

Sami Alaverronen & Jussi Pohjola

PANDAS IN ACTION

**ANALYSIS OF CHINA RELATED ADVANCED
PERSISTENT THREAT ACTORS'
TACTICS, TECHNIQUES & PROCEDURES**



UNIVERSITY OF JYVÄSKYLÄ
FACULTY OF INFORMATION TECHNOLOGY
2023

ABSTRACT

Alaverronen, Sami

Pohjola, Jussi

Pandas in Action: Analysis of China Related Advanced Persistent Threat Actors' Tactics, Techniques & Procedures

Jyväskylä: University of Jyväskylä, 2023, 130 pp.

Cyber Security, Master's Thesis

Supervisor: Lehto, Martti

For long, the United States has been the single first tier power in cyberspace, but there is a rising contender from the east. As China is trying to advance their reach in cyberspace, China related Advanced Persistent Threat cyber-attacks are growing in numbers. These Advanced Persistent Threat cyber-attacks target both the government and companies alike in order to gain valuable information or perform other desired actions. Cyber security actors can in turn analyse cyber-attacks to gain valuable cyber threat intelligence from different indicators of compromise to used techniques, tactics, and procedures. This information is further refined by categorizing it for example to a form of a taxonomy.

This thesis consisted of an analysis of 41 different cyber security companies' reports that had been attributed to China related Advanced Persistent Threat cyber-attacks and identified different procedures with content analysis. Lockheed Martin's Cyber Kill Chain and MITRE ATT&CK frameworks were used to discover China related Advanced Persistent Threat cyber-attack tactics, techniques, and procedures. The results showed that the Chinese APT cyber-attacks relied first on gathering the victim organizations information, then developed capabilities to attack and delivered the weapon to the target by utilizing phishing, usually spear phishing. Once the weapon was delivered, command and scripting interpreter was utilized to exploit the target system. After the exploitation, the attack continued with installation of a web shell, backdoor or something similar and contacted the command-and-control network utilizing application layer protocols. Finally, the attack was concluded using different remote access tools to exfiltrate data or to expand the attack.

Keywords: China, Advanced Persistent Threat, Cyber Threat Intelligence, Cyber Kill Chain, MITRE ATT&CK Framework, Content Analysis

TIIVISTELMÄ

Alaverronen, Sami

Pohjola, Jussi

Pandat vauhdissa: Kiinaan liitettyjen APT-ryhmien taktiikoiden, tekniikoiden ja toimintamallien analyysi

Jyväskylä: Jyväskylän yliopisto, 2023, 130 s.

Kyberturvallisuus, pro gradu -tutkielma

Ohjaaja: Lehto, Martti

Yhdysvallat on jo pitkään ollut kyberavaruuden johtava valtio, mutta idästä on nousemassa haastaja. Kiinan yrittäessä kasvattaa valtaansa kyberavaruudessa on siihen liitettyjen kohdistettujen haittaohjelmahyökkäysten määrä kasvussa. Nämä kyberhyökkäykset kohdistuvat sekä valtion hallintoon että yrityksiin päämääränään tiedon kerääminen tai muun halutun toimenpiteen suorittaminen. Kyberturvallisuustoimijat voivat puolestaan analysoida kyberhyökkäyksiä saadakseen arvokasta kyberuhkatietoa tunkeutumisindekatooreista sekä käytetyistä tekniikoista, taktiikoista ja toimintamalleista. Näitä tietoja voidaan tarkentaa edelleen esimerkiksi erilaisin luokitteluin.

Tässä tutkimuksessa analysoitiin 41 eri kyberturvallisuusyrityksen raporttia, joissa attribuutio oli määritetty Kiinaan liittyviin kohdistettuihin haittaohjelmahyökkäyksiin ja tunnistettiin erilaisia toimintamalleja sisällönanalyysin avulla. Lockheed Martinin Cyber Kill Chain- ja MITRE ATT&CK-viitekehysiä käytettiin Kiinaan liitettyiden kohdistettujen haittaohjelmahyökkäysten taktiikoiden, tekniikoiden ja toimintamallien havaitsemiseen. Tulokset osoittivat, että kiinalaiset kohdistetut haittaohjelmahyökkäykset suorittivat ensin kohdeorganisaatioiden tietojen keräämisen. Tämän jälkeen kehittivät hyökkäyskyvyn ja toimittivat kyberaseen kohteeseen hyödyntämällä tietojenkalastelua, yleensä keihäskalastelua. Kun kyberase oli toimitettu, komentotulkkia käytettiin kohdejärjestelmän hyödyntämiseen. Hyödyntämisen jälkeen hyökkäys jatkui webshell-takaportin, takaoven tai vastaavan asentamisella, minkä jälkeen yhteys komento- ja kontrolliverkkoon avattiin sovelluskerroksen protokollia käyttäen. Lopuksi hyökkäys vietiin päätökseen käyttämällä erilaisia etäkäyttötyökaluja tietojen varastamiseen tai hyökkäyksen laajentamiseen.

Avainsanat: Kiina, Kohdistettu haittaohjelmahyökkäys, Kyberuhkatiedustelu, Kyberhyökkäyksen estoketju, MITRE ATTA&CK-viitekehys, Sisällönanalyysi

FIGURES

FIGURE 1 Mandiant’s Attack Lifecycle Model (According to Mandiant, 2013, p.27)	27
FIGURE 2 General Cyber-Attack Model (According to Lehto, 2022, p. 126).....	28
FIGURE 3 Lockheed Martin’s Cyber Kill Chain (According to Lockheed Martin, 2023).....	31
FIGURE 4 An Example of MITRE ATT&CK Framework Matrix (According to MITRE ATT&CK, 2023)	38
FIGURE 5 Abstraction Comparison of Models and Threat Knowledge Databases (According to Strom et al., 2020, p. 23)	40
FIGURE 6 The Intelligence Process (According to Joint Publication 2-0, 2013, p. I-6).....	43
FIGURE 7 Relationship of Data, Information and Intelligence (According to Joint Publication 2-0, 2013, p. I-2)	43
FIGURE 8 Thesis’ Topic Fitted in Intelligence Process	45
FIGURE 9 A Comparison of TTP Definition by NIST and MITRE	46
FIGURE 10 Cost of Exposure (According to Markstedter, 2020).....	47
FIGURE 11 The Role and Usage of Taxonomies in Cyber Threat Intelligence...	48
FIGURE 12 "The Pyramid of Pain" (According to Bianco, 2014, p. N/A)	53
FIGURE 13 Extended Diamond Model (According to Caltagirone et al., 2013, p. 19).....	55
FIGURE 14 Pivoting Between the Core Features (According to Caltagirone et al., 2013, p. 27)	56
FIGURE 15 General Cyber-Attack Model (According to Lehto, 2022, p. 126)....	60
FIGURE 16 Chinese Military Hackers (According to Orinx & de Swielande, 2019, p. 62)	68
FIGURE 17 The Content Analysis of the Cyber World (Adapted from Kuusisto & Kuusisto, 2015, p. 37).....	78
FIGURE 18 Theory-Driven Content Analysis of China Related APT Actors’ Tactics, Techniques and Procedures	79
FIGURE 19 TTP Mapping.....	80
FIGURE 20 ThaiCERT Portal	81
FIGURE 21 Result List of Chinese APT Groups.....	82
FIGURE 22 APT Actor’s Threat Group Card.....	83
FIGURE 23 Procedure Categorization.....	85
FIGURE 24 Tactics Summary	85
FIGURE 25 Tactic & Technique Categorization.....	86
FIGURE 26 Technique Categorization.....	86
FIGURE 27 Summary Spreadsheet	87
FIGURE 28 The Difference Between the Most and the Second-most Used Techniques in CKC Phases.....	99
FIGURE 29 Number of Procedures Retrieved per Report Categorized by CKC-Phases	100

FIGURE 30 CKC-Phases Present in the Research Material	101
FIGURE 31 Chinese APT Actors' TTPs	105

TABLES

TABLE 1 Indicators at Each Stage of the Intrusion Kill Chain for APT1 Intrusion Attempts (According to Grooby, Dargahi & Dehghantanha, 2019, p. 236)	12
TABLE 2 Historical Definition of the Term APT	18
TABLE 3 APT Definition Wordplay	21
TABLE 4 Comparison Between Traditional Cyber-attacks and APT-Attacks (Chen et al., 2014, p. 65)	21
TABLE 5 Cyber Attack Models.....	29
TABLE 6 MITRE ATT&CK Framework Use Cases (Strom et al., 2020, pp. 3-4)	39
TABLE 7 Cyber Concepts of the Thesis Part 1	41
TABLE 8 Cyber Concepts of the Thesis Part 2	49
TABLE 9 The Spectrum of State Responsibility Explained (According to Healey, 2011, p. 62)	59
TABLE 10 Typology of Cyber Threat Narratives (According to Lindsay, 2015, p. 12).....	65
TABLE 11 Alleged Chinese Drivers/Motivations for Action in Cyberspace	66
TABLE 12 Details of Chinese APT Operations.....	73
TABLE 13 A Part of the Spectrum of State Responsibility (According to Healey, 2011, p. 62)	73
TABLE 14 Appearances of CKC-phases in the Research Material.....	90
TABLE 15 Ten Most Used Reconnaissance Techniques	91
TABLE 16 Ten Most Used Weaponization Techniques	93
TABLE 17 Ten Most Used Delivery Techniques.....	94
TABLE 18 Ten Most Used Exploitation Techniques	95
TABLE 19 Ten Most Used Installation Techniques	96
TABLE 20 Ten Most Used Command & Control Techniques.....	97
TABLE 21 Ten Most Used Actions on Objectives Techniques.....	99

TABLE OF CONTENTS

ABSTRACT

TIIVISTELMÄ

FIGURES

TABLES

1	INTRODUCTION	8
2	RELATED RESEARCH	11
3	THEORY & CONCEPTS	15
3.1	Cyberspace and Advanced Persistent Threat.....	15
3.1.1	Cyberspace	15
3.1.2	Origins of the Term APT	17
3.1.3	Variety of APT Definitions.....	18
3.2	Cyber-Attack Modelling	24
3.2.1	Cyber-Attack Models.....	25
3.2.2	Lockheed Martin's Cyber Kill Chain.....	31
3.2.3	MITRE's ATT&CK Framework.....	37
3.3	Cyber Threat Intelligence	41
3.3.1	From Traditional Intelligence to Cyber Threat Intelligence.....	42
3.3.2	Cyber-Threat Information.....	45
3.3.3	Cyber Threat Intelligence and the Role of Taxonomies.....	48
4	ATTRIBUTION ISSUE & CHINA AS AN ACTOR IN CYBERSPACE	51
4.1	Cyberspace and the Issue of Attribution.....	51
4.1.1	Performing Attribution	52
4.1.2	Actors of Attribution.....	57
4.1.3	International Law and Cyber Attribution	58
4.2	Drivers and Anomalies of Chinese Cyberspace Activities	60
4.2.1	Domestic Influence.....	61
4.2.2	Mentality and Motivations for Action.....	63
4.2.3	Actors and Operations.....	67
5	RESEARCH METHODOLOGY	75
5.1	Scientific Research Basis	75
5.2	Theoretical Framework.....	77
5.3	Execution of the Study	80
5.3.1	Research Material.....	80
5.3.2	CKC & MITRE ATT&CK Driven Content Analysis.....	84

5.3.3	Reliability and Validity.....	87
6	RESULTS.....	90
6.1	Overall CKC-phases	90
6.1.1	Reconnaissance	91
6.1.2	Weaponization.....	92
6.1.3	Delivery	93
6.1.4	Exploitation	94
6.1.5	Installation.....	95
6.1.6	Command and Control.....	96
6.1.7	Actions on Objective	98
6.2	Additional Findings	99
7	ANALYSIS AND DISCUSSION	102
8	CONCLUSION	106
	REFERENCES	108
	APPENDIX 1 PRIMARY RESEARCH MATERIAL	123
	APPENDIX 2 CHINESE CYBERSPACE ACTORS.....	126
	APPENDIX 3 CHINA’S CYBER ESPIONAGE APPARATUS.....	130

1 INTRODUCTION

The Finnish Parliament's internal IT systems were breached during a cyber-attack in 2020. The National Bureau of Investigation (NBI) and the Finnish Security and Intelligence Service's (SUPO) have officially stated that China affiliated actor APT31 is suspected to be behind the cyber-attack (Cimpanu, 2021). It is stated in SUPO's National Security Overview 2021 report that especially China and Russia are running cyber espionage operations in Finland (SUPO, 2021).

According to Diotte (2020) modern China has the most Internet users in the world, many giant Internet-based companies and thriving business culture which already challenges the US as the hub for technological innovation.¹ During its development process, cyber espionage has become an important part of China's toolset and according to the US estimations, China is responsible for 50-80% of cross border property theft worldwide and the trade secret theft costs for the US alone billions of dollars annually.

The International Institute for Strategic Studies (IISS) (2021) places China as a second tier of cyber powers in its report on Cyber Capabilities and National Power. China sits on the second tier together with countries like Australia, Canada, France, Israel, Russia and the United Kingdom. The US being the sole country in tier one. However, China is evaluated of being the closest to rise to the first tier of cyber powers.

China's ambition is to become a "Cyber Great Power" as the Chinese themselves express it and President Xi has personally devoted himself to make this happen (Doshi, de la Bruyère, Picarsic & Ferguson, 2021). It is somehow logical, that a nation with a population of 1.4 billion inhabitants wants to have its saying both on the world stage and also in cyberspace. Raud (2016) has brought up two essential factors. First, the most inhabitants in the world gives China the upper hand of experts with potential value for government and its cyber operations. Second, China is ruled by the communist party and in the end the party controls all the actions related to cyberspace.

