01/john-perry-barlow-20` (visited on 01/25/2018).

[94] C. Doman, *The "Great Cannon"has been deployed again*, 2019. [Online]. Available: `https://cybersecurity.att.com/blogs/labs-research/the-great-cannon-has-been-deployed-again` (visited on 01/20/2020).

[95] P. Dombrowski and C. C. Demchak, "Cyber War, Cybered Conflict, and the Maritime Domain", *Naval War College Review*, vol. 67, no. 2, pp. 71–97, 2014.

[96] M. J. Druzdzel and L. C. van der Gaag, "Elicitation of Probabilities for Belief Networks: Combining Qualitative and Quantitative Information", in *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, 1995, pp. 141–148. arXiv: `1302.4943`.

[97] I. Duyvesteyn, "Between Doomsday and Dismissal: Cyber war, the Parameters of War, and Collective Defense", *Atlantisch Perspectief*, pp. 20–24, 2014.

[98] A. Dwyer and J. A. M. Silomon, *Dangerous Gaming: Cyber-Attacks, Air-Strikes and Twitter*, 2019. [Online]. Available: `https://www.e-ir.info/2019/09/23/dangerous-gaming-cyber-attacks-air-strikes-and-twitter/`.

[99] A. Echevarria, *Fourth-Generation War and Other Myths*. 2005.

[100] J. Eggenschwiler and J. A. M. Silomon, "Challenges and opportunities in cyber weapon norm construction", *Computer Fraud & Security*, no. December, pp. 11–17, 2018.

[101] ——, "Three Measures That Could Pave the Way to Building Successful Cyber Norms", *Council on Foreign Relations*, 2018. [Online]. Available: `https://www.cfr.org/blog/three-measures-could-pave-way-building-successful-cyber-norms`.

[102] J. Eriksson and G. Giacomello, "The Information Revolution, Security, and International Relations: (IR)relevant Theory?", *International Political Science Review/ Revue internationale de science politique*, vol. 27, no. 3, pp. 221–244, 2006.

[103] N. Falliere, L. O. Murchu, and E. Chien, "W32. stuxnet dossier", Tech. Rep., 2010.

[104] R. C. Feenstra, R. Inklaar, and M. P. Timmer, "The Next Generation of the Penn World Table", *American Economic Review*, vol. 105, no. 10, pp. 3150–3182, 2015. [Online]. Available: www.ggdc.net/pwt.

[105] N. Fenton and M. Neil, *Risk Assessment and Decision Analysis with Bayesian Networks*. CRC Press, 2012.

[106] M. P. Fischerkeller and R. J. Harknett, "Deterrence is Not a Credible Strategy for Cyberspace", *Orbis*, vol. 61, no. 3, pp. 381–393, 2017.

[107] S. Flank, "Exploding the Black Box: The Historical Sociology of Nuclear Proliferation", *Security Studies*, vol. 3, no. 2, pp. 259–294, 1993.

[108] L. Floridi, "The latent nature of global information warfare", *Philosophy and Technology*, vol. 27, no. 3, pp. 317–319, 2014.

[109] F. J. Fowler Jr, *Survey research methods*. Sage Publications Limited, 2013.

[110] C. R. Freeman, "Bayesian Network Analysis of Nuclear Acquisitions", PhD thesis, Texas A&M University, 2008.

[111] W. Fucks, *Formeln zur Macht: Prognosen über Völker, Wirtschaft, Potentiale*. Deutsche Verlags-Anstalt, 1965.

[112] F. Fukuyama, "The End of History?", *The National Interest*, pp. 1–18, 1989. arXiv: 1107.3081v1.

[113] D. Garcia, "Future arms, technologies, and international law: Preventive security governance", *European Journal of International Security*, vol. 1, no. 1, pp. 94–111, 2016.

[114] E. Gartzke, "The Myth of Cyberwar: Bringing War in Cyberspace Back Down to Earth", *International Security*, vol. 38, no. 2, pp. 41–73, 2013.

[115] K. Geers, "Cyber weapons convention", *Computer Law and Security Review*, vol. 26, no. 5, pp. 547–551, 2010.

[116] ——, "Cyber Weapons Convention", *Computer Law & Security Review*, vol. 26, no. 5, pp. 547–551, 2010.

[117] F. C. German, "A tentative evaluation of world power", *Journal of Conflict Resolution*, vol. 4, no. 1, pp. 138–144, 1960.

[118] J. Ghosh, M. Delampady, and T. Samanta, *An Introduction to Bayesian Analysis Theory and Method*. New York: Springer, 2006.

[119] D. M. Gibler, *International military alliances, 1648-2008*. CQ Press, 2009. [Online]. Available: http://cow.dss.ucdavis.edu/data-sets/formal-alliances.

[120] Global Cyber Security Capacity Centre, "Cybersecurity Capacity Maturity Model for Nations (CMM). Revised Edition", *Oxford Martin School*, no. CMM, pp. 1–60, 2017. [Online]. Available: https://www.sbs.ox.ac.uk/cybersecurity-capacity/content/cybersecurity-capacity-maturity-model-nations-cmm-revised-edition.

[121] M. A. Gomez, "Identifying Cyber Strategies vis-a-vis Cyber Power", *World Cyberspace Cooperation Summit IV*, 2013.

[122] Google, *Google Ngram Viewer - Cyber Security*, 2019. [Online]. Available: https://books.google.com/ngrams/graph?content=cyber+security%2Ccybersecurity&case_insensitive=on&year_start=1950&year_end=2008&corpus=15&smoothing=3&share=&direct_url=t4%3B%2Ccybersecurity%3B%2Cc0%3B%2Cs0%3B%3Bcybersecurity%3B%2Cc0%3B%3BCyberSecu (visited on 01/10/2019).

[123] ——, *Google Trends - Cyber Security*, 2019. [Online]. Available: https://trends.google.com/trends/explore?date=all{\&}q=cybersecurity,cybersecurity (visited on 01/10/2019).

[124] A. Gostev and I. Soumenkov, *Stuxnet/Duqu: The Evolution of Drivers*, 2011. (visited on 11/29/2018).

[125] S. Graham, "The End of Geography or the Explosion of Place?", *Progress in Human Geography*, vol. 22, no. 2, pp. 165–185, 1998.

[126] T. Graves, *Active Cyber Defense Certainty Act*, 2017. [Online]. Available: https://www.congress.gov/bill/115th-congress/house-bill/4036.

[127] A. Greenberg, *Petya Ransomware Epidemic May Be Spillover From Cyberwar*, 2017. [Online]. Available: https://www.wired.com/story/petya-ransomware-ukraine/ (visited on 12/29/2019).

[128] ——, *Shopping for zero-days: A price list for hackers' secret software exploits*, 2012. (visited on 01/25/2018).

[129] ——, *The Untold Story of NotPetya, the Most Devastating Cyberattack in History*, 2018. [Online]. Available: https://www.wired.com/story/notpetya-cyberattack-ukraine-russia-code-crashed-the-world/ (visited on 01/20/2020).

[130] P. Hakkarainen, *Cyber Weapon Target Analysis*. BoD, 2014.

[131] J. Y. Halpern, *Reasoning about Uncertainty*, 2nd. MIT press, 2017.

[132] M Hamilton, "A rating scale for depression", *Journal of neurology, neurosurgery, and psychiatry*, vol. 23, pp. 56–62, 1960.

[133] M. Hansel, M. Mutschler, and M. Dickow, "Taming cyber warfare: lessons from preventive arms control", *Journal of Cyber Policy*, vol. 3, no. 1, pp. 44–60, 2018.

[134] J. Hart, "Three approaches to the measurement of power in international relations", *International Organization*, vol. 30, no. 02, p. 289, 1976.

[135] Harvard Nuclear Study Group, *Living With Nuclear Weapons*. Bantam Books, 1983.

[136] N. A. Hassan, "Ransomware Families", in *Ransomware Revealed*, Berkeley, CA: Apress, 2019, pp. 47–68.

[137] U. Häussler, "Cyber Security and Defence from the Perspective of Articles 4 and 5 of the NATO Treaty", *International Cyber Security Legal & Policy Proceedings*, no. June, pp. 100–125, 2010.

[138] D Heckerman, D Geiger, and D. Chickering, "Learning Bayesian Networks: The Combination of Knowledge and Statistical Data", *Machine Learning*, vol. 20, pp. 197–243, 1995.

[139] M. Henrion, "Practical issues in constructing a Bayes belief network", *International Journal of Approximate Reasoning*, vol. 2, no. 3, p. 337, 1988. arXiv: 1304.2725. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/0888613X88901466.

[140] T. Herr, "Malware Counter-proliferation and the Wassenaar Arrangement", in *2016 8th International Conference on Cyber Conflict (CyCon)*, vol. 2016-Augus, IEEE, 2016, pp. 175–190.

[141] ——, "PrEP : A Framework for Malware & Cyber Weapons", *The Journal of Information Warfare*, vol. 13, no. 1, 2014.

[142] ——, *Proliferation in Cybersecurity*, 2017. [Online]. Available: http://magazine.milcyber.org/stories/proliferationincybersecurity%0AProliferation (visited on 01/08/2018).

[143] T. Herr and E. Armbrust, "Milware: The Implications of State Authored Malicious Software", *Social Science Research Network*, 2015.

[144] T. Herr, B. Schneier, and C. Morris, "Taking Stock - Estimating Vulnerability Rediscovery", 2017.

[145] M. Hof, "Questionnaire Evaluation with Factor Analysis and Cronbach's Alpha", *Citeseer*, pp. 1–11, 2012. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.297.6430&rep=rep1&type=pdf.

[146] K. H. Höhn, "Geopolitics and the Measurement of National Power", PhD thesis, Universität Hamburg, 2011.

[147] R. Holcombe, "Development of a Bayesian Network to Monitor the Probability of Nuclear Proliferation", no. 1998, 2008.

[148] M. C. Horowitz, "The Diffusion of Military Power: Causes and Consequences for International Politics", Doctoral, Harvard University, 2006.

[149] ——, *The Diffusion of Military Power: Causes and Consequences for International Politics*. New Jersey: Princeton University Press, 2010.

[150] M. C. Horowitz, "When Speed Kills: Autonomous Weapon Systems, Deterrence, and Stability", *SSRN Electronic Journal*, vol. 42, no. 6, pp. 764–788, 2019.

[151] M. C. Horowitz and N. Narang, "Poor Man's Atomic Bomb? Exploring the Relationship between "Weapons of Mass Destruction"", *Journal of Conflict Resolution*, vol. 58, no. 3, pp. 509–535, 2014.

[152] J. Hunker, B. Hutchinson, and J. Margulies, "Role and challenges for sufficient cyber-attack attribution", *Institute for Information Infrastructure Protection*, no. 2003, 2008. [Online]. Available: http://www.thei3p.org/docs/publications/whitepaper-attribution.pdf.

[153] J. E. C. Hymans, "Theories of Nuclear Proliferation", *The Nonproliferation Review*, vol. 13, no. 3, pp. 455–465, 2006.

[154] International and Red Cross Committee, "A Guide to the Legal Review of New Weapons, Means and Methods of Warfare: Measures to Implement Article 36 of Additional Protocol I of 1977", *International Review of the Red Cross*, vol. 88, no. 864, pp. 931–956, 2007.

[155] V. Isachenkov, *Russia presents info on missile US says violates pact*, 2019. [Online]. Available: `https://apnews.com/d2d41dccab534618a982966c1001110b` (visited on 01/20/2020).

[156] G. R. Iversen, *Bayesian Statistical Inference*. Sage Publications Limited, 1984.

[157] M. Ivezic, *Stuxnet: the father of cyber-kinetic weapons | CSO Online*, 2018. [Online]. Available: `https://www.csoonline.com/article/3250248/cyberwarfare/stuxnet-the-father-of-cyber-kinetic-weapons.html` (visited on 01/10/2019).

[158] S. Jasper, "US Strategic Cyber Deterrence Options", PhD, University of Reading, 2018.

[159] D.-j. Jo and E. Gartzke, "Weapons Proliferation", *Journal of Conflict Resolution*, pp. 167–194, 2009.