¹ For general China statistics see e.g., <https://www.cia.gov/the-world-factbook/countries/china/>

When talking about cyber-attack and a nation-state in the same sentence, the used term is Advanced Persistent Threat (APT). APT actors, often referred also as groups, have the needed resources and time to execute long-lasting cyber operations. Chinese APT groups are one of the most examined cyberspace adversaries and the Internet is full of information of these groups (see e.g., Lemay, Calvet, Menet & Fernandez, 2018). However, in Finland many of the county-related cyber security studies have concentrated especially on Russia (see e.g., Bunda, 2020; Vatanen, 2020). Therefore, this thesis concentrates on China as an actor in cyberspace and more specifically on cyber-attacks allegedly executed by Chinese APT groups.

The number of cyber-attacks has grown exponentially, and we have probably seen only the tip of the iceberg. This means that the defenders need to understand how their adversaries are operating in order to harden their security posture. A chance to change one's position from reactive defender to proactive actor has been one of the main driving forces for cyber security actors – to be able to act proactively, one needs to understand how the adversaries are operating. Traditional intelligence thinking and methods have also landed on the cyber security field of practice. Often, when referring to intelligence, cyber security practitioners use the term cyber threat intelligence (CTI). In a nutshell, the main goal for intelligence is to generate knowledge for the decision makers. If the goal is to get cyber security matters on a strategic level of decision-making, more descriptive language and report-style briefing are needed.² According to Launius taxonomies might work as a problem solver on how to get the cyber security issues on the table of senior executives (Launius, 2020). Therefore, it is important to have simple, realistic, and comparable models to form a common operational picture.

This thesis continues the work of Bahrami et al. (2019) by utilizing Lockheed Martin's Cyber Kill Chain (CKC) seven phase: Reconnaissance, Weaponization, Delivery, Exploitation, Installation, Command & Control, and Actions on Objectives cyber-attack modelling/taxonomy in APT research. By analysing 41 private cyber security companies' reports consisting of 11 China affiliated APT groups, the authors try to form a picture of the tactics, techniques, and procedures (TTP) Chinese APT groups have used during APT cyber-attacks, and answer the research question of the thesis: What are Chinese APT actors' tactics, techniques, and procedures? The theoretical model of the thesis is based on the CKC framework that is enriched with MITRE ATT&CK framework for more specific analysis of the CKC phases. The CKC model is used to describe the possible phases of cyber-attack from the adversary's perspective. MITRE ATT&CK framework provides technique categorization that allows the identification of used techniques.

The ontological stance of this thesis is constructivism. Thus, reality is seen as something that every person is constructing themselves by the way they see the world and interact with it. Because our thesis concentrates especially on texts, we see that the reality and understanding of things are formulated via text. The

² See chapter 3.3 of this study.

chosen research method is theory-driven content analysis and therefore this research is deductive in nature.

This thesis discovered that the Chinese APT cyber-attacks rely first on gathering the victim organizations information, then developing capabilities to attack and deliver the weapon to the target by utilizing phishing, usually spear phishing. Once the weapon is delivered, it will utilize command and scripting interpreter to exploit the target system. After the exploitation, the attack will continue with installation of a web shell, backdoor or something similar and then the weapon will contact the C2 network utilizing application layer protocols. Finally, the cyber-attack will be concluded using different remote access tools to exfiltrate data or to expand the attack. The analysed material also revealed that in certain CKC phases there is a preferred tactic for China related APT-groups. Delivery phase is usually performed using phishing, server software component is most utilized technique during installation phase and command and control is done with application layer protocol. It seems that these three phases are the main ones, when analysing the attack.

We start our journey in Chapter 2 by examining the scarce research on Chinese APT actors, and how taxonomies have been used in earlier APT studies. In Chapter 3 we dive more deeply into the concepts used in this thesis. We have divided the terms and concepts under three main categories: Cyberspace and Advanced Persistent Threat, Cyber-Attack Modelling, and Cyber Threat Intelligence. The idea is first to describe the environment where cyber threats lurk and then explain how the defenders have modelled adversaries' actions in cyberspace ending up going through the preventive actions in the form of cyber threat intelligence. In Chapter 4, we discuss how cyber threat intelligence is used in possibly one of the hardest subjects in cyberspace, cyber-attack attribution. Even though performing attribution is hard, it can be done to a certain level. Thus, Chapter 4 delves also into the country that has been affiliated to the APT actors investigated in this thesis, China. We examine what China and Chinese APT actors have done in cyberspace and especially take look into the motivations behind these actions. In chapter 5, we go through the chosen scientific research basis, and the unique research methodology of our thesis is presented. Also, reliability and validity of the thesis is discussed. In Chapter 6, the results of the thesis are revealed. Chapter 7 is dedicated to analysis and discussion of the results. The final chapter concludes the thesis with some future research ideas.

Due to the fact that this thesis has had two writers, it is in place to open the writing process. The whole process has been a team effort, and it is hard to name particular chapters made by one of the writers. However, Chapters 1, 3, 4 and 8 have been written mainly together. Pohjola has been in charge of writing Chapters 2 and 5, and Alaverronen Chapters 6 and 7.

2 RELATED RESEARCH

It is a well-known issue that cyber security industry has a monopoly on APT primary source information and academic APT research involving empirical research material is scarce (Lemay et al., 2018).³ This causes certain challenges for researchers. First, the evaluation of available information is hard without knowing how it is gathered and what is possibly left out. Therefore, it is quite natural that researchers in their efforts to gain objectivity have not proceeded with this line of research too eagerly. Second, the APT investigation is time consuming, and the time span of doing academic research can be quite short. Third, because of the attribution issue, it might be even harder to concentrate the examination to a particular nation or instance without losing objectivity. However, there are non-technical (see e.g. Alshamrani, Myneni, Chowdhary & Huang, 2019; Quintero-Bonilla & Martín del Rey, 2020) and more technically oriented research papers related to APTs (see e.g. Chen, Su, Yeh & Yung 2018).⁴ Overall, the non-technical studies tend to focus on describing APT as a phenomenon whereas the more technically oriented research is concentrated on building defensive mechanisms e.g. by showing anomalies from a certain data stream.

The research on China related APT actors is also scarce and most of the available information is produced by private cyber security companies. However, there are a few non-technical studies where the researchers have focused at least partially on Chinese APT actors even though the used research materials have been produced by commercial entities. Chen, Desmet and Huygens (2014) have studied four APT campaigns attributed to China using six-stage cyber-attack model: Recon. and Weaponization, Delivery, Initial Intrusion, Command and Control, Lateral Movement, and Data Exfiltration, to illustrate adversaries' tactics and techniques. The results showed how employee emails, zero-day exploits, backdoors, trojanized docs, remote access tools, weaknesses in a website, and C2 tools were used in the first stage of the attacks. Delivery was done by spear phishing and watering hole attacks. Drive-by download, xls vulnerability, and opening of an executable file were seen as initial intrusions. Command and control were executed with custom C2 protocols based on HTTP protocol and operating on TCP port 443, but also special exploiting tools were used such as Poison Ivy RAT, ZxShell, and Gh0st RAT. Lateral movement was made possible by compromising supply chain management and obtaining source code, performing privilege escalation and gathering Secure ID data, and compromising internal systems and collecting data. Finally, data exfiltration was performed by uploading data to C2 servers, and with RAR files transferred via file transfer protocol (FTP).

³ Our examination is focused on studies that are published mainly in academic journals and conference proceedings and are non-technical in nature.

⁴ The authors recognise that the reviewed research papers in this chapter are not among the Senior Scholars' Basket of Journals in Information Security Management research field.

Ussath, Jaeger, Cheng and Meinel (2016) have analysed 22 separate APT reports with three stage attack model: Initial Compromise, Lateral Movement, and C2. Out of the 22 APT groups or campaigns seven are China affiliated. Spear phishing was the most used initial compromise method and also server attack was used once. Lateral movements were executed using standard OS tools, and hash and password dumping. For C2 channel HTTP(s) protocol was used in addition to custom protocols.

In their study of open-source publications of APT actors and their activities Lemay et al. (2018) have gathered altogether 15 Chinese APT groups' open-source information. They argue that China is the origin of the largest number of APT groups, and it is common for Chinese APT groups to share tools and services among each other. Among the Chinese APT groups spear phishing was the most used initial compromise method. Lateral movement was executed by standard operating system (OS) tools and hash/password dumping. HTTP(s) was the most used protocol to communicate with C2 servers.

Grooby, Dargahi and Dehghantanha (2019) have examined the modus operandi of one Chinese APT group using the Diamond Model and kill chain analysis. The results are presented in table 1.

TABLE 1 Indicators at Each Stage of the Intrusion Kill Chain for APT1 Intrusion Attempts (According to Grooby, Dargahi & Dehghantanha, 2019, p. 236)

Phase	Intrusion Attempt 1	Intrusion Attempt 2	Intrusion Attempt 3
Reconnaissance	Recipient list Stolen email information	Google dorks for web facing industrial control systems Recipient email address	Lure file: "Global A&D outlook 2012.pdf"
Weaponization	N/A	N/A	N/A
Delivery	Email subject Email body Sender email address Sender IP address	Email subject Email body Sender email address Sender IP address CITYREQUEST.doc attachment	Email subject: "2012 Global aerospace and defence industry outlook" Email body Sender email address Sender IP address Global A&D outlook 2012.pdf attachment

(continues)

TABLE 1 (continues)

Exploitation		tomb-keeper@126.com Unreadable text string	Adobe Flash Player CVE-2011-0611 'SWF' File Remote Memory Corruption Vulnerability
Installation	Dropper malware Remote Access Trojan	Ai.exe Gh.exe "tthackfas@#"\$	svchost.exe ntshrui.dll Backdoor.Barkiofork
C2	Exploited middle-man webserver IP Base64 encoded instructions Configuration file headers: [ListenMode] [MServer] [BServer] Etc. Port 80 or 443 connection	C2 server IP address	osamu.update.ikwb.com
Action on Objectives	N/A	N/A	N/A

One thing is common for the above-mentioned China related APT studies; the results are presented in a form of taxonomy based on the attack-model used in each study. Therefore, the chosen model of analysis has a central role and often determines how the results are highlighted for the wider audience. Because the results of our thesis are presented in a form of a taxonomy, it is convenient to review how taxonomies have been used among the cyber security researchers.

In general, cyber security companies, -researchers, -practitioners, -governmental and non-governmental organisations all have used taxonomies to present various cyber security-related information to tackle the malicious actions in cyberspace (see e.g., Cho et al., 2018; Mavroeidis & Bromander, 2017; Derbyshire, Green, Prince, Mauthe & Hutchison, 2018; ENISA, 2018; NIST SP 800-30, 2012; WASC, 2010). Chapman, Leblanc and Partington (2011) have created a taxonomy of cyber-attacks based on the level of access needed on the target system to launch the attack. Virvilis and Gritzalis (2013) taxonomy showed certain characteristics and identified common patterns and techniques of Stuxnet, Duqu, Flame and Red October cyber-attacks. Yadav and Rao (2015) formed taxonomy based on Hutchins, Cloppert and Amin (2011) CKC-framework to present methodologies, techniques, and tools involved in targeted cyber-attacks. Ussath et al., (2016) taxonomy was based on a three-stage attack-model consisting of initial compromise, lateral movement, and command and control channel (C2). The taxonomy of Lemay et al., (2018) presented threat actor, content, and type of open-source publications related to APT actors and their activities.

The use of taxonomies in studies handling cyber-attacks is complex. As Cho et al. have pointed out: "..., there is no standardized and commonly used cyber threat taxonomy." (Cho et al., 2018, p.2). On many occasions, researchers have first created their own taxonomy based on their research interest before research material exploration and analysis. The existing taxonomies, especially among academic field, are narrow in their scope and are not used consistently (Bahrami et al., 2019). The lack of common threat language often distorts both the creation of taxonomies and the presented results (Launius, 2020). When thinking about the use of taxonomies in cyber-attack analysis, the possibility to reuse the selected categorization in other cyber-attack examinations is essential.

Bahrami et al., (2019) have used CKC-based taxonomy for the analysis of 40 APT campaigns. The authors present the idea of using CKC-framework consistently in APT cyber-attack investigation. In their study CKC is utilised to structure complex cyber-attacks into stages and these stages are the taxonomy's categories under which their analysis of 40 APT campaigns are divided. According to the authors, CKC-based taxonomy's value is that it identifies relevant attack characteristics, facilitate incident response, aid in cyber threat hunting and analysis sharing, and help in future security policy creation and mitigation strategy formulation. This thesis follows Bahrami et al.'s CKC-based research bringing in MITRE ATT&CK framework to enrich the analysis.

3 THEORY & CONCEPTS

This chapter defines terms, concepts, and phenomena which can be seen as relevant in examining China related APT actors' cyber-attacks. First, the operating environment that we have ended up calling Cyberspace is explained. Then we move to a cyber threat that is especially advanced, persistent and complex in nature. After we have described the sphere where all the action happens and the threat that lurks in this environment, we go through some countermeasures to tackle these threats proactively: APT attack modelling, and cyber threat intelligence (CTI) in which special attention is given to taxonomies in CTI efforts. All in all, cyber-related concepts are on the move constantly and our examination should be considered as a snapshot and live document that should be updated frequently.⁵

3.1 Cyberspace and Advanced Persistent Threat

According to the Finnish Security Committee (FSC) (2018) the word cyber is often used together with other words to form a compound word with a special meaning. When referring to the word cyber, the meaning is often related to handling of information in a digital form. Merriam-Webster defines cyber as: "of, relating to, or involving computers or computer networks (such as the Internet)" (Merriam-Webster, 2021).⁶

3.1.1 Cyberspace

The definition of cyberspace highlights the comprehensive and complicated nature of the environment where cyber-related issues happen. According to the FSC (2018) cyberspace is an operational environment formed out of one or multiple digital information systems. In cyberspace electronics and electromagnetic spectrum are used with the assistance of communication networks to store, adapt and transfer data and information. The environment consists also the physical structures involved in data and information handling.

Laari, Flyktman, Härmä, Timonen and Tuovinen (2019) in the Finnish National Defence University's manual defines cyberspace as an operational environment build-up of digital information systems including physical structures and all the actors in the operational environment. The definition is further analysed by highlighting the geographically unrestricted nature of the cyberspace, its vulnerable and worldwide structure without an owner, and industrial control

⁵ See e.g., Lehto's (2021) work in defining cyberworld phenomena.