[160] S. R. Johnson, G. A. Tomlinson, G. A. Hawker, J. T. Granton, H. A. Grosbein, and B. M. Feldman, "A valid and reliable belief elicitation method for Bayesian priors", *Journal of Clinical Epidemiology*, vol. 63, no. 4, pp. 370–383, 2010.

[161] M. Kaldor, *New and Old Wars: Orginised Violence in a Gobal Era*. Stanford: Stanford University Press, 1999, p. 38.

[162] V. Kannan, *What Really Happened in the Cyber Command Action Against Iran?*, 2019. [Online]. Available: `https://www.lawfareblog.com/what-really-happened-cyber-command-action-against-iran` (visited on 12/12/2019).

[163] C. Kathryn, "Elicitation of prior distributions", in *Bayesian Biostatistics*, D. A. Berry and D. Stangl, Eds., New York, 1996, ch. 4, pp. 141–156.

[164] J. Keegan, *A history of warfare*. London: Pimlico, 1994.

[165] R. L. Keeney and D. von Winterfeldt, "Eliciting Probabilities from Experts in Complex Technical Problems", *IEEE Transactions on Engineering Management*, vol. 38, no. 3, pp. 191–201, 1991.

[166] L. Kello, "The Meaning of the Cyber Revolution", *International Security*, vol. 38, no. 2, pp. 7–40, 2013.

[167] ——, *The Virtual Weapon and International Order*. Yale University Press, 2017, p. 336.

[168] S. Khattak, N. R. Ramay, K. R. Khan, A. A. Syed, and S. A. Khayam, "A Taxonomy of botnet behavior, detection, and defense", *IEEE Communications Surveys and Tutorials*, vol. 16, no. 2, pp. 898–924, 2014.

[169] D. Kimball, *The Intermediate-Range Nuclear Forces (INF) Treaty at a Glance*, 2019. [Online]. Available: `https://www.armscontrol.org/factsheets/INFtreaty` (visited on 01/20/2020).

[170] O. Kipersztok and H. Wang, "Another Look at Sensitivity of Bayesian Networks to Imprecise Probabilities", *AI and Statistics*, 2001.

[171] U. Kjærulff and L. van der Gaag, "Making Sensitivity Analysis Computationally Efficient", *Proceedings of the Sixteenth conference on Uncertainty in artificial intelligence*, pp. 317 –325, 2000.

[172] M. Klare, "Light Weapons Diffusion and Global Violence in the Post-Cold War Era", *Light Weapons and International Security*, pp. 1–40, 1995.

[173] G. D. Koblentz and B. M. Mazanec, "Viral Warfare: The Security Implications of Cyber and Biological Weapons", *Comparative Strategy*, vol. 32, no. 5, pp. 418–434, 2013.

[174] B. M. Kramer, S. M. Kalick, and M. A. Milburn, "Attitudes Toward Nuclear Weapons and Nuclear War : 1945-1982", *Journal of Social Issues*, vol. 39, no. I, pp. 7–24, 1983.

[175] R. C. Kramer and R. J. Michalowski, "War, aggression and state crime", *British Journal of Criminology*, vol. 45, no. 4, pp. 446–469, 2005.

[176] K. Krause, *Small Arms and Light Weapons: Proliferation Processes and Policy Options*. Department of Foreign Affairs and International Trade, 2000.

[177] ——, *Small arms and light weapons: Towards global public policy*. International Peace Academy, 2007.

[178] M. Kroenig, "Importing the Bomb", *Journal of Conflict Resolution*, vol. 53, no. 2, pp. 161–180, 2009.

[179] M. Kwan, K.-P. Chow, F. Law, and P. Lai, "Reasoning About Evidence Using Bayesian Networks", in *Advances in Digital Forensics IV*, Boston, MA: Springer US, 2008, pp. 275–289.

[180] M. Y. K. Kwan, "The Research of Using Bayesian Inferential Network in Digital Forensic Analysis", PhD thesis, University of Hong Kong, 2011.

[181] M. Y. K. Kwan, K. P. Chow, F. Y. W. Law, and P. K. Y. Lai, "Computer Forensics using Bayesian Network : A Case Study", Hong Kong, Tech. Rep., 2007, pp. 1–21. [Online]. Available: `http://i.cs.hku.hk/cisc/forensics/papers/BayesianNetwork.pdf`.

[182] R. Langner, "Stuxnet: Dissecting a Cyberwarfare Weapon", *IEEE Security & Privacy Magazine*, vol. 9, no. 3, pp. 49–51, 2011.

[183] ——, "To Kill a Centrifuge: A Technical Analysis of What Stuxnet's Creators Tried to Achieve", Tech. Rep. November, 2013.

[184] A. Lanoszka, "The INF Treaty: Pulling Out in Time", *Strategic Studies Quarterly*, vol. 13, no. 2, pp. 48–67, 2019.

[185] P.-S. Laplace, *Essai philosophique sur les probabilitiés*. Brussels: Chez H Remy, 1829.

[186] K. B. Laskey, "Sensitivity analysis in Bayesian networks", in *IEEE Transactions on Systems, Man and Cybernetics*, 6, vol. 25, 1995, pp. 901–906.

[187] J. Lauder, *Stuxnet: The real life sci-fi story of 'the world's first digital weapon'*, 2016. [Online]. Available: `https://www.abc.net.au/triplej/programs/hack/the-worlds-first-digital-weapon-stuxnet/7926298` (visited on 01/10/2019).

[188] S. L. Lauritzen, "Propagation of probabilities, means, and variances in mixed graphical association models", *Journal of the American Statistical Association*, vol. 87, no. 420, pp. 1098–1108, 1992.

[189] P. R. Lavoy, "Nuclear Myths and the Causes of Nuclear Proliferation", *Security Studies*, vol. 2, no. 3-4, pp. 192–212, 1993.

[190] J. S. Levy, "Interstate War and Peace", in *Handbook of International Relations*, W. Carlsnaes, T. Risse, and B. A. Simmons, Eds., SAGE Publications, 2013, ch. 23, pp. 581–606.

[191] ——, "War and Peace", in *Handbook of International Relations*, 2010, pp. 350–368.

[192] M. C. Libicki, *Cyberdeterrence and Cyber War*. RAND Corporation, 2009, p. 240.

[193] J. R. Lindsay and L. Kello, "Correspondence: A Cyber Disagreement", *International Security*, vol. 39, no. 2, pp. 181–192, 2014.

[194] D. A. MacKenzie, *Inventing accuracy: A historical sociology of nuclear missile guidance*. MIT press, 1993.

[195] D. Makrushin, *The cost of launching a DDoS attack*, 2017. [Online]. Available: `https://securelist.com/the-cost-of-launching-a-ddos-attack/77784/` (visited on 01/21/2018).

[196] Z. Maoz and E. A. Henderson, "The World Religion Dataset, 1945-2010: Logic, Estimates, and Trends", *International Interactions*, no. 39, pp. 265–291, 2013. [Online]. Available: `http://cow.dss.ucdavis.edu/data-sets/world-religion-data`.

[197] B. Marczak, D. Fifield, J. Scott-railton, R. Deibert, and V. Paxson, "An Analysis of China's 'Great Cannon'", *5th USENIX Workshop on Free and Open Communications on the Internet (FOCI 15)*, 2015.

[198] M Marshall, T Gurr, and K Jaggers, *Polity IV Project: olitical Regime Characteristics and Transitions, 1800-2015*, 2015. [Online]. Available: `http://www.systemicpeace.org/polityproject.html`.

[199] T. Maurer, "A Dose of Realism: The Contestation and Politics of Cyber Norms", *Hague Journal on the Rule of Law*, no. 0123456789, 2019.

[200] ——, *The Case for Cyberwarfare*, 2011. [Online]. Available: `https://foreignpolicy.com/2011/10/19/the-case-for-cyberwarfare/` (visited on 12/12/2019).

[201] M. Maybaum and J. Tolle, "Arms control in cyberspace-architecture for a trust-based implementation framework based on conventional arms control methods", *International Conference on Cyber Conflict, CYCON*, vol. 2016-Augus, pp. 159–173, 2016.

[202] McAfee, *What Is Petya and NotPetya Ransomware?* [Online]. Available: `https://www.mcafee.com/enterprise/en-gb/security-awareness/ransomware/petya.html` (visited on 12/05/2019).

[203] R. R. McCrae and P. T. Costa, "Validation of the Five-Factor Model of Personality Across Instruments and Observers", *Journal of Personality and Social Psychology*, vol. 52, no. 1, pp. 81–90, 1987.

[204] G. McGraw, "Cyber war is inevitable (unless we build security in)", *Journal of strategic studies*, vol. 36, no. 1, pp. 109–119, 2013.

[205] R. McKemmish, *What is Forensic Computing?*, 118. 1999.

[206] J. Mearsheimer, *The Tragedy of Great Power Politics*. Norton, 2001, p. 555.

[207] S. Mele, "Cyber-weapons: Legal and Strategic Aspects v2", *Italian Institute of Strategic Studies Niccolo Machiavelli*, no. June, p. 22, 2013.

[208] S. M. Meyer, *The dynamics of nuclear proliferation*. University of Chicago Press, 1984.

[209] R. A. Miller, H. E. Pople, and J. D. Myers, "Internist-I, an Experimental Computer-Based Diagnostic Consultant for General Internal Medicine", *New England Journal of Medicine*, vol. 307, no. 8, pp. 468–476, 1982.

[210] S. E. Miller and S. D. Sagan, "Nuclear power without nuclear proliferation", *Dædalus*, pp. 1–13, 2009.

[211] A. Montgomery, "Stop Helping Me : When Nuclear Assistance Impedes Nuclear Programs", in *The Nuclear Renaissance and International Security*, A. N. Stulberg, Ed., 2013, ch. 7, pp. 177–202.

[212] A. H. Montgomery, "Ringing in Proliferation", *International Security*, vol. 30, no. 2, pp. 153–187, 2005.

[213] H. J. Morgenthau and K. W. Thompson, *Politics Among*, 7th. McGraw Hill, 1948.

[214] P. Muncaster, *China's Great Cannon Fires on Hong Kong Protesters*, 2019. [Online]. Available: `https://www.infosecurity-magazine.com/news/chinas-great-cannon-fires-on-hong/` (visited on 01/20/2020).

[215] A. H. Murphy and R. L. Winkler, "Reliability of Subjective Probability Forecasts of Precipitation and Temperature", *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, vol. 26, no. 1, pp. 41–47, 1977.

[216] R. Naraine, *Stuxnet attackers used 4 Windows zero-day exploits | ZDNet*, 2010. [Online]. Available: `https://www.zdnet.com/article/stuxnet-attackers-used-4-windows-zero-day-exploits/` (visited on 01/08/2019).

[217] N. Narang, "All together now? Questioning WMDs as a useful analytical unit for understanding chemical and biological weapons proliferation", *Nonproliferation Review*, vol. 22, no. 3-4, pp. 457–468, 2015.

[218] National Research Council, "Technology, Policy, Law, and Ethics Regarding U.S. Acquisition and Use of Cyberattack Capabilities", p. 391, 2009. [Online]. Available: `http://www.nap.edu/catalog.php?record_id=12651`.

[219] NATO, "AJP-3.9 Allied Joint Doctrine for Joint Targeting", Tech. Rep. April, 2016. [Online]. Available: `https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/628215/20160505-nato_targeting_ajp_3_9.pdf`.

[220] ——, *NATO - Topic: NATO and the INF Treaty*, 2019. [Online]. Available: `https://www.nato.int/cps/en/natohq/topics_166100.htm` (visited on 09/11/2019).

[221] Net Politics and Digital and Cyberspace Policy Program, *The Relationship Between the Biological Weapons Convention and Cybersecurity*, 2015. [Online]. Available: `https://www.cfr.org/blog/relationship-between-biological-weapons-convention-and-cybersecurity`.

[222] S.-L. T. Normand, R. G. Frank, and T. G. McGuire, "Using Elicitation Techniques to Estimate the Value of Ambulatory Treatments for Major Depression", *Medical Decision Making*, vol. 22, no. 3, pp. 245–261, 2002.