⁶ For more philosophical examination of cyber-related issues and origin of the word in cyber security field of expertise see e.g., Kuusisto & Kuusisto, 2015; Lehto, 2021.

systems being part of it. Physical world and digital world are more and more intertwined, and the physical world is extremely dependent on the digital world - cyberspace being the humanmade digital world.

The US Committee on National Security Systems (CNSS) (2017) has also used the term cyberspace. CNSS defines cyberspace as “The interdependent network of information technology infrastructures, and includes the Internet, telecommunications networks, computer systems, and embedded processors and controllers in critical industries”. There are similarities with the Finnish versions of cyberspace definitions, but infrastructure such as the Internet is defined more concretely, and especially critical infrastructure is mentioned.

Kuusisto and Kuusisto use the term cyber world instead of cyberspace and their cyber world definition is: “The earth with its inhabitants and all things upon it related to or involving computers and computer networks.” (Kuusisto & Kuusisto, 2015, p. 33). The researchers emphasize that it is important to understand the complete and comprehensive system of human existence as part of the cyber world and this understanding enables the examination of human social behaviour supported by information technical solutions (Kuusisto & Kuusisto, 2015).

Lehto’s (2021) definition says that cyberspace merges all networks, databases and information sources to a global virtual system. It is supranational and therefore it causes sovereignty and security disputes between nation-states. Thus, legal and ethical questions play a role in cyberspace definitions.

Even though the US versions define a few of the interdependent network of information systems infrastructures, the Finnish versions are more thorough in explaining the complex matter. The physical aspect is more present in the Finnish definitions and actors are brought up to highlight that cyberspace is not something that is apart from its creators. Because data and information are of essence in cyberspace, the Finns seem to have wanted to emphasize how these important elements are moving in the intertwined networks when referring to electronics and electromagnetic spectrum. When bringing to the table human social behaviour, legal and ethical issues, cyberspace starts to get touching points for us humans to take control of the environment and actions happening in that environment. At the moment, one could say that cyberspace is a wild west without any concrete rules, guidance, and surveillance. Therefore, it is fertile ground for wrongdoers to utilize the opportunities it gives in more and more digitalized world.

In our thesis cyberspace is defined as a humanmade digital system in which data and information is transferred via interconnected networks using electronics and electromagnetic spectrum needing the physical infrastructure, humans, and digitalization for its existence. When examining China related APT actors’ actions in cyberspace, we must understand the complexity, novelty, and irregularity of the environment where all the possible procedures happen. Uncertainty is present all the time, but actions can be analysed by trying to understand e.g., the motives behind actions and learning from historical review. In chapter 4 of our thesis, we dig deeper into this thematic.

Because cyberspace is humanmade, we must assume that everyone does not want to play by the rules. In the cyberspace malicious activity is often even easier than in the physical world due to the lack of norms, rules, and laws. The threats in cyberspace can be physical and non-physical (see e.g., ENISA, 2018). Going through all the possible threats and threat scenarios is out of the scope of our thesis. We put effort in trying to understand one threat called Advanced Persistent Threat or APT because it is in the centre of attention in our thesis.

3.1.2 Origins of the Term APT

Our examination starts by going through the memory lane by gaining understanding of present from the past. Many times, when looking back to history we get answers to modern issues and actions. In our case, trying to find out the origins of the term APT might help us in taking control of a complex issue.

One possible starting point for the use of the term APT can be seen when the United States Air Force (USAF) analysts used it in 2006. Their motive was to make the overall discussion of intrusion activities easier with civilians. Hence, if the analysts would come up with a common term, they could discuss different attack features without revealing classified identities (Binde, McRee & O'Connor, 2011; see also Jeun, Lee & Won, 2012; Xiang, Guo & Li, 2020; Quintero-Bonilla & Martín del Rey, 2020). In his book Cole (2012) gives an idea to which classified identities USAF analysts were referring to when writing that the term was originally developed as a code-name for Chinese cyber-attacks against US military organizations.

However, officially the term seems to appear for the first time in a US patent application filed in 2007 and published in 2008 (Ahmad, Webb, Desouza & Boorman, 2019). In their patent application Schmidt, Rattray and Fogle (2008) are describing a threat the following way (emphasis made by the authors):

The next phase in threat evolution is a more *advanced, persistent threat*. It is characterized by greater sophistication and skill, rapid collaboration, and increasingly structured relationships to overwhelm complex network security mechanisms – oftentimes from the inside. Their motivation is becoming increasingly profit-focused, and their modus operandi includes persistence and stealth. It includes possible *state-sponsored actors* whose effects contribute to long-term influence and exploitation campaigns, as well as devastating effects to facilitate military action. Their signatures include the use of zero-day exploits, distributed agent networks, advanced social engineering techniques such as spear phishing, and long-term data mining and exfiltration. Their flexibility and robust kitbag of tools and techniques makes the advanced threats particularly difficult to successfully defeat with today's technology-heavy network security focus. (p. 3)

Ahmad et al. (2019) have studied APT's origin and evolution of the concept. According to them, even though the patent application mentions "possible state-sponsored actors" the term was already used to depict a threat scenario instead of an attacker. In the early days of the phenomenon, media had taken a stance

that the term APT was originally created to describe state-sponsored actors, especially Chinese ones. By the end of 2010, APT was often referred to as long-lasting hacking campaigns executed by well-funded adversaries against certain targets using complex tactics, techniques, and procedures.

What makes the issues complicated is that inside the information security industry the term has been used to describe different things right from the early days of the term which can be seen in the following definitions: APT is a campaign of intellectual property theft using cyber-methods executed by human beings or organizations (Hoglund, 2009); APT is a group of sophisticated, determined and coordinated attackers (Mandiant, 2010). In table 2 a small summary of the memory lane is presented.

TABLE 2 Historical Definition of the Term APT

Year	Source	Definition
2006	USAF	A common term used for anonymity purposes
2007	US patent application	State-sponsored actors
2009	Hoglund	A campaign of intellectual property theft
2010	Mandiant, Inc.	A group of sophisticated, determined and coordinated attackers
2012	Cole	Code name for Chinese cyber-attacks
2019	Ahmad et al.	A threat scenario

The history shows us that there is not a single starting point for the APT term and phenomenon as such. With the help of this small investigation, we could say that from the early days of APT it has meant both the actual attacker behind cyber-attacks and the cyber-attack itself. As the US patent application shows, APT can be seen as a broad concept. For the general understanding and familiarizing cyberworld phenomena this is quite tricky. For example, it would be easier if one could say that a sophisticated cyber-attacker is called APT, and the actual attack would be called something else. For our purposes Cole's definition is of special interest – if Chinese were behind the cyber-attacks which gave name to a threat that is in the centre of attention in cyber security field of expertise. We are at the kernel of discussion considering APT and it seems that we must investigate more.

3.1.3 Variety of APT Definitions

As mentioned earlier, we are dealing with a term that has many definitions. According to FSC, APT is considered as a multistage cyber-attack (network attack) towards a specific target using malwares and other functions (FSC, 2018). But after the above specification the fun begins. There are plenty of comments to clarify the term and its multifunctional meaning. The following points from one to six are modified from FSC's definitions (FSC, 2018, p. 53):

- 1) APT cyber-attack targets e.g., businesses, different industries, state organizations or limited group of people. The goal is often to gain critical information of a specific target or alter the target's operation.
- 2) The actor(s) behind APT cyber-attack frequently gathers information on their target and use this information to infiltrate malware in the target's systems. The attacker tries to function undetected and the way that all its traces are cleared making attribution difficult.
- 3) APT cyber-attacks are long-lasting, and the malwares used during the attack might be individually designed.
- 4) APT cyber-attack can be both a cyber operation and part of the cyber operation.
- 5) APT cyber-attacks are designed and executed by APT groups. An APT group is an organized group of hackers which are operating on their own or in guidance of a nation-state. APT groups are examined and possibly identified by analysing their tactics, techniques, and procedures.
- 6) APT cyber-attacks are often called campaigns.

FSC's definition of the APT is thorough but also shows the complexity of it. This definition elucidates target, goals, modus operandi, attack description and actors of APT in general. For our efforts, the broad definition has value. The most intriguing parts are the points number five and six. The point five describes our thesis's overall motivation well and only word China is missing from the explanation. If APT is considered as a cyber-attack executed by specific *APT groups*, it makes sense to examine how these groups have executed their cyber-attacks to find some common denominators. Also, we are trying to analyse specifically the tactics, techniques, and procedures of *campaigns* under investigation.⁷

When comparing FSC's definition of APT to its US counterpart's equivalent brings up controversial aspect of still relatively young cyber security field and especially the used language - no common language and understanding of phenomena under scrutiny (see e.g., Launius, 2020, pp. 6-7). CNSS defines advanced persistent threat as an adversary with sophisticated levels of expertise and significant resources with one of the longest phrases we have ever seen (CNSS, 2017). Thus, we are not giving the whole definition. Instead, it is important to mention that CNSS is referring to National Institute of Standards and Technology's (NIST) definition given in Special Publication (SP) 800-39 (see NIST SP 800-39, 2011). The main observation is however that US officials are referring APT to an *adversary* with almost the same features as described in FSC's six points. Now, we have a contradiction because the Finns are saying that APT is a form of a *cyber-attack* albeit also explaining the adversary aspects as well. Without going too deep into

⁷ We want to highlight that from the cyber security student's perspective examining specified ATP groups' tactics, techniques and procedures is educating and throws us in the middle of cyber threat intelligence efforts. Thus, turning the table from reactive cyber-attack victim to an initiative-taking cyber security defender.

the above mentioned thematic, one way to solve the issue is to look deeper into it before making a conclusion i.e., authors' definition of APT in this thesis.

As we already know, the term consists of three words: advanced, persistent and threat. The term can be opened by investigating the meanings given to those words separately. Cole (2012) describes the APT in general as well-funded, organized groups that are systematically compromising government and commercial entities. For him the word *advanced* does not stand for the sophistication of the attack rather the sophistication of the attacker. The *persistent* part refers to the nature of the attack. The attackers are prepared to work the target for a long time. The *threat* part is left out of the examination.

If we think of the term threat in the context of cyberspace, NIST describes it the following way:

A threat is any circumstance or event with potential to adversely impact organizational operations and assets, individuals, other organizations, or the Nation through an information system via unauthorized access, destruction, disclosure, or modification of information, and/or denial of service. (NIST, 2012, p. 8)

According to Jeun et al. (2012) *advanced* means that the adversary is familiar with computer intrusion techniques and tools and can develop custom exploits. *Persistent* is the adversary's intention to fulfil a mission and *threat* refers to the motivated, organized, and funded nature of the adversary. Alshamrani et al. (2019) states that *advanced* means the APT actors have necessary resources to access advanced tools and methods required to perform an APT attack. The attackers are highly determined and *persistent* i.e., they do not give up and once they get into the system, they try to stay there as long as they can. The *threat* part in APT attack is often loss of sensitive data or impediment of critical components or mission. For Tankard (2011) the word *advanced* means both the nature of the *threats* and the exploits used by the hackers. *Persistent* refers to APT's goal which is to stay on the targeted system to gain long-term control and enable data collection. Thus, what does the wordplay tell us?

The following table gathers information acquired from these few studies above (see table 3).

TABLE 3 APT Definition Wordplay

Source	Advanced	Persistent	Threat
Tankard (2011)	Nature of threats and exploits	Ability to stay stealth	n/a
Cole (2012)	Sophistication of the attacker	Nature of the attack	n/a
Jeun et al. (2012)	Sophistication of the attacker	Intention to fulfil a mission	Nature of the adversary
Alshamrani et al. (2019)	Possession of necessary resources	Nature of the attacker	Loss of data or impediment of components or mission

Actually, this examination is quite revealing because we can observe that the contradiction if APT is an actor or a cyber-attack is hidden in the inspected three words – the answer is that APT is both. The actor and attack are mingled together. When examining APTs the researcher has to take this into consideration. Therefore, we are one step closer to our APT definition when stating that it has to take into consideration both the attacker and the attack. Next, we will take a closer look at how researchers, nation-states and cyber security industry define APT.

The academia examination starts with a comparison between APT and other types of cyber-attacks used as a method to clarify APT. The characteristics that distinguish APT from more traditional cyber-attacks are specific targets and clear objectives; well-resourced and organized attackers; a long-lasting campaign with repeated attempts; and stealthy and deceitful attack techniques (Chen et al., 2014; see table 4). Concisely one could say that APT cyber-attack takes time, is well-planned and has often new, unseen features enabling stealthy action despite of preventive cyber security measures.

TABLE 4 Comparison Between Traditional Cyber-attacks and APT-Attacks (Chen et al., 2014, p. 65)

	Traditional Attacks	APT Attacks
Attacker	Mostly single person	Highly organized, sophisticated, determined and well-resourced group
Target	Unspecific, mostly individual systems	Specific organizations, governmental institutions, commercial enterprises
Purpose	Financial benefit, demonstrating abilities	Competitive advantages, strategic benefits
Approach	Single-run, "smash and grab", short period	Repeated attempts, stays low and slow, adapts to resist defences, long term

According to Alshamrani et al. if there could have been more than one way to prevent the attack, the attack did not require a lot of adaptation by attackers, and the attack did not involve any new techniques in its variants, then the attack is targeted attack, not an APT attack (Alshamrani et al., 2019). In other words, targeted cyber-attack is not considered as APT attack even though the target could be specific. More weight is put on the lack of prevention possibilities, adaptive nature of the attacker(s) and novelty of attack techniques. Nevertheless, APTs are also seen as targeted attacks and targeted attacks, unlike worms and viruses, involve intelligent planning. The attitude of the attacker makes a difference when targeting specific individuals or groups (Sood & Enbody, 2012). Hence, APT is seen as something that does not work with trial-and-error tactics. Intelligence is based on well-defined process as we will see later when cyber threat intelligence is examined. The attitude can be different e.g., when someone gets a salary by executing cyber-attacks even though the attack fails compared to someone who is not liable to anyone. APT is stealthy, targeted and data focused, and these features differentiates it from a traditional threat (Cole, 2012).