[223] J. S. Nye, "Deterrence and Dissuasion in Cyberspace", *International Security*, vol. 41, no. 3, pp. 44–71, 2017.

[224] J. S. Nye, "Soft Power", *Foreign Policy*, no. 80, p. 153, 1990.

[225] J. S. Nye Jr., "Cyber Power", *Belfer Center for Science and International Affairs*, no. May, pp. 1–31, 2010.

[226] ——, "Nuclear Lessons for Cybersecurity?", *Strategic Studies Quarterly*, vol. 19, no. 1, pp. 18–38, 2011.

[227] ——, "Preface, What is Power, Diffusion and Cyberpower, Power Transition", in *Future of Power: Its Changing Nature and Use in the Twenty-First Century*, New York: PublicAffairs, 2011, pp. ix–204.

[228] K. G. Olesen, U. Kjaerulff, F. Jensen, F. V. Jensen, B. Falck, S. Andreassen, and S. K. Andersen, "A munin network for the median nerve—A case study on loops", *Applied Artificial Intelligence*, vol. 3, no. 2-3, pp. 385–403, 1989.

[229] R. E. Overill, J. A. M. Silomon, M. Y. K. Kwan, K.-P. Chow, F. Y. W. Law, and P. K. Y. Lai, "Sensitivity Analysis of a Bayesian Network for Reasoning about Digital Forensic Evidence", *2010 3rd International Conference on Human-Centric Computing*, pp. 1–5, 2010.

[230] R. E. Overill, E. P. Zhang, and C. Kam-Pui, "Multi-Parameter Sensitivity Analysis of a Bayesian Network From a Digital Forensic Investigation", *Proceedings of the Conference on Digital Forensics, Security & Law*, pp. 129–140, 2012.

[231] "Technical and Operational Considerations in Cyberattack and Cyberexploitation", in *Technology, Policy, Law, and Ethics Regarding U.S. Acquisition and Use of Cyberattack Capabilities*, W. A. Owens, K. W. Dam, and H. S. Lin, Eds., 2009, pp. 79–158.

[232] P. Paganini, *The thin line between BlackEnergy, DragonFly and TeamSpy attacks*, 2017. [Online]. Available: `https://securityaffairs.co/wordpress/66867/apt/blackenergy-dragonfly-teamspy-attacks.html`.

[233] G. Palmer, V. D'Orazio, M. Kenwick, and M. Lane, "The Mid4 Dataset, 2002–2010: Procedures, Coding Rules and Description", *Conflict Management and Peace Science*, vol. 32, pp. 222–42, 2015. [Online]. Available: `http://cow.dss.ucdavis.edu/data-sets/MIDs`.

[234] J. Pearl, *Causality: Models, Reasoning, and Inference*, 2nd. Cambridge University Press, 2009.

[235] Y. Peng and J. Reggia, "Plausibility of Diagnostic Hypotheses The Nature of Simplicity", *Proceedings of the 5th National Conference on Artificial Intelligence. Volume 1*, pp. 140–147, 1986. arXiv: `0504060 [physics]`.

[236] G. Perkovich and W. Hoffman, "From Cyber Swords to Plowshares", in *Think Peace: Essays for an Age of Disorder*, T. de Waal, Ed., The Carnegie Endowment for International Peace, 2019. [Online]. Available: https://carnegieendowment.org/2019/10/14/from-cyber-swords-to-plowshares-pub-80035.

[237] D. Peterson, "Offensive Cyber Weapons: Construction, Development, and Employment", *Journal of Strategic Studies*, vol. 36, no. 1, pp. 120–124, 2013.

[238] I. R. Porche III, C. Paul, C. C. Serena, C. P. Clarke, E.-e. Johnson, and D. Herrick, *Tactical Cyber - Building a Strategy for Cyber Support to Corps and Below*. 2017.

[239] Press and Communication Directorate, *Cybersecurity: Paris Call of 12 November 2018 for Trust and Security in Cyberspace*, 2019. [Online]. Available: https://www.diplomatie.gouv.fr/en/french-foreign-policy/digital-diplomacy/france-and-cyber-security/article/cybersecurity-paris-call-of-12-november-2018-for-trust-and-security-in.

[240] C. G. J. Putman, "Business Model of Botnets", p. 12, 2017.

[241] V. J. Radunovic, "DDoS - available weapon of mass disruption", *Telecommunications Forum (TELFOR), 2013 21st*, pp. 5–8, 2013.

[242] A. Rapoport, Ed., *Game Theory as a Theory of a Conflict Resolution*. Dordrecht: Springer Netherlands, 1974.

[243] R. Reagan, "National Security Directive 32: U.S. National Security Strategy", Washington, DC, Tech. Rep., 1982, p. 8.

[244] R. Rebonato, *Coherent stress testing: A Bayesian approach to the analysis of financial stress*. John Wiley & Sons, 2010.

[245] M. Reiss, *Without the bomb : the politics of nuclear nonproliferation*. New York: Columbia University Press, 1988.

[246] S. Renooij, "Probability elicitation for belief networks: Issues to consider", *Knowledge Engineering Review*, vol. 16, no. 3, pp. 255–269, 2001.

[247] T. Rid, "Cyber War Will Not Take Place", *Journal of Strategic Studies*, vol. 35, no. 1, pp. 5–32, 2012. arXiv: 1011.1669v3.

[248] T. Rid and B. Buchanan, "Attributing Cyber Attacks", *Journal of Strategic Studies*, vol. 38, no. 1-2, pp. 4–37, 2015.

[249] T. Rid and P. McBurney, "Cyber-Weapons", en, *The RUSI Journal*, vol. 157, no. 1, pp. 6–13, 2012.

[250] F. Rohrbein, J. Eggert, and E. Korner, "Child-Friendly Divorcing: Incremental Hierarchy Learning in Bayesian Networks", in *Proceedings of International Joint Conference on Neural Networks*, 2009, pp. 2711–2716.

[251] S. Romanosky and Z. Goldman, "Cyber Collateral Damage", *Procedia Computer Science*, vol. 95, pp. 10–17, 2016.

[252] J. Rowland, M. Rice, and S. Shenoi, "The anatomy of a cyber power", *International Journal of Critical Infrastructure Protection*, vol. 7, no. 1, pp. 3–11, 2014.

[253] ——, "Whither cyberpower?", *International Journal of Critical Infrastructure Protection*, vol. 7, no. 2, pp. 124–137, 2014.

[254] S. D. Sagan, *The Limits of Safety: Organizations, Accidents, and Nuclear Weapons*. Princeton, New Jersey: Princeton University Press, 1993.

[255] ——, "Why do states build nuclear weapons? Three models in search of a bomb", *International Security*, vol. 21, no. 3, pp. 54–86, 1997.

[256] S. D. Sagan and K. N. Waltz, *The spread of nuclear weapons: A debate*. 1995.

[257] ——, *The Spread of Nuclear Weapons: A Debate Renewed*, 2002.

[258] A. Saltelli, D. Gatelli, F. Campolongo, J. Cariboni, M. Ratto, M. Saisana, S. Tarantola, and T. Andres, *Global Sensitivity Analysis: The Primer*. Wiley, 2008.

[259] D. E. Sanger, *The Perfect Weapon: War, Sabotage, and Fear in the Cyber Age*. New York: Crown, 2018, p. 384.

[260] M. R. Sarkees and F. Wayman, *Resort to War: 1816 - 2007 v4.0*. Washington DC: CQ Press, 2010. [Online]. Available: `http://cow.dss.ucdavis.edu/data-sets/COW-war`.

[261] K. Sasikumar and C. Way, "Testing theories of proliferation in South Asia", in *Inside Nuclear South Asia*, Stanford University Pres, 2009, pp. 68–105.

[262] K. M. Sayler, "Hypersonic Weapons : Background and Issues for Congress Hypersonic Weapons : Background and Issues for Congress", 2019.

[263] T. C. Schelling, *Bargaining, communication, and limited war*. Oxford University Press, 1960.

[264] T. C. Schelling and M. H. Halperin, *Strategy and Arms Control*. New York: Twentieth Century Fund, 1961.

[265] M. Schmitt, "Classification of Cyber Conflict", *Journal of Conflict and Security Law*, vol. 17, no. 2, pp. 245–260, 2012.

[266] M. N. Schmitt, Ed., *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations*. Cambridge: Cambridge University Press, 2017.

[267] M. N. Schmitt, Ed., *Tallinn manual on the international law applicable to cyber warfare*. Cambridge University Press, 2013, p. 282.

[268] B. Schneier, *Could your plane be hacked?*, 2015. [Online]. Available: `https://edition.cnn.com/2015/04/16/opinions/schneier-hacking-airplanes/index.html` (visited on 02/01/2020).

[269] M. Schulze, "Quo Vadis Cyber Arms Control? – A Sketch of an International Vulnerability Equities Process and a 0-Day Emissions Trading Regime", in *Proceedings of the Interdisciplinary Conference on Technical Peace and Security Research*, C. Reuter, J. Altmann, M. Göttsche, and M. Himmel, Eds., Darmstadt, 2019, pp. 21–40.

[270] Seven Security Group, *DDoS Stress Testing*. [Online]. Available: `https://www.7sec.com/testing/ddos-stress-testing/`.

[271] J. Sheldon, "Deciphering Cyberpower: Strategic Purpose in Peace and War", *Strategic Studies Quarterly*, pp. 95–112, 2011.

[272] N. W. Shock, R. C. Gruelich, R. Andres, D. Arenberg, P. T. Costa Jr, E. G. Lakatta, and J. D. Tobin, "Normal human aging: the Baltimore Longitudinal Study Of Aging", *National Institutes of Health*, no. 84-2450, 1984. [Online]. Available: `https://eric.ed.gov/?id=ED292030`.

[273] M. A. Shwe, B Middleton, D. E. Heckerman, M Henrion, E. J. Horvitz, H. P. Lehmann, and G. F. Cooper, "A reformulation of the metal-electrolyte double layer problem", *Methods of Information in Medicine*, vol. 30, no. 4, pp. 241–255, 1991.

[274] Siemens, *Charter of Trust*, 2019. [Online]. Available: `https://new.siemens.com/global/en/company/topic-areas/cybersecurity.html`.

[275] J. A. M. Silomon, "A Bayesian Network Approach to the Proliferation of Software as a Weapon", in *ICCWS 2019 14th International Conference on Cyber Warfare and Security*, Academic Conferences and Publishing Limited, 2019, pp. 377–387.

[276] ——, "Cyberwar vs. Clausewitz: Harm done to and by digital systems", University of Oxford, Oxford, Tech. Rep., 2015.

[277] ——, "Factors Contributing to the Proliferation of Software as a Weapon", in *Proceedings of the 17th European Conference on Cyber Warfare and Security*, Oslo: ACPI, 2018, pp. 471–179.

[278] ——, "Software as a Weapon: Factors Contributing to the Development and Proliferation", *Journal of Information Warfare*, vol. 17, no. 3, pp. 106–123, 2018.

[279] J. A. M. Silomon and M. P. Roeling, "Assessing Opinions on Software as a Weapon in the Context of (Inter)national Security", in *Transactions on Computational Science XXXII*, M. L. Gavrilova, C. J. K. Tan, and A. Sourin, Eds., Springer Berlin Heidelberg, 2018, ch. 4, pp. 43–56.

[280] J. A. M. Silomon and A. W. Roscoe, "Attitudes towards Software as a Weapon", in *IADIS International Conference ICT, Society and Human Beings*, Lisbon, 2017.

[281] ——, "Software and Malware Capabilities: Opinions on (Inter)National Security", in *International Conference on Cyberworlds*, 2017, pp. 96–102.

[282] J. D. Singer, "Reconstructing the correlates of war dataset on material capabilities of states, 1816–1985", *International Interactions*, vol. 14, no. 2, pp. 115–132, 1988.

[283] J. D. Singer, S. Bremer, and J. Stuckey, "Capability Distribution, Uncertainty, and Major Power War, 1820-1965", in *Peace, War, and Numbers*, B. Russett, Ed., Beverly Hills: Sage, 1972, pp. 19–48.

[284] J. D. Singer and M. Small, *The wages of war, 1816-1965: a statistical handbook*. John Wiley & Sons, 1972.