Making difference between APT and a campaign has been used to explain APT. This comparison is promising and according to Quintero-Bonilla and Martin del Rey (2020) APT is a selective attack to gain access to confidential data or cause damage to a company, industry, or government organization. Campaigns are considered as the customized techniques, methods, and actions an actor uses against a target to execute an APT attack. The above definition makes a difference between the attack (APT) and the functions used in the attacks (campaign). This is similar kind of thinking to which the authors have already ended when stating that the APT definition should take into consideration both: the attacker and the attack. Although the above description makes a difference between the attack and the techniques used in the attack. Therefore, we end up stating that there are three things to consider when formulating APT definition: the attacker, the attack, and the attack techniques.

Ahmad et al. present their definition for the utility of both researchers and practitioners: "An entity that engages in a malicious, organized, and highly sophisticated long-term or reiterated network intrusion and exploitation operation to obtain information from a target organization, sabotage its operations, or both." (Ahmad et al., 2019, p. 406). The use of word entity might refer to an actor and the description highlights the actor's activities and goals. Thus, also this definition intertwines attacker and attack together. Finally, we end up with an explanation based on the most relevant characteristics for APTs according to Ussath, Jaeger, Cheng and Meinel (2016):

- Specific targets
- Use of high-end tactics, techniques, and procedures
- Constantly evolving attack steps; mainly network exfiltration
- Repeating attack attempts
- Maintain long-term presence in the target environment

The characteristics are already familiar from the previous examinations and again the features of the attacker and the attack are listed without making a clear difference between them in APT definition. As a summary, one could say that researchers do not have a clear definition for the meaning of advanced persistent threat. APT is considered either the attacker or the attack based on the researchers' research goals. Nevertheless, the academia examination brought up the idea to include attack techniques to our definition. What are cyber security companies saying about the topic?

Russian cyber security company Kaspersky (2023a) defines APT as continuous, clandestine, and sophisticated hacking techniques used to gain access to a system and the ability to remain inside the system with destructive consequences. The US equivalent Mandiant, Inc. (2010) states that APT is a group of attackers who are sophisticated, determined, and coordinated. The line is also drawn between hackers and APT. APT's motivation, techniques and tenacity are different. They are more professional and successful in their intrusion attempts. Mandiant clearly wants to name advanced threat actors as APT based on their experience in investigating computer security breaches (Mandiant, 2013). It is interesting that even the cyber security industry does not have a common definition of APT despite many similarities between different cyber security companies. As an US company, Mandiant is stating that APT is an actor like US officials. Maybe all comes down to the USAF's analysts' early definition and the need to have a common denominator for a threat without revealing adversary's identity. For the Russian company Kaspersky, it might be more convenient to define APT as an attack instead of attacker because APT actors have been often attributed to Russia albeit mainly by US cyber security companies (see e.g., Mandiant, 2023; MITRE, 2023b).

Nowadays, many nation-states have their own cyber security strategy. It can be stated that cyber security strategies reveal the countries' cyber security maturity and vision. After reviewing the cyber security strategies of the United States, Canada, Australia, United Kingdom and New Zealand; the countries belonging to Five Eyes intelligence alliance (National Cyber Security Strategy, 2018; National Cyber Security Strategy 2016-2021; Australia's Cyber Security Strategy 2020; New Zealand's cyber security strategy 2019) and Finland's strategy (Finland's Cyber Security Strategy, 2019) respectively, we can say that nation states do not mention or define APT in any way. Thus, it is difficult to evaluate how states see APTs officially. Thus, nation-states do not give us any help in our definition efforts. The caution in using words and defining unfamiliar issue for wider audience is understandable. On the other hand, because physical world and cyberspace are more and more mixed together, would it be beneficial to raise cyberspace phenomena more on the nation-states' official rhetoric now and in the future especially?

In general, the definitions of APT vary. APT is seen, among others, as a campaign, groups of attackers, cyber-attack, part of a cyber operation, targeted attack and the list goes on. As a simple summary of our examination, one could say that the information security industry considers APT either an actor or a

campaign/operation, academia defines the term based on their research goals and nation states are too afraid to give any meaning to it. FSC has tackled the difficulty by simply stating that APT is a multistage cyber-attack (network attack) towards a specific target using malwares and other functions; and after the definition numerous clarifications are given. Those clarifications are all mentioned in our academia, industry, and nation-state examinations. Hence, we end up defining APT as: a cyber-attack consisting of multiple separate attack phases. APT is executed by an actor(s) which have the resources and skills to execute long-lasting and stealth cyber-attack operations. The tactics, techniques and procedures used in APT are often sophisticated and custom-made.

Due to commercialization of cyber security and broader media coverage, APT term has started to have life of its own pushing it away from the original definitions. Quite often APT is considered automatically as a nation-state's cyber-attack or if a cyber-attack is in some parts more sophisticated than most of the everyday attacks in cyberspace it is considered as APT. As Markstedter (2020) has brought up, it is important to remember that not only nation-states can be interested in political documents because of the monetary value of e.g., political, and military secrets. There are other institutions with the same interests as nation-states for example financial speculators, political analysts, private consulting, and political opposition. The focus on specific type of data does not indicate that the adversary is absolutely a nation-state.

As the examination in this chapter has also brought up, the lack of mutual understanding, language and definitions makes the APT observation even more complicated. Next, we will move our review from APT definition to more specific examination. We will have a closer look on how APT cyber-attacks have been modelled as part of countermeasures to tackle APT cyber-attacks. At the same time, the modelling functions as a tutor for person trying to learn and understand what has happened when APT cyber-attack has occurred.

3.2 Cyber-Attack Modelling

This chapter investigates the anatomy of an APT cyber-attack by studying how different instances have modelled APTs and what we can learn from these models. We start by first looking into different APT cyber-attack models to get an overview of attack modelling. As attack modelling is quite often described with consecutive phases of the attack, we start our examination from the least amount of cyber-attack phases onwards in numerical order. Because we have chosen to use deductive content analysis method in our thesis, our predetermined theory relies especially on two cyber-attack models and therefore special attention is given to Lockheed Martin's Cyber Kill Chain and MITRE's ATT&CK framework models.

3.2.1 Cyber-Attack Models

Different models have been created to describe cyber-attacks (see e.g., Lehto, 2022). Common to these models is that they consist of consecutive cyber-attack phases. The phases illustrate different aspects of an attack. According to Lehto (2022), the motives and objectives of various cyber-attacks are the building blocks for different attack models. In addition, it is important to separate a traditional cyber-attack from APT attack because APT attack can be seen more sophisticated with a high degree of customization involved. Sophisticated cyber-attacks are often described with a lifecycle model.

Quintero-Bonilla and Martin del Rey (2020) notes that during the last few years researchers have modelled APTs with lifecycles organized in phases. The phases identify techniques, methods, and tools which the attackers are using when executing a targeted intrusion. There are different approaches which determine the number of phases in an attack.

Ussath et al. (2016) have analysed different methods and techniques of APTs and ended up in three stage attack model. The model consists of initial compromise, lateral movement, and command and control stages. In the initial compromise stage attackers try to get access to the target system. They use techniques such as spear-phishing, watering-hole, server-side attacks, and infected storage media⁸. The lateral movement refers to an attempt to compromise services on the target system or network⁹. The goal is to get access to or change credentials to persist in the system. The used techniques can be e.g., standard operating system tools and vulnerability exploitation. When the system has been compromised command and control steps in to exfiltrate data via external connection. HTTP, HTTPS or FTP services are used in addition to tools like virtual network computing (VNC) or remote desktop protocol (RDP).

Zhang, Huo, Liu and Weng (2017) four stage attack model specifies the behaviours and purposes of APT. The four stages are: information collection, intrusion, latent expansion, and information theft. In information collection stage

⁸ *Phishing* is the act when someone tries to steal your personal data e.g., credentials using ethical and technical artifices for malicious practices (see e.g., Pande & Voditel, 2017). An example clarifies the difference between phishing and spear phishing: wrongdoer sends large number of hoax emails to trick people to write their credentials is phishing but if the same procedure is done to a single or well-determined receivers using time and effort to formulate the email, it is *spear-phishing*. *Watering-hole* is a cyber-attack technique in which the attacker infects website(s) with a malware. The goal is to trick website visitors e.g., to click a malicious link or banner and infect their computers. The odd name comes from wildlife where predators wait for their prey near watering holes (see e.g., Allen et al., 2020). *Server-side attacks* are executed directly from the attacker to a listening service. These attacks do not require user interaction. For example, in a server-side request forgery (SSRF) attack, the attacker can supply and modify URL which the target's server will read or submit data to. If succeeded, the attacker can read server configuration, connect to internal services, or perform post requests (see e.g., OWASP, 2021). *Infected storage media* is e.g., a device plugged into the computer's USB port which contains a malware (see e.g., Walters, 2012).

⁹ Lateral movement is a cyber-attack phase which is often executed after the attacker has gained access to the target's network (see e.g., AI Amin, Shetty, Njilla, Tosh & Kamhoua, 2021).

scanning or social engineering tools are used for recognition of the network. During intrusion stage spear-phishing techniques, malicious email attachments or backdoors are used to acquire access privileges¹⁰. In latent expansion stage the attacker tries to maintain control and acquire data enabling attack expansion in the target network. Finally, in information stage the attacker establishes a connection to a server and transfers the stolen data. In many cases different encryption techniques are used to hide the extracted data.

Sexton, Storlie and Neil (2015) calls their five-stage attack model as “Attack Chain”. In delivery stage spear-phishing emails are send to recipients within the network. The exploitation stage means that vulnerabilities of the services, system or applications are exploited. The installation stage refers to a possibility to install malware e.g., remote access tool (RAT) to the target’s system¹¹. In the command-and-control stage the attacker has remote access to a compromised host or server. The last stage is called “actions”, and it refers to the actions made to widen access to other hosts and servers on the same network to exfiltrate confidential information.

Ghafir and Prenosil (2016) use six phase lifecycle model. First phase is information gathering and the goal of this phase is to gather information via public social network profiles to find out the structure of an organization. Second is point of entry in which social engineering, spear-phishing and zero-day exploits are used to gain access to the target environment. Third is command-and-control server. To maintain the connection to the target, a connection is established from the compromised host to command-and-control server e.g., using secure socket layer (SSL) encryption¹². Fourth phase lateral movement is described as the attacker’s actions to find a vulnerable host inside the network. The fifth phase is data of interest which is the identification of critical information on hosts or servers. The sixth phase is external service where the data is transferred to the attacker’s command and control servers.

Vukalovic and Delija (2015) present a seven stage APT attack approach. First stage is called “research”, and it means that the attackers are seeking publicly available information about the target. Second stage is “preparation” which refers to an initial exploitation for custom exploits creation. Third stage is “intrusion”, and it means the first attack made usually with spear-phishing. Fourth stage is “conquering” the network and it means that at least one host is compromised, and this host is used to install RATs or backdoors to control the system. In the fifth, “hiding the presence” stage an attacker stays hidden for a long time. Sixth stage is “data gathering” in which an attacker searches the data of interest and camouflages the data among legitimate traffic and extracts it slowly. Finally,

¹⁰ CNSS defines backdoor as an undocumented way of gaining access to computer system (CNSS, 2017).

¹¹ Remote access tool is a software used to remotely access or control a computer (see e.g., Mazerik, 2014).

¹² SSL is a method created for secure data transfer between a customer and a server over a web. SSL utilizes public key encryption and digital certificates or IDs created by certificate authorities to establish secure, encoded data transfer (see e.g., Shaikh, Bhat & Moharir (2017)).

the seventh stage is “maintaining the access” where the attacker extends its access to the network with modified or new exploits, RATs, and command and control servers.

The Lockheed Martin’s Cyber Kill Chain (CKC) is comprised of seven different attack stages: Reconnaissance, Weaponization, Delivery, Exploitation, Installation, Command & Control, Actions on Objectives. More detailed overview of the model is presented in chapter 3.2.2.

Mandiant (2013) has been the architect of the eight-stage model (see Figure 1). The stages are initial recon which means initial recognition of the target, initial compromise that refers to the methods used for the first intrusion such as spear-phishing, establish foothold with command and control servers from outside the target network, escalate privileges is seeking for credentials that permit wider access, internal recon stage refers information collection the attacker executes, move laterally means that the attacker can connect and share resources using legitimate credentials, maintain presence means that the attacker stays undetected by performing necessary actions, and in complete mission stage the data of interest is compressed and send to the command and control servers.

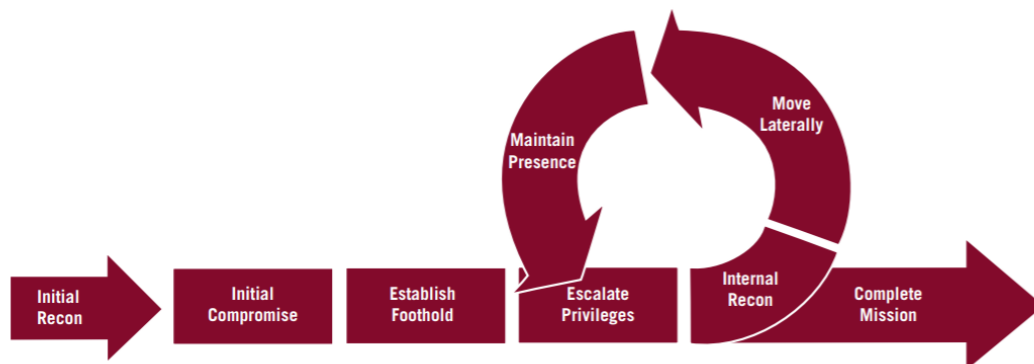


FIGURE 1 Mandiant’s Attack Lifecycle Model (According to Mandiant, 2013, p.27)

Lehto (2022) has ended up creating a general cyber-attack model which has eight different stages after reviewing five different attack models: MITRE ATT&CK, Mandiant Attack Lifecycle Model, Lockheed Martin’s CKC, Unified Kill Chain and Hybrid Cyber Kill Chain (see Figure 2). To the traditional echelon of attack stages, Lehto has added actions before and in the end of cyber-attack lifecycle by raising up two novel stages: Strategic Decision-Making and End State. In addition, the whole attack model is divided into three high level categories Early-Attack Phase comprising from Strategic Decision-Making, Pre-Attack Phase including Reconnaissance and Weaponization and actual Attack Phase comprised of Access, Lateral Movement, Command and Control, Execution and End State.

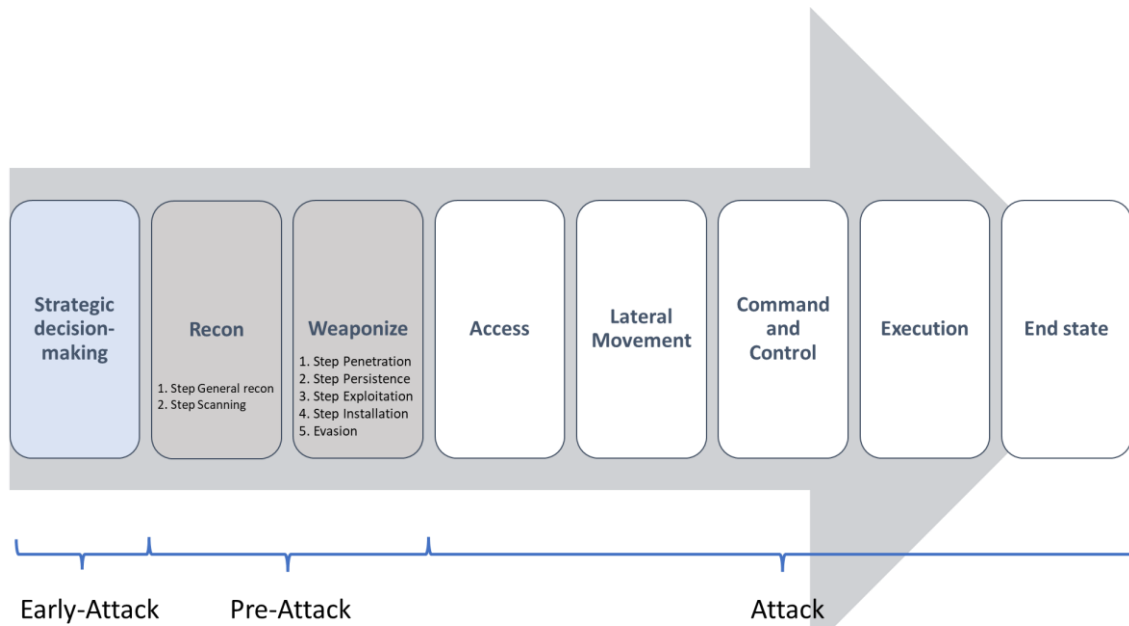


FIGURE 2 General Cyber-Attack Model (According to Lehto, 2022, p. 126)

Zeng and Germanos (2019) created a Hybrid Cyber Kill Chain (HCKC) framework based on the critique of LM's CKC model. According to Zeng and Germanos referring also to other CKC efficiency studies, CKC model lacks the ability to tackle insider threats and is too focused on the perimeter and malware prevention. Therefore, their framework consists of two additional stages to CKC: Persistence and Lateral Movement. Altogether the model has nine stages Reconnaissance, Weaponization, Delivery, Exploitation, Installation, Command & Control and Action on Objectives from the CKC model and the above-mentioned Persistence and Lateral Movement respectively. Persistence expresses the attacker's persistence to complete its mission and Lateral Movement denotes attacker's actions to e.g., escalate privileges once the attacker has already gained access to target network.

In the eleven-stage attack model (Swisscom, 2019), different stages of an attack analyse the stages a threat actor works through to gain the strategic goal: Initial Access: First contact with the target to search for patient zero, Persistence: Efforts to maintain long-lasting access in the target, Privilege Escalation: An attacker needs privileges to install malware or gain access to confidential data, Discovery: Find relevant information from the target e.g. system location or usernames, Lateral Movement: Attacker's movement within the network to search for services of important information, Collection: Collect relevant information, Exfiltration: Extract the collected data, Execution: Execute malware through remote connections established between phases one and five, Defence Evasion: The attacker stays undetected, Credential Access: Accessing the target system with valid credentials, Command and Control: create communication channel between the attacker's servers and target systems (pp. 22–23.)

The MITRE corporation's ATT&CK framework consists of 14 stages: Reconnaissance, Weaponization, Delivery, Social Engineering, Exploitation, Persistence, Defence Evasion, Command & Control, Pivoting, Discovery, Privilege

Escalation, Execution, Credential Access, Lateral Movement, Collection, Exfiltration, Impact, Objectives. The model is presented more closely in chapter 3.2.3.

Pols (2017) has created the Unified Kill Chain model which combines several other frameworks and models such as CKC and ATT&CK. The model is divided into 18 stages: Reconnaissance, Weaponization, Delivery, Social Engineering, Exploitation, Persistence, Defence Evasion, Command & Control, Pivoting, Discovery, Privilege Escalation, Execution, Credential Access, Lateral Movement, Collection, Exfiltration, Impact, Objectives. The 18 stages are grouped into three overarching areas: Initial Foothold as the first, Network Propagation the second and Actions on Objectives the third.

APT is a complicated phenomenon and quite often practitioners' topic under investigation affects the way APTs are examined and how many attack phases are named and examined more closely (see table 5). The number of defined phases highlights how precisely a cyber-attack has been decided to investigate. All the above models have similarities: get the necessary information to gain access to the target's systems, escalate privileges, establish remote connection, stay undetected, exfiltrate or destroy information. Overall, Lehto's two novel attack phases: Strategic Decision-Making and End State opens new doors for discussion of the motives behind an attack and puts more attention how the attackers have terminated their attacks and what the defenders can learn from the end phase of an attack.

TABLE 5 Cyber Attack Models

Model Name	Source	Phases #	Phase Names
N/A	Ussath et al. (2016)	3	Initial Compromise, Lateral Movement, Command and Control
N/A	Zhang et al. (2017)	4	Information Collection, Intrusion, Latent Expansion, Information Theft
Attack Chain	Sexton et al. (2015)	5	Delivery, Exploitation, Installation, Command and Control, Actions
N/A	Ghafir & Prenosil (2016)	6	Information Gathering, Point of Entry, Command and Control Server, Lateral Movement, Data of Interest, External Service
N/A	Vukalovic & Delija (2015)	7	Research, Preparation, Intrusion, Conquering, Hiding the Presence, Data Gathering, Maintaining the Access
Cyber Kill Chain (CKC)	Hutchins et al. (2011)	7	Reconnaissance, Weaponization, Delivery, Exploitation, Installation, Command & Control, Actions on Objectives
Attack Lifecycle Model	Mandiant (2013)	8	Initial Recon, Initial Compromise, Establish Foothold, Escalate Privileges, Internal Recon, Move Laterally, Maintain Presence, Complete Mission

(continues)

TABLE 5 (continues)

General APT Cyber-Attack Model	Lehto (2022)	8	Strategic Decision-Making, Reconnaissance, Weaponize, Access, Lateral Movement, Command and Control, Execution, End State
Hybrid Cyber Kill Chain (HCKC)	Zeng & Germanos (2019)	9	Reconnaissance, Weaponization, Delivery, Persistence, Exploitation, Installation, Command & Control, Lateral Movement, Action on Objectives
N/A	Swisscom (2019)	11	Initial Access, Persistence, Privilege Escalation, Discovery, Lateral Movement, Collection, Exfiltration, Execution, Defence Evasion, Credential Access, Command and Control
MITRE ATT&CK	MITRE Corporation (2023)	14	Reconnaissance, Resource Development, Initial Access, Execution, Persistence, Privilege Escalation, Defence Evasion, Credential Access, Discovery, Lateral Movement, Collection, Command and Control, Exfiltration, Impact
Unified Kill Chain (UCK)	Pols (2017)	18	Reconnaissance, Weaponization, Delivery, Social Engineering, Exploitation, Persistence, Defence Evasion, Command & Control, Pivoting, Discovery, Privilege Escalation, Execution, Credential Access, Lateral Movement, Collection, Exfiltration, Impact, Objectives

APT modelling is not something that is discovered and left aside. It is a developing process as our examination above shows. For example, Mandiant, Inc. changed its APT lifecycle model after it gained more experience and knowledge of APT examination (see Mandiant, 2010; Mandiant, 2013). The APT cyber-attack modelling also reveals the specific features of APT cyber-attacks and can be seen informative for the attacker and the target as such – when an APT study is published the attacker can acquire information on the knowledge-level of the defenders. On the other hand, the targets can share revealed information and reinforce their defensive posture e.g., in the form of taxonomies (some more known than others). Certain attack models are used more often than others, but researchers or cyber security industry do not have a standard model to follow. As mentioned earlier, the fragmented cyber security field and numerous cyber-attack tracking models can be sometimes confusing combined with the lack of common language.

Two cyber-attack frameworks have been often used among cyber security practitioners and researchers: MITRE corporation’s ATT&CK and Lockheed Martin’s Cyber Kill Chain (CKC) (see MITRE, 2023a; Lockheed Martin, 2023). Next, these two cyber-attack models/frameworks are presented more thoroughly due to the reason that the methodological part of our thesis is based on the combination of these two models.

3.2.2 Lockheed Martin's Cyber Kill Chain

Lockheed Martin's Cyber Kill Chain framework is widely used method to describe cyber-attack phases and it is used in various ways within the cyber security research field (see e.g., Bahrami et al., 2019; Cho et al., 2018; Mohsin & Anwar, 2016; Assante & Lee, 2015; van der Watt & Slay, 2021). Lockheed Martin's analysts Hutchins, Cloppert and Amin (2011) introduced CKC to tackle the growing risks of APTs - instead of acting as a passive and reactive intrusion defender, more proactive measures were needed. Hutchins et al. saw that with more developed analysis process based on the knowledge of the threats, it would be possible to anticipate and mitigate future intrusions.

The logic of CKC is taken from US military and it is based on a chain of sequentially executed actions. If one the action is disturbed the entire process will be interrupted. Thus, CKC brings this logic to cyberspace by defining different actions in the chain with actions possibly needed in network intrusion operations. These actions are Reconnaissance, Weaponization, Delivery, Exploitation, Installation, Command & Control and Actions on Objective (see Figure 3). Next, the APT cyber-attack phases are described more specifically with illustrative examples.

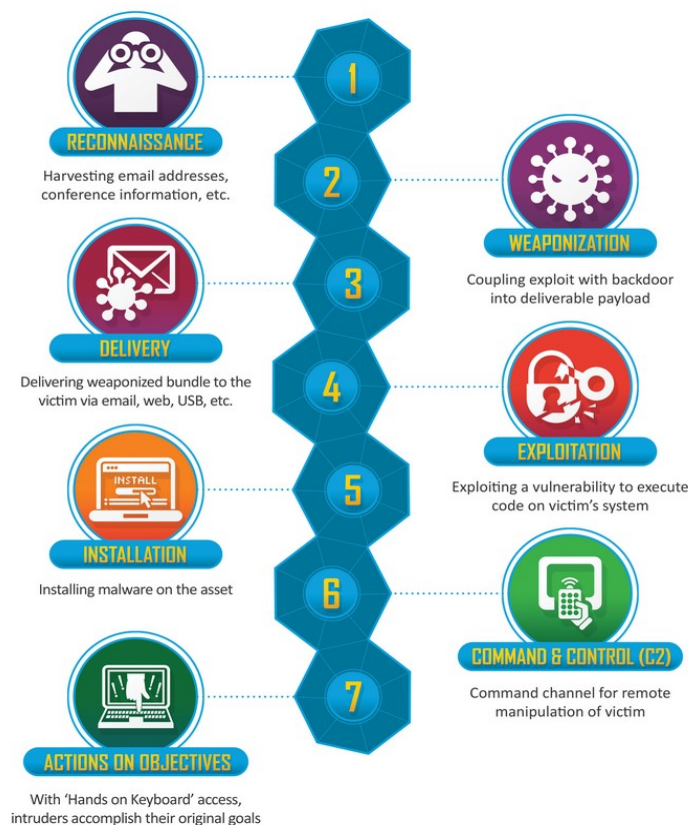


FIGURE 3 Lockheed Martin's Cyber Kill Chain (According to Lockheed Martin, 2023)

Reconnaissance is the intelligence phase of the cyber-attack. Information is gathered from multiple sources, with different techniques and possibly during a long

period of time. The idea is to gather data as much as possible from the target. The data can be anything from the people working in the target all the way to the equipment, which is used on the target. Generally, the APT attackers research, or reconnaissance, their target to maximize the impact and to minimize the risk of exposure (Bahrami et al., 2019). There are both passive and active reconnaissance, that the attackers can carry out against the target and its network. The difference between active and passive reconnaissance can be described that in passive, the aim is to gather as much data without encountering the target's network, people or assets. In active, there is no limitation on what can be done. Bhatnagar, Som and Khatri (2019) argue that in general, APT attacks involve active reconnaissance to get the necessary information. APT actors are likely to spend more time on this phase if not detected.

Once there is enough data gathered from the target, some patterns may emerge. These patterns can lead into new discoveries and enlarge the attack surface against the target. The more carefully this stage is done, the better it will be for the later stages of the attack, especially for the weaponization and delivery stages. Based on the information gathered during this stage, the attacker will construct the attacking plan and prepare the necessary tools (Chen et al., 2014). Reconnaissance will speed up the cyber-attack by eliminating dead ends and reducing the chance of detection (Bahrami et al., 2019).

Weaponization is about preparing the weapon used in the future cyber-attack. A malware is designed including remote access trojan integrated with an exploit code into deliverable payload. Once the target's assets have been identified, the attacker can start to create a necessary exploits and backdoors. Exploit can be out of the box, meaning that the exploit targets known vulnerabilities and the used exploit mechanism is known, or it can be custom built, if the attacker has enough necessary resources in their use. Custom build exploits are harder to make but can also be harder to detect by the target's security.