[285] P. W. Singer and A. Friedman, *Cybersecurity: What everyone needs to know*. Oxford University Press, 2014.

[286] S. Singh and C. R. Way, "The correlates of nuclear proliferation: A quantitative test", *Journal of Conflict Resolution*, vol. 48, no. 6, pp. 859–885, 2004.

[287] M. Smeets, *Cyber Conflict and International Relations: Where to get started*, 2019. [Online]. Available: `http://maxsmeets.com/blog/` (visited on 01/20/2020).

[288] ——, "U.S. cyber strategy of persistent engagement & defend forward: implications for the alliance and intelligence collection", *Intelligence and National Security*, vol. 35, no. 3, pp. 444–453, 2020.

[289] E. Solingen, "The Domestic Sources of Nuclear Postures Influencing "Fence-Sitters" in the Post-Cold War Era", *IGCC Policy Papers*, no. 8, 1994.

[290] E. M. Spiers, *A History of Chemical and Biological Weapons*. Cromwell Press Group, 2010.

[291] P. Spirtes, C. Glymour, and R. Scheines, *Causation, Prediction, and Search*, ser. Lecture Notes in Statistics. New York, NY: Springer New York, 1993, vol. 81.

[292] R. E. Stake, *The art of case study research*. Sage Publications Limited, 1995.

[293] T. Starks, *Questions about U.S. cyberattack blowback*, 2019. [Online]. Available: `https://www.politico.com/newsletters/morning-cybersecurity/2019/07/10/questions-about-us-cyberattack-blowback-677108` (visited on 11/25/2019).

[294] T. Stevens, *Cyber Security and the Politics of Time*. Cambridge University Press, 2016, p. 282.

[295] ——, "Cyberweapons: an emerging global governance architecture", *Palgrave Communications*, vol. 3, no. 1, 2017.

[296] P. N. Stockton and M. Golabek-Goldman, "Curbing the Market for Cyber Weapons", *Yale Law and Policy Review*, vol. 32, no. 1, pp. 240–266, 2011.

[297] J. Stone, "Cyber War Will Take Place!", *Journal of Strategic Studies*, vol. 36, no. 1, pp. 101–108, 2013.

[298] H. Strachan and A. Herberg-Rothe, *Clausewitz in the Twenty-first Century*. Oxford University Press, 2007.

[299] H. Strachan and S. Scheipers, *The changing character of war*. Oxford: Oxford University Press, 2011.

[300] A. M. Sukumar, *The UN GGE Failed. Is International Law in Cyberspace Doomed As Well?*, 2017. [Online]. Available: `https://www.lawfareblog.com/un-gge-failed-international-law-cyberspace-doomed-well` (visited on 01/22/2020).

[301] Symantec, "Advanced Persistent Threats : A Symantec Perspective", *Symantec*, p. 12, 2011. [Online]. Available: `http://www.symantec.com/content/en/us/enterprise/white_papers/b-advanced_persistent_threats_WP_21215957.en-us.pdf`.

[302] D. Tarakanov, *Shamoon The Wiper: Further Details*, 2012. [Online]. Available: `https://securelist.com/shamoon-the-wiper-further-details-part-ii/57784/` (visited on 11/29/2018).

[303] ——, *Shamoon the Wiper in details*, 2012. [Online]. Available: `https://securelist.com/shamoon-the-wiper-in-details-40/34369/` (visited on 11/29/2018).

[304] Thales, *Distributed Denial-of-Service Stress Testing Service*, 2019. [Online]. Available: `https://www.thalesgroup.com/en/critical-information-systems-and-cybersecurity/distributed-denial-service-stress-testing-service`.

[305] The Council of Europe, "Convention on Cybercrime", *European Treaty Series - No. 185*, 2001.

[306] The Guardian, *The Cambridge Analytica Files | The Guardian*, 2018. [Online]. Available: https://www.theguardian.com/news/series/cambridge-analytica-files (visited on 01/10/2019).

[307] "The Paris Call for Trust and Security in Cyberspace", 2018. [Online]. Available: https://etno.eu/datas/press_corner/ParisCallText-EN.pdf.

[308] N. Tsagourias, "Cyber attacks, self-defence and the problem of attribution", *Journal of Conflict and Security Law*, vol. 17, no. 2, pp. 229–244, 2012.

[309] E Tyugu, "Command and control of cyber weapons", *Cyber Conflict (CYCON), 2012 4th International Conference on*, pp. 1–11, 2012.

[310] ——, "Situation awareness and control errors of cyber weapons", *2013 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*, pp. 143–148, 2013.

[311] M. Uma and G. Padmavathi, "A survey on various cyber attacks and their classification", *International Journal of Network Security*, vol. 15, no. 5, pp. 390–396, 2013.

[312] UN General Assembly, "A/70/174", Tech. Rep. July, 2015, p. 17. [Online]. Available: http://www.un.org/ga/search/view_doc.asp?symbol=A/70/174.

[313] United States Air Force, "Air Force Instruction 51-402: Legal Reviews of Weapons and Cyber Capabilities", no. July, p. 7, 2011.

[314] University of Groningen, *Maddison Historical Statistics | Historical Development | University of Groningen*, 2018. [Online]. Available: https://www.rug.nl/ggdc/historicaldevelopment/maddison/releases/maddison-project-database-2018 (visited on 01/11/2019).

[315] ——, *The Database | Penn World Table | Productivity | University of Groningen*, 2018. [Online]. Available: https://www.rug.nl/ggdc/productivity/pwt/ (visited on 01/11/2019).

[316] US Senate, "Nominations Before The Senate Armed Services Committee (hearing, 10/03/2010)", pp. 1–731, 2010. [Online]. Available: https://fas.org/irp/congress/2010_hr/alexander.html.

[317] US State Department, "Adherence to and Compliance with Arms Control, Nonproliferation, and Disarmament Agreements and Commitments", Tech. Rep. July, 2014. [Online]. Available: https://2009-2017.state.gov/documents/organization/230108.pdf.

[318] B. Valeriano and R. C. Maness, "The dynamics of cyber conflict between rival antagonists, 2001-11", *Journal of Peace Research*, vol. 51, no. 3, pp. 347–360, 2014.