Attackers can also utilize 0-day vulnerabilities during their attack. A 0-day vulnerability is a vulnerability that has no patches or fix released at the time of its discovery (Bahrami et al., 2019). This makes it almost invisible against cyber security products. 0-day usage is attackers' privilege and the amount used during an attack reveals the APT actors' technical proficiency. Thus, APT actors are very careful when using 0-day exploits as they are very scarce. Most of the APT attacks don't utilize 0-days (Ussath et al., 2016). Once the 0-day is discovered, the attackers will try to utilize it as much as they can before its patched. During the weaponization stage the attackers also need to make plans on how Command and Control of the backdoor will be handled during the attack. The C&C network used in an attack should be unique for the attack, as if the network is reused and the server IP-addresses are known, there is a possibility that the target's security technology will be blocking or at least monitoring these addresses. A website, even a legitimate one, can be compromised and equipped with an exploit that can give attackers access to the target's computer (Bahrami et al., 2019). But the biggest problem with a compromised website-based attack is that the victim needs to visit the site. If the attackers want to be successful, they must prepare

multiple different exploits for different attack vectors, in the case that if one attack vector is blocked the attacker can adapt more efficiently (Chen et al., 2014).

Delivery reveals how the above-mentioned payload (weapon) is delivered into the target system. Once the exploit and backdoor have been chosen and packaged correctly, it is time to deliver them to the target organization. For this, the attackers have many different possibilities to choose from and this is where the work that has been done on the reconnaissance comes into fruition. There are two different types of delivery mechanism used: direct and indirect (Chen et al., 2014). The most used way to deliver the exploit is spear phishing. This method is a direct delivery. According to Bahrami et al. (2019) about 90% of the targeted attacks were using spear phishing in their delivery method. Spear-phishing is different from the regular phishing that the email is customized specifically to the target user. This information gathered during the reconnaissance stage will increase the likelihood of attack's success.

Afterwards, the user is either lured into downloading the attached exploit or clicking a link in the mail which in turn leads to downloading the malicious exploit or to require the user to enter credentials. According to Chen et al. (2014) APT attacks tend to use malicious attachments more than malicious links as the people normally share files via email. Ussath et al. (2016) found out that out of 22 APT campaigns analysed in their study, 15 used attachments and eight embedded links. Most common attachment file types used in spear-phishing attacks are PDF files, Flash files, Microsoft office files and portable executable files.

Some of the more modern security technologies have a way to open the malicious files via a sandbox and thus figure out if the file is malicious or not. Compressing the malicious attachment might make it more difficult for AV-scanner to identify the malicious file. Spear-phishing can also be done via social media sites. False profiles in a social network site like Facebook or LinkedIn can contact the targeted user and lure the target to click on to a link, which will lead the user into a malicious website (Bahrami et al., 2019).

Attackers can also try to attack the Internet facing servers of the target. This is useful, if other delivery mechanisms are unusable. Ussath et al. (2016) noticed that out of 22 studied APT-campaigns, only one has used this technique as initial compromise vector. The same amount of usage was in delivery with infected storage devices. Another delivery method is watering hole attack where a website is compromised to provide the exploit to desired targets. The vulnerabilities exploited are usually browser or browser plug-in related. The chance of a watering hole attack succeeding is greater than the spear phishing attack, because revisiting a website that is trusted is more likely than opening an attachment received via email (Bahrami et al., 2019). Ussath et al. (2016) found out that only four campaigns out of 22 used this as an initial compromise and from the four, two of them used 0-days to gain access to the target. This might be because compromising a website is not necessarily easy and might require a large number of resources from the attacker (Bahrami et al., 2019). Watering hole attack is not to be confused with a supply chain attack, as when the former concentrates on the website the latter goes on to the software. In a supply chain attack a supplier,

which can be a software or a hardware supplier, is compromised and their updates are poisoned with malicious code, which lets the attacker gain a foothold in the target's network. This delivery type relies heavily on the fact that the supplier, which is used to attack, is used on the target and the updates from the supplier's site are trusted. A good example of a supply chain delivery method is NotPetya, where the attacker compromised M.E.Docs updates and used it to deliver a destructive NotPetya code to machines using the software in question and this attack spiralled out of control infecting computers worldwide causing billions in costs (Greenberg, 2018).

Exploitation refers to the action after the payload (weapon) has been delivered to the system and exploitation trickers intruders' code. During this stage the wanted vulnerability is exploited and initial access to the target's network is obtained. Chen et al. (2014) notes that this exploitation can also be done with credentials obtained either through a leak or using social engineering, or the more traditional way of executing malicious code that exploits vulnerability in the target's computer. The exploitation is normally done using a vulnerability in a common software like Microsoft Office products or Adobe acrobat/reader, or even against a vulnerability in a browser.

The presence of zero-days are not common as mentioned before and the used exploits can also be older, but there is always the danger that if exploited software is patched, it will not do anything. Bahrami et al. (2019) notes that depending on what the target is, the attackers can use SQL injection to try to execute their code on the backend server. Even if the target is too well defended for a straightforward code execution, the attacker might be able to retrieve usernames and passwords from the attacked database server.

These credentials can be valuable later, especially if the users have a habit of reusing credentials. Credential reuse creates the possibility of exploitation against the target. Bahrami et al. (2019) states that the most common way to retrieve credentials is phishing emails but utilizing social media and other online fake profiles are also a possibility. There is a huge dark web market for credentials, from where the attacker can purchase the needed credentials. The prices go according to the potential profitability of the credentials. Instead of credential phishing, the phishing emails might contain a link where the user will, willingly or unwillingly, download a malicious executable, which in turn will lead to exploitation.

This drive-by-download attack has seen its evolution from the days, where the executable was shared via email. Modern email clients can distinguish the executables and most organizations have some sort of email security on them that will clean up the malicious attachments. A successful exploitation leads to the installation phase (Chen et al., 2014).

Installation of a remote access trojan or a backdoor enables persistent access to the target's system or network. After the exploitation is completed, the malware starts first installing a downloader or a dropper to the memory and after that downloading a backdoor on the compromised computer. Backdoor, also known as remote access trojan or RAT, is a program that allows remote access to

the computer in question. Installation of a backdoor might fail, if the vulnerability used has not acquired high enough rights to the system although Chen et al. (2014) argues that successful exploitation will typically lead to successful installation. Once the backdoor has been installed, the dropper will delete itself to remove any traces it had created to obfuscate the defenders (Bahrami et al., 2019). There are different installation techniques depending on the target platform. Bahrami et al. (2019) list registry key modifying, DLL search order hijacking, DLL side-loading, startup folder modification, using scheduled tasks and component object model hijacking as preferred installation methods on Windows platforms. For also non-windows platforms, the installation techniques that the attackers prefer to use are e.g., bootkits, usage of local job scheduling or to create an account with valid credentials. Creating an account with valid credentials is probably the most reliable way to achieve persistence.

In *Command and Control* (C2) phase communication between compromised host(s) and C2 server is established. Once the backdoor has been installed, the attacker can start controlling the computer and using it as a beachhead in the target's network to expand the control and enable further persistence. This will ultimately create network traffic and thus can alert the defenders (Chen et al., 2014). Command and Control traffic, from this point on C2, can be done in different ways. Most common method is to utilize normal HTTP or other common network protocols like SMTP, ICMP or DNS (Bahrami et al., 2019). The key is to masquerade the C2 traffic to the normal traffic to remain under the radar as long as possible. Chen et al. (2014) state that the attackers tend to use various legitimate services and publicly available tools for the C2 traffic. The HTTP can be directed for example to various social media sites, where the attackers can send commands via blog posts or status messages.

Bahrami et al. (2019) notes that the usage of email protocols such as SMTP and POP3 is common with C2 traffic. The attackers can use SMTP for an external mail server, or if that type of connection is not allowed, they can try to leverage the target's own email server for C2 traffic. DNS protocol is also one of the protocols the attackers can utilize to send C2 traffic, and they tend to use it in low bandwidth mode as their aim is to stay undetected.

Though listed, these networking protocols are not the only ones utilized and the limit is only the attacker's creativity. The biggest problem for the attackers with the C2 servers is that once they are found, the law enforcement tends to take them quickly down or block them (Bahrami et al., 2019). There might be a situation where the target network is air-gapped, meaning that the equipment in the network is not connected to the Internet. The only way for the attacker to get to the network is to manually insert the malware into it. Bahrami et al. (2019) notes that to circumvent air-gapped networks, attacker should use removable media like USB-sticks as it only requires someone from the users of the target network to insert the malicious USB stick to the network. They also note that C2 traffic is done similarly, inserting a specially designed USB stick into the air-gapped network. On the USB-stick the C2 data is stored in a special partition that is not visible to the user.

After the prior steps have been executed successfully can *Actions on Objectives* be executed i.e., the attacker takes action to achieve her goal(s). Once the attacker has achieved the foothold on the computer, they attempt to move laterally to infect more devices and gain access to high-value targets (Bahrami et al., 2019). Chen et al. (2014) divides the lateral movement to three different parts. First is running an internal reconnaissance to map out the network and acquire intelligence so the attackers can plan their next step carefully. Second part is to compromise additional systems and harvest credentials that have escalated privileges to make sure that the foothold to the network is strong and that the attackers have high enough credentials to meet their objective. Third part is to identify and collect valuable assets. The lateral movement and actions on the objective phase in the network take a long time as the attackers want to harvest as much data as they can, and their activities are designed to run slow and undetected. Attackers tend to use legitimate tools and OS features, which can be seen used by legitimate IT administrators.

By modelling cyber-attack with the above-described actions, it is possible to gather information from previous documented cyber-attacks and learn the attackers' modus operandi. The idea is to gather as much information as possible from different stages. If the intruder is caught in certain stage for the first time, one must assume that preceding stages were utilised also. To be able to react effectively in the following intrusion attempts, the defender must learn the tactics and techniques used in the previous stages to be effective in the future. The philosophy behind CKC is to use the same persistence attackers have on the defensive efforts by analysing every stage of the CKC carefully which enforces attackers to change every phase of their intrusion. It is important to analyse also failed intrusions and analyse what might have happened after mitigated intrusions. One of the main motivators for our thesis is expressed in Hutchins et al. (2011):

At a strategic level, analyzing multiple intrusion kill chains over time will identify commonalities and overlapping indicators... The principle goal of campaign analysis is to determine the patterns and behaviors of the intruders, their tactics, techniques, and procedures (TTP), to detect "how" they operate rather than specifically "what" they do. (p.7)

Even though CKC-model was introduced over a decade ago, it is still used in private sector and among cyber security researchers (see e.g., Lockheed Martin, 2023; Bahrami et al., 2019). However, the term cyber kill chain has also gained popularity and it is quite often used as a general term to model a cyber-attack in a particular way which the researcher has chosen or created (see e.g., Kim, Kwon & Kim, 2019). As Assante and Lee notify about the usage of the original CKC-model: "...it serves as a great foundation and concept on which to build." (Assante & Lee, 2015, p. 2). It is understandable that the situation and the nature of the cyber-attack quite often determines what model is used or even created to capture the features of cyber-attack. On the other hand, researchers have utilized the original CKC-model literally in their research of different cyber security phenomena such as APT cyber-attacks, banking Trojans, and crypto-ransomware

features (see Bahrami et al., 2019; Kiwia, Dehghantanha, Choo & Slaughter 2018; Dargahi et al., 2019).

3.2.3 MITRE's ATT&CK Framework

Strom et al. (2020) describe MITRE's ATT&CK (Adversarial Tactics, Techniques and Common Knowledge) framework as a curated database that displays different tactics, techniques and procedures found on different cyber-attacks. The aim of the ATT&CK framework is to model the behaviour of the attacker and how they could operate in the defender's environment. The ATT&CK framework has grown from its first iteration of 2015, when it only had 9 stages and 96 different techniques, to its current form of 14 stages and 193 techniques and 401 sub-techniques. It also has branched out with the creation of ATT&CK for Mobile, concentrating on Android and iOS, and ATT&CK for ICS, used in industrial control systems. In our thesis, we will concentrate on the original version, now known as ATT&CK for enterprise.

In their whitepaper, Strom et al. (2020) state that the information in ATT&CK framework is mainly obtained from open-source material on suspected advanced persistent threat group behaviour. This provides grounding for the framework as it tries accurately portraying the activity in the wild. Also, information received from techniques discovered and reported through offensive research are used. The mentioned sources of information are threat intelligence reports, conference presentations, webinars, social media, blogs, open-source code repositories and malware samples. This information is then prioritized and curated by MITRE to ensure it aligns with ATT&CK framework.

Strom et al. (2020) note that the current ATT&CK framework is divided into 14 stages which the company calls "tactics" referring to short-term tactical goals during an attack. The tactics contain several "techniques" defining the methods to achieve tactical goals. The techniques are further divided into sub-techniques when seen as necessary. Thus, tactics answer the question why a certain technique is used, and it gives the defender information on what phase the attack might be in. A single technique is not unique to a single tactic, but tactics are treated as tags for techniques in ATT&CK framework. A single tactic is an abstract container that holds in certain techniques that could be used to complete that certain tactical stage. The number of stages or "tactics" are not fixed but can be changed in the future to portray adversary objectives more accurately. In Figure 4, a part of the ATT&CK framework's matrix is displayed. There are four different stages out of 14 visible, Initial Access, Execution, Persistence and Privilege Escalation. These "tactics" have in them multiple different techniques and further sub-techniques.

Like stated earlier, a single technique is not unique to a single tactic as Figure 4 shows. Scheduled Task/Job, which is a technique where attacker abuses task scheduling functionality, is present in the three visible tactic, execution, persistence, and privilege escalation (Strom et al., 2020). And when the attacker is

using the previously mentioned tactic, what is the aim of the usage defines the tactic it falls in.