[319] L. C. Van Der Gaag, S. Renooij, C. L. Witteman, B. M. Aleman, and B. G. Taal, "Probabilities for a probabilistic network: A case study in oesophageal cancer", *Artificial Intelligence in Medicine*, vol. 25, no. 2, pp. 123–148, 2002.

[320] S. Vanheule, M. Desmet, H. Groenvynck, Y. Rosseel, and J. Fontaine, "The factor structure of the Beck Depression Inventory-II: An evaluation", *Assessment*, vol. 15, no. 2, pp. 177–187, 2008.

[321] J. A. Vasquez, *The power of power politics*. Cambridge: Cambridge University Press, 1999, p. 470.

[322] U. von Waldow and F. Roehrbein, "Structure Learning in Bayesian Networks with Parent Divorcing", in *Proceedings of the EuroAsianPacific Joint Conference on Cognitive Science*, 2015, pp. 146–151.

[323] C. D. Wallace, "Cyber Weapon Reviews under International Humanitarian Law: A Critical Analysis", *Tallinn Paper*, p. 11, 2018. [Online]. Available: https://ccdcoe.org/library/publications/cyber-weapon-reviews-under-international-humanitarian-law-a-critical-analysis/.

[324] S. M. Walt, "The Renaissance of Security Studies", *International Studies Quarterly*, vol. 35, no. 2, p. 211, 1991.

[325] S. M. Walt, "Theory and Policy in International Relations: Some Personal Reflections", *Yale Journal of International Affairs*, vol. 7, no. 2, pp. 33–43, 2012.

[326] ——, *What role should IR scholars play in policy-making?*, 2012. [Online]. Available: https://foreignpolicy.com/2012/09/26/what-role-should-ir-scholars-play-in-policy-making/.

[327] K. N. Waltz, *Man, the State and War*. New York: Columbia University Press, 1954.

[328] ——, "The Origins of War in Neorealist Theory", *The Journal of Interdisciplinary History*, vol. 18, no. 4, pp. 615–628, 1988.

[329] ——, *Theory of International Politics*. Reading, Mass: Addison-Wesley, 1979.

[330] Wassenaar.org, *Summary of Changes List of Dual-Use Goods & Technologies and Munitions List*, 2017. [Online]. Available: http://www.wassenaar.org/wp-content/uploads/2017/12/Summary-of-Changes-to-2017-Lists-Website.pdf.

[331] M. C. Waxman, "Cyber-Attacks and the Use of Force: Back to the Future of Article 2 (4)", *Yale Journal of International Law*, vol. 36, pp. 421–459, 2011.

[332] C. Way and J. L. P. Weeks, "Making It Personal: Regime Type and Nuclear Proliferation", *American Journal of Political Science*, vol. 58, no. 3, pp. 705–719, 2014.

[333] D. A. Wheeler and G. N. Larsen, "Techniques for Cyber Attack Attribution", Institute for Defense Analyses, Alexandria, Tech. Rep., 2003, p. 84. [Online]. Available: http://www.gpo.gov/fdsys/pkg/CHRG-111hhrg57603/pdf/CHRG-111hhrg57603.pdf.

[334] P. Whitney and S. Walsh, "Calibrating Bayesian Network Representations of Social-Behavioral Models", in *Advances in Social Computing*, vol. 6007 LNCS, 2010, pp. 98–107. arXiv: 9780201398298.

[335] WHO, *WHOQOL User Manual*, 1998. arXiv: 1011.1669v3. [Online]. Available: http://apps.who.int/iris/bitstream/handle/10665/77932/WHO_HIS_HSI_Rev.2012.03_eng.pdf?sequence=1&isAllowed=y.

[336] C. Whyte, "Power and Predation in Cyberspace", *Strategic Studies Quarterly*, vol. 9, no. 1, pp. 100–118, 2015.

[337] C. Wilson, *Cyber weapons: 4 defining characteristics*, 2015. [Online]. Available: `https://gcn.com/articles/2015/06/04/cyber-weapon.aspx` (visited on 07/14/2016).

[338] D. Winder, *China Fires 'Great Cannon' Cyber-Weapon At The Hong Kong Pro-Democracy Movement*, 2019. [Online]. Available: `https://www.forbes.com/sites/daveywinder/2019/12/05/china-fires-great-cannon-cyber-weapon-at-the-hong-kong-pro-democracy-movement/#2a41c75b7c85` (visited on 01/20/2020).

[339] J. J. Wirtz, *Planning the Unthinkable: how new powers will use nuclear, biological, and chemical weapons*, P. R. Lavoy, S. D. Sagan, and J. J. Wirtz, Eds. Cornell University Press, 2000, pp. 1–15.

[340] L. Wittgenstein, *Philosophische Untersuchungen*, Reprint 20. Frankfurt: Suhrkamp Verlag, 1953.

[341] A. Wolfers, *"National Security as an Ambiguous Symbol." Discord and Collaboration: Essays on International Politics*. The Johns Hopkins Press, 1962, p. 150.

[342] A Zagorecki, "Local Probability Distributions in Bayesian Networks: Knowledge Elicitation and Inference", PhD, University of Pittsburgh, 2010. [Online]. Available: `http://d-scholarship.pitt.edu/6542/`.

[343] A. Zagorecki and M. J. Druzdzel, "Knowledge engineering for bayesian networks: How common are noisy-MAX distributions in practice'", *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans*, vol. 43, no. 1, pp. 186–195, 2013.

[344] J. P. Zanders, "Assessing the risk of chemical and biological weapons proliferation to terrorists", *The Nonproliferation Review*, vol. 6, no. 4, pp. 17–34, 1999.

[345] K. Zetter, *Countdown to Zero Day: Stuxnet and the launch of the world's first digital weapon*. Broadway Books, 2014.

[346] Y. Zhou, N. E. Fenton, and M. Neil, "An Extended MPL-C Model for Bayesian Network Parameter Learning with Exterior Constraints", *Probabilistic Graphical Models: 7th European Workshop. PGM 2014, Utrecht. The Netherlands, September 17-19, 2014*, no. 61273322, pp. 581–596, 2014.