Initial Access	Execution	Persistence	Privilege Escalation
9 techniques	13 techniques	19 techniques	13 techniques
Drive-by Compromise	Command and Scripting Interpreter (8)	Account Manipulation (5)	Abuse Elevation Control Mechanism (4)
Exploit Public-Facing Application	Container Administration Command	BITS Jobs	Access Token Manipulation (5)
External Remote Services	Deploy Container	Boot or Logon Autostart Execution (14)	Boot or Logon Autostart Execution (14)
Hardware Additions	Exploitation for Client Execution	Boot or Logon Initialization Scripts (5)	Boot or Logon Initialization Scripts (5)
Phishing (3)	Inter-Process Communication (3)	Browser Extensions	Create or Modify System Process (4)
Replication Through Removable Media	Native API	Compromise Client Software Binary	Domain Policy Modification (2)
Supply Chain Compromise (3)	Scheduled Task/Job (5)	Create Account (3)	Escape to Host
Trusted Relationship	Serverless Execution	Create or Modify System Process (4)	Event Triggered Execution (16)
Valid Accounts (4)	Shared Modules	Event Triggered Execution (16)	Exploitation for Privilege Escalation
	Software Deployment Tools	External Remote Services	
	System Services (2)	Hijack Execution Flow (12)	
	User Execution (3)	Implant Internal Image	
	Windows Management Instrumentation	Modify Authentication Process (7)	
		Office Application Startup (6)	
		Pre-OS Boot (5)	
		Scheduled Task/Job (5)	
		Server Software Component (5)	
		Traffic Signaling (2)	
		Valid Accounts (4)	
			DLL Search Order Hijacking
			DLL Side-Loading
			Dylib Hijacking
			Executable Installer File Permissions Weakness
			Dynamic Linker Hijacking
			Path Interception by PATH Environment Variable
			Path Interception by Search Order Hijacking
			Path Interception by Unquoted Path
			Services File Permissions Weakness
			Services Registry Permissions Weakness
			COR_PROFILER
			KernelCallbackTable
			Process Injection (12)
			Scheduled Task/Job (5)
			Valid Accounts (4)

FIGURE 4 An Example of MITRE ATT&CK Framework Matrix (According to MITRE ATT&CK, 2023)

If tactics were an answer to “why” something is happening, technique answers to the question “how” (Strom et al., 2020). How does the attacker reach their tactical objective, what is the performed action? When doing a hijack execution flow technique, MITRE ATT&CK ID T1574, the attacker tries to gain persistence on the target machine. Thus, the answer here for the questions of why and how could be to gain persistence, the tactic, and by hijacking execution flow. But as stated earlier, some techniques can be used to achieve a different goal. In that case the defender needs to ask “what” the attacker would gain by doing this

technique (Strom et al., 2020). Previously mentioned hijack execution flow can also be utilized to gain privilege escalation or to evade defences. Thus, in some cases the answer to question “why” the technique was used depends on “what” would the attacker gain when utilizing the technique. In Strom et al. (2020) there are some techniques where the level of the technique is too abstract, thus sub-techniques have been created to give the defender more information. These sub-techniques give the defender a more detailed answer to the question of how the attack was carried out and how it could be detected and mitigated. Previously mentioned hijack execution flow, T1574, has multiple different sub-techniques, that achieve similar results but in different ways and thus they have different detection and mitigation. Hijack execution flow and its sub-techniques can be seen in Figure 4. Tactics and techniques are nothing without procedures. Strom et al. (2020) describe that procedures are the single instance where the attacker implements one or multiple different techniques to achieve their tactical objectives. They are observed in-the-wild use of technique.

There are several different possible use cases for ATT&CK framework, that are stated in the MITRE ATT&CK whitepaper. These are examined in table 6.

TABLE 6 MITRE ATT&CK Framework Use Cases (Strom et al., 2020, pp. 3-4)

Use Case	Explanation	Example
Adversary Emulation	Utilize a known TTP of an attacker to verify detection and/or mitigation for the defender.	Create a profile utilizing known Turla TTPs and test and verify the defenses against techniques that Turla has used.
Red teaming	Can help red team to create plans that help them accomplish the end objective without detection	Red team can utilize ATT&CK to avoid certain defensive measures and thus show defenders room for improvement
Behavioral Analytics Development	Using behavioral mechanisms instead of traditional IoCs can reveal previously unknown malicious activity	The attacker is using a previously unknown benign tool to achieve malicious goals.
Defensive Gap Assessment	Allows the defenders to determine possible weak points and blind spots on the defense.	Use ATT&CK as a model to assess existing defensive tools, mitigation and monitoring.
SOC Maturity Assessment	Assessing how mature the Security Operations Center (SOC) in detecting threats against the network	ATT&CK can be used as a metric how well SOC detects, analyzes and responds to intrusions.
Cyber Threat Intelligence Enrichment	Knowledge about cyber threats and threat actor groups. Includes information about malware, tools, TTPs and other indicators	ATT&CK can be used to educate defenders about the common behaviors of different attacking groups.

MITRE itself also utilizes ATT&CK framework to evaluate different cyber security product vendors on how well they perform against different adversary emulation scenarios. The aim is not to compare the products to each other, but to help the vendors to see what their products detect and how they could be improved. (MITRE engenuity, 2023). Many of these cyber security products like Microsoft sentinel (Marsiano, 2022) and Palo Alto Networks Cortex XSOAR (Laufer, 2021) have incorporated ATT&CK framework into itself to produce more context and information for the defender.

According to Strom et al. (2020) ATT&CK framework differs from other types of threat models in the level of abstraction for adversary tactics and techniques. For example, CKC framework is considered as a high-level model which is useful at understanding adversary goals and high-level processes. On the other end of the equation, is low level examination such as exploit databases and models (see figure 5). MITRE's ATT&CK is considered as a mid-level adversary model that ties high- and low-level components together providing "A common taxonomy of individual adversary actions and goals understood by both offence and defence and an appropriate level of categorization to relate adversary's action and *specific* ways of defending against it." (Strom et al., 2020, p. 23). In general, one could say that cyber security community refers MITRE's framework more technically orientated and as mentioned above it brings up more detailed attack aspects. Nevertheless, as observed in previous chapters, quite often it is about the nature of the examination and the effort researcher puts into her investigation that defines how closely cyber-attacks are examined – not the model used.

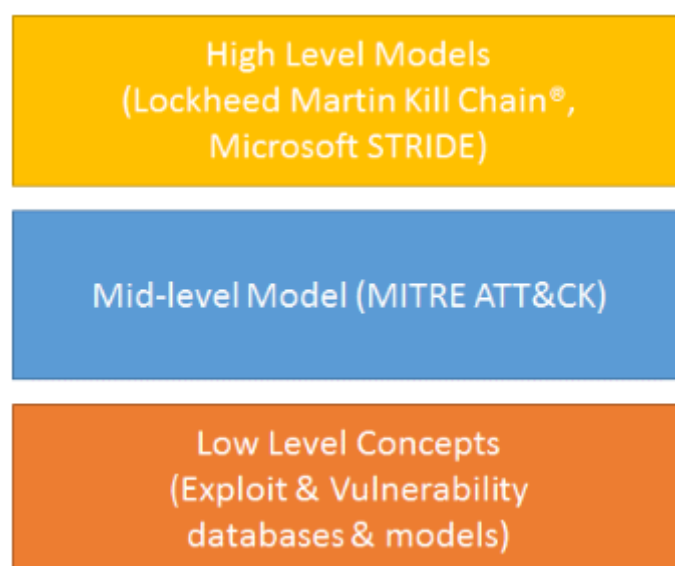


FIGURE 5 Abstraction Comparison of Models and Threat Knowledge Databases (According to Strom et al., 2020, p. 23)

In chapter 3.2 we have defined the environment where our thesis's action happens, APT cyber-attack as one of the biggest threats facing this environment and presented different cyber-attack models to describe APTs more specifically. Table 7 gathers cyber concepts which will help us in understanding and investigating China related APT campaigns.

TABLE 7 Cyber Concepts of the Thesis Part 1

Concept/Term/Model	Description
Cyberspace	A humanmade digital system in which data and information is transferred via interconnected networks using electronics and electromagnetic spectrum needing the physical infrastructure, humans, and digitalization for its existence.
APT Cyber-Attack	A cyber-attack consisting of multiple separate attack phases. APT is executed by an actor(s) which have the resources and skills to execute long-lasting and stealth cyber-attack operations. The tactics, techniques and procedures used in APT are often sophisticated and custom-made.
Lockheed Martin's Cyber Kill Chain framework	Actions: Reconnaissance, Weaponization, Delivery, Exploitation, Installation, Command & Control, and Actions on Objectives. CKC framework is considered as a high-level model which is useful at understanding adversary goals and high-level processes.
MITRE's ATT&CK (Adversarial Tactics, Techniques and Common Knowledge) framework	Divided into 14 stages which MITRE calls "tactics" referring to short-term tactical goals during an attack: Reconnaissance, Resource Development, Initial Access, Execution, Persistence, Privilege Escalation, Defence Evasion, Credential Access, Discovery, Lateral Movement, Collection, Command and Control, Exfiltration, and Impact. Is considered as a mid-level adversary model that ties high- and low-level components together

Chapter 3.3 will end our theoretical journey and clarifies the rest of the cyber concepts of our thesis. The missing piece of our puzzle is a concept that ties all the elements together: Cyber Threat Intelligence.

3.3 Cyber Threat Intelligence

Our thesis is constructed on the assumption that we are examining China related APT campaigns to be able to defend more efficiently against cyber-attacks in the future. Therefore, we need to be able to make informed decision based on knowledge gained by cyber threat intelligence. Intelligence is seen still as a relatively new information security subfield and therefore often tech-savvy cyber incident analysts without formal intelligence training can make mistakes leading to wrong "intelligence-driven" decisions (see e.g., Kime, 2016; Lee & Brown,

2021). Traditionally intelligence has been something that government and military organisations have practiced. However, someone might wonder if it is even necessary to adapt traditional intelligence process to cyber security field. The complexity of cyberspace and fast-changing cyber threat landscape imposes challenges to often long-lasting intelligence processes involving company or government policies. But if we want to fight against a threat that is advanced and persistent, we have to use methods that are advanced and especially more persistent which takes time anyway. The following examination is aimed at showing how we position our thesis in cyber threat intelligence process and how use of taxonomies can bring value especially to data collection, and when simplified enough even inform strategic-level decision makers. In the end, we finalise our Cyber Concepts table with the relevant, remaining terms of our thesis.

3.3.1 From Traditional Intelligence to Cyber Threat Intelligence

Intelligence has been around for centuries before cyber security community adapted it as part of its lexicon. Therefore, it is important to understand the features of traditional intelligence – no need to invent the cycle all over again (Lee & Brown, 2021). Cambridge Dictionary defines intelligence: “the ability to learn, understand, and make judgments or have opinions that are based on reason” (Cambridge Dictionary, 2023a). NIST Glossary (2023) has multiple definitions for the word intelligence. If we take one of them under scrutiny, one can see that intelligence definition is divided into two: a product that is a result of a *process*, and *information and knowledge* about an adversary. US military Joint Publication (JP) 2-0 describes the intelligence process as a cycle consisting of five different phases: Planning and Direction, Collection, Processing and Exploitation, Analysis and Production and Dissemination and Integration (JP 2-0, 2013; see figure 6).



FIGURE 6 The Intelligence Process (According to Joint Publication 2-0, 2013, p. I-6)

The JP 2-0 has another illustrative figure for our purposes (see figure 7). It shows how data is collected from operational environment, how it is processed and transformed into information and finally the information is refined to intelligence through analysis process.

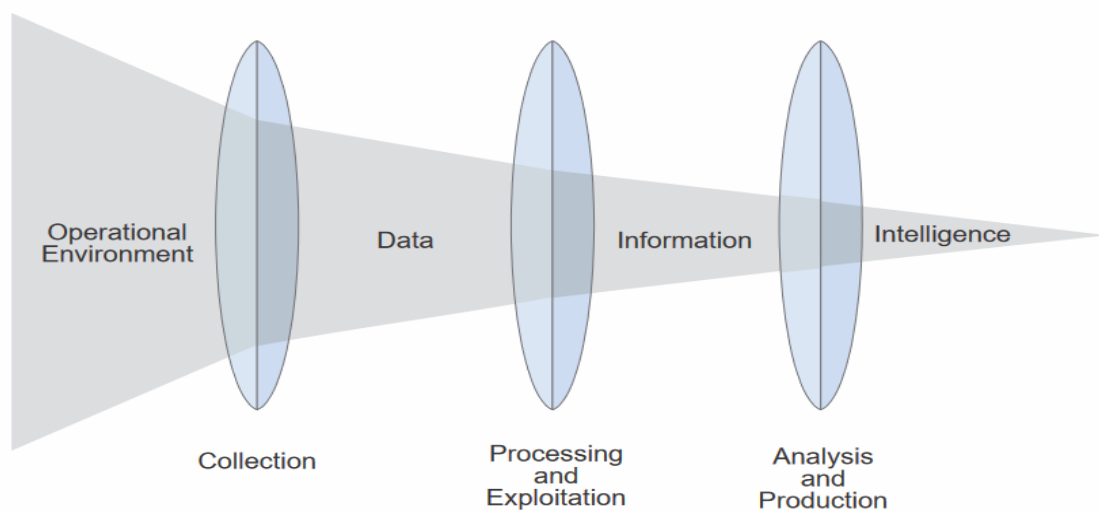


FIGURE 7 Relationship of Data, Information and Intelligence (According to Joint Publication 2-0, 2013, p. I-2)

This process shows how threat data is collected and transformed into threat intelligence (Launius, 2020). McMillan describes threat intelligence: "... as a task of gathering evidence-based knowledge about an existing or emerging menace or hazard to assets that can be used to inform decisions regarding the subject's response to that menace or hazard." (McMillan, 2013). According to NIST's Glossary, threat intelligence is threat information that has been aggregated, transformed, analysed, interpreted, or enriched to generate knowledge for the decision-makers (NIST, 2023). Often in cyber security field of expertise e.g., indicators of compromise (IOCs) are considered as intelligence even though an IOC is only a single data point.¹³ As we can observe from the figure 7, data is still two steps away from intelligence. According to Kime (2016) observed malicious activity of cyberspace added to IOC is just information and often cyber security firms are selling threat information as CTI. Albeit this can be useful in network defence efforts, it is still not intelligence.

It is common in cyber security field to call technical security practitioners as analysts. This is misleading because in most of the cases the analysts are lacking analytical way of thinking (see e.g., Heuer, 1999). Especially, if we take the stance that intelligence is not a data feed or does not come from a tool, the analyst's ability to think analytically of an adversary's intent, opportunity, and capability to do harm is of essence (Lee & Brown 2021).

Referring to military intelligence, Kime (2016) states that CTI should also incorporate tactical, operational, and strategic level intelligence of cyber security in its operations and products. In the Joint Publication 2-0 (2013), strategic level is the highest level in organisational hierarchy and strategic intelligence is produced for the executives of this level. It is used to develop national strategy and policy. In the corporate world this level could be the company's board. Operational level comes down one step and operational intelligence helps to keep executives informed of events within their area of interest. In private sector this level could be the board of directors. Finally, tactical level is the level where all the action happens. Tactical intelligence seeks to define adversary's tactical operations and gain advantage. In private companies this level can be seen as the mid-management (pp. I-23-I-25).

Because CTI is in its early years of development - the situation is the same as with APT or cyberspace definitions - the cyber security field does not have either complete or common definition for it.¹⁴ Thus, a thorough examination of traditional intelligence process adapted to CTI is missing in which all the phases of intelligence process are examined.¹⁵ One white paper goes through collection of threat data from IT infrastructure using intelligence process defined in US

¹³ IOCs are e.g., hash values, IP addresses, network and host artifacts and the tools used during cyber-attack. See chapter 3.3.2.

¹⁴ For example, NIST (2023) has the term Cyber Threat Intelligence but under headline "Definition(s)" is stated "None". See https://csrc.nist.gov/glossary/term/cyber_threat_intelligence.

¹⁵ This remark is based on our insufficient examination of published research articles. On the other hand, e.g., private company CrowdStrike Inc. has implemented traditional intelligence cycle and process to their product repertoire (Baker, 2023).

military publications although concentrating mainly on Planning and Direction phase of the intelligence cycle (see Kime, 2016). Figure 8 shows how our thesis making is fitted into intelligence process.

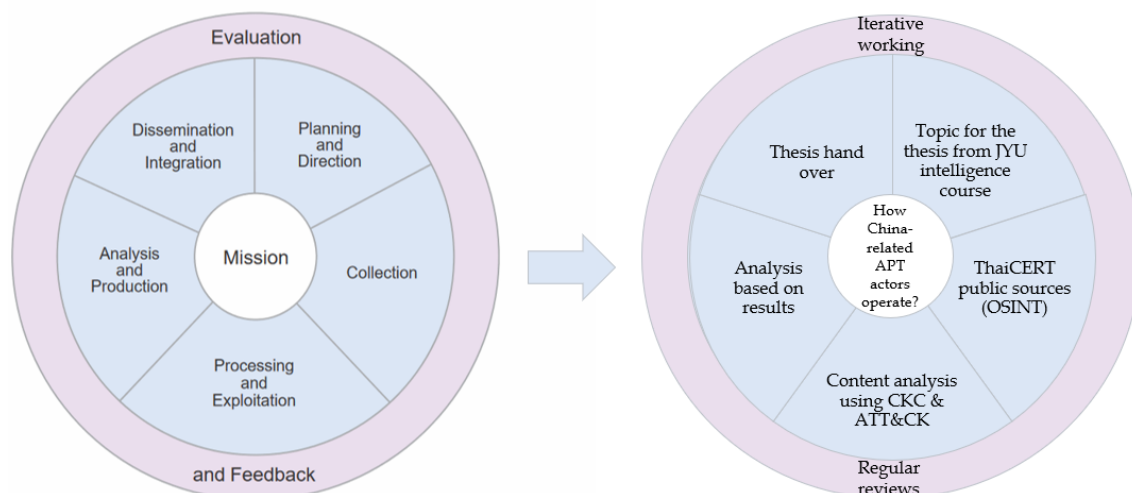


FIGURE 8 Thesis' Topic Fitted in Intelligence Process

Partially, the focus of CTI has been on taxonomies, standards, and ontologies rather than on the whole process of intelligence gathering (Mavroeidis & Bromander, 2017)¹⁶. But if we have these different forms of information collection, what is the data/information collected with the use of taxonomies, standards, or ontologies.

3.3.2 Cyber-Threat Information

Johnson, Feldman and Witte (2017) state that cyber-threat information includes any information to assist an organization to identify, assess, monitor, and respond to cyber threats. Organizations should establish relationships for cyber-treat information sharing and use cyber security community's experience to advance their security posture. According to NIST SP 800-150, major threat information consists of: indicators, TTPs, security alerts, threat intelligence reports, and tool configurations (NIST SP 800-150, 2016).

In NIST SP (2016) technical artifacts or observables known as indicators can indicate that an attack is about to take place, is already in progress, or that a compromise may have already taken place. Indicators can be used to spot potential threats and mount an effective defence. The Internet Protocol (IP) address of a suspected command and control server, a dubious Domain Name System (DNS)

¹⁶ The Structured Threat Information Expression (STIX™) is one example of a standard for sharing structured threat information (OASIS Open, 2023). Mavroeidis & Bromander (2017) have noted that the scientific community is lacking a general security ontology since the work of Blanco et al. (2008). They define ontology as a form of knowledge representation that can integrate information coming from different sources.

domain name, a Uniform Resource Locator (URL) that links to malicious content, a file hash for a malicious executable, or the subject line of a malicious email message are a few examples of indicators. This kind of data and information have been part of CTI for the longest period and often threat intelligence efforts are based upon them (Mavroeidis & Bromander, 2017). For Hutchins et al. (2011) indicator is the fundamental element of intelligence. Indicators are divided into three types: atomic, computed, and behavioural. Atomic indicators such as IP and email addresses and vulnerability identifiers cannot be broken down into smaller parts. Computed indicators are e.g., hash values and regular expression and these indicators are part of the data involved in an incident. Behavioural indicators are formed using collected computed and atomic indicators to describe intruder actions during a cyber-attack. NIST defines *TTP* as:

The behavior of an actor. A tactic is the highest-level description of this behavior, while techniques give a more detailed description of behavior in the context of a tactic, and procedures an even lower-level, highly detailed description in the context of a technique. (NIST, 2021)

In Strom et al. (2020) MITRE defines TTP slightly differently. The procedure is how a technique is used and it can span multiple techniques. Techniques tell how a tactical object is to be achieved and tactic is why the attacker is doing a certain action (see figure 9). In a simple way TTP is about adversary behaviour telling what has been done and what techniques have been used and how these techniques are utilised. TTP characterise adversaries' skills, actions and methods, i.e., modus operandi, used to gain their goals (Launius, 2020). The term can be used as a means of profiling certain threat actor and to describe an approach of analysing an APT operation (Markstedter, 2020).

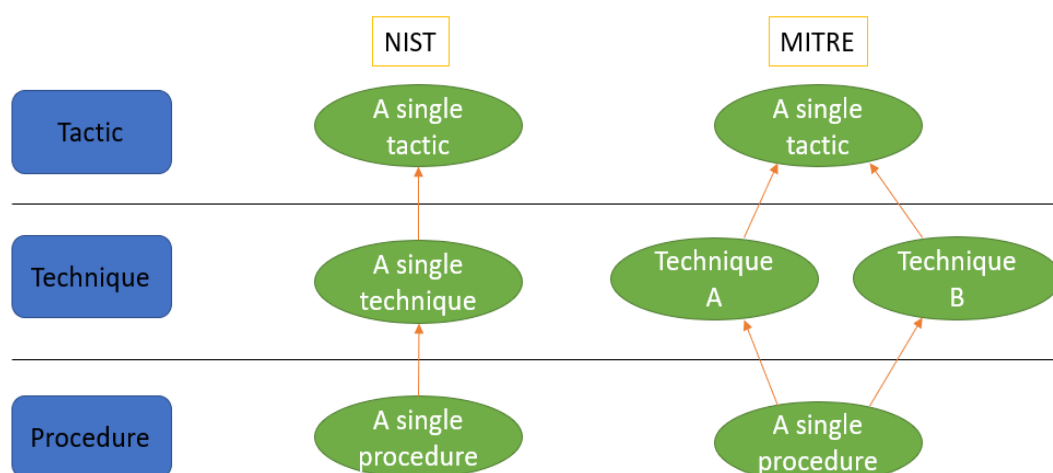


FIGURE 9 A Comparison of TTP Definition by NIST and MITRE

Markstedter (2020) sees that cyber security industry is concentrating too heavily on IOC even though TTP should be the priority number one. The automated security solutions are updated based on the newly found IOCs and this action is

reactive in nature. It is important to be able to understand adversary's TTPs especially for the target i.e., what the attacker would do in different stages of the cyber-attack when facing defensive obstacles as the defender has. This is not possible when concentrating only on static IOCs. A behaviour-oriented detection and incident response methodology is seen as more efficient in stopping or recovering from cyber-attacks. The cost of TTP exposure is considerably higher than IOC exposure (see figure 10).

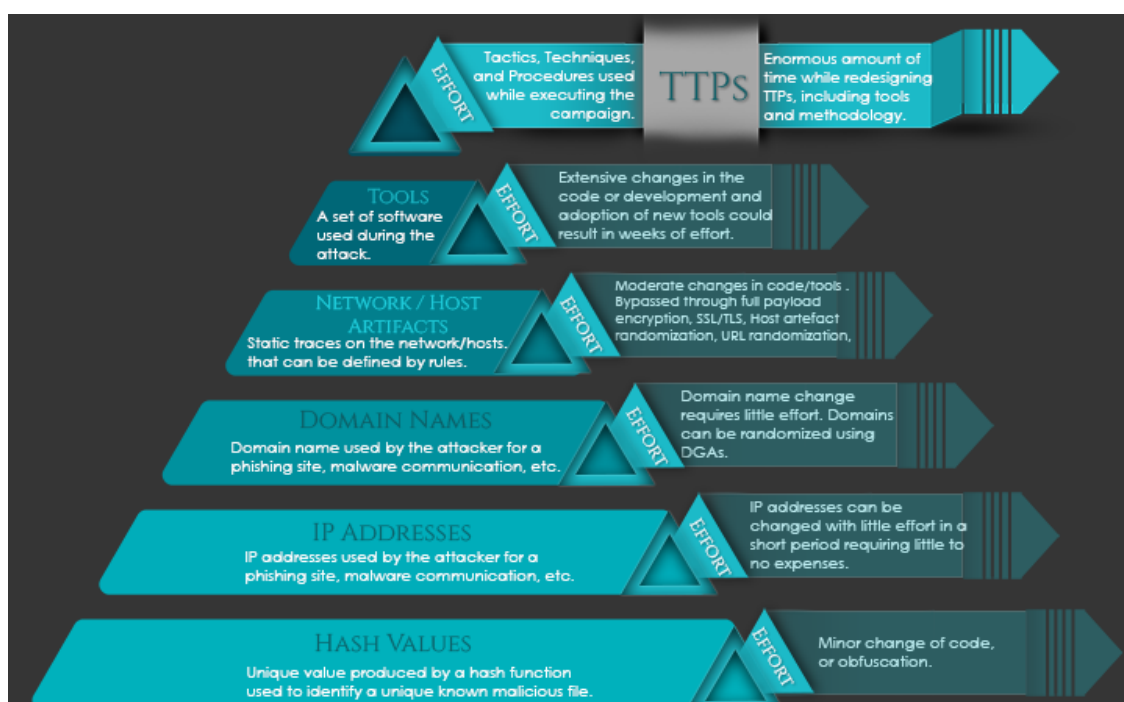


FIGURE 10 Cost of Exposure (According to Markstedter, 2020)

In NIST SP (2016) security alerts, also referred to as advisories, bulletins, and vulnerability notes, are succinct, typically human-readable technical notifications about the most recent exploits, vulnerabilities, and security issues. Sources of security alerts include the National Vulnerability Database (NVD), the United States Computer Emergency Readiness Team (US-CERT), Information Sharing and Analysis Centers (ISACs), Product Security Incident Response Teams (PSIRTs), commercial security service providers, and security researchers. Threat intelligence reports, which give an organization a better sense of the threat environment, are typically prose documents that describe TTPs, actors, the kinds of systems and information that are targeted, and other threat-related information. Threat intelligence is threat information that has been compiled, altered, examined, summarized, or enriched to offer the necessary context for decision-making procedures. Recommendations for how to set up and use tools (mechanisms) that support the automated collection, exchange, processing, analysis, and use of threat information are known as tool configurations. For instance, instructions on how to install and use a rootkit detection and removal tool, or how to create and alter intrusion detection signatures, router access control lists (ACLs), firewall

rules, or web filter configuration files, are examples of tool configuration information.

3.3.3 Cyber Threat Intelligence and the Role of Taxonomies

According to the Cambridge Dictionary, the meaning of taxonomy in English is a system for naming and organizing things, especially plants and animals, into groups that share similar qualities (Cambridge Dictionary, 2023b). Merriam-Webster defines taxonomy as the process or system of describing the way in which different living things are related by putting them in groups (Merriam-Webster, 2023). NIST's glossary gives a simple definition: "A scheme of classification" (NIST Glossary, 2023). Overall, one could say that taxonomy is about determining certain categories and distributing relevant information to these categories. According to Launius (2020) CTI standards are for modelling, storing, and sharing cyber-attack investigations' threat data and these standards are used to capture IOCs or TTPs. Taxonomy, on the other hand, is an ordered classification system with often hierarchical structure.

Considering the nature of our thesis, it is relevant to address how taxonomies fit into CTI process. According to Launius (2020) threat taxonomies are useful in structuring data to produce information during cyber threat intelligence process. When gathering information with standards and ontologies often taxonomies are the starting point of these processes. Threat taxonomies fit especially on strategic level of intelligence because they can be used to capture a consistent threat perspective (see figure 11).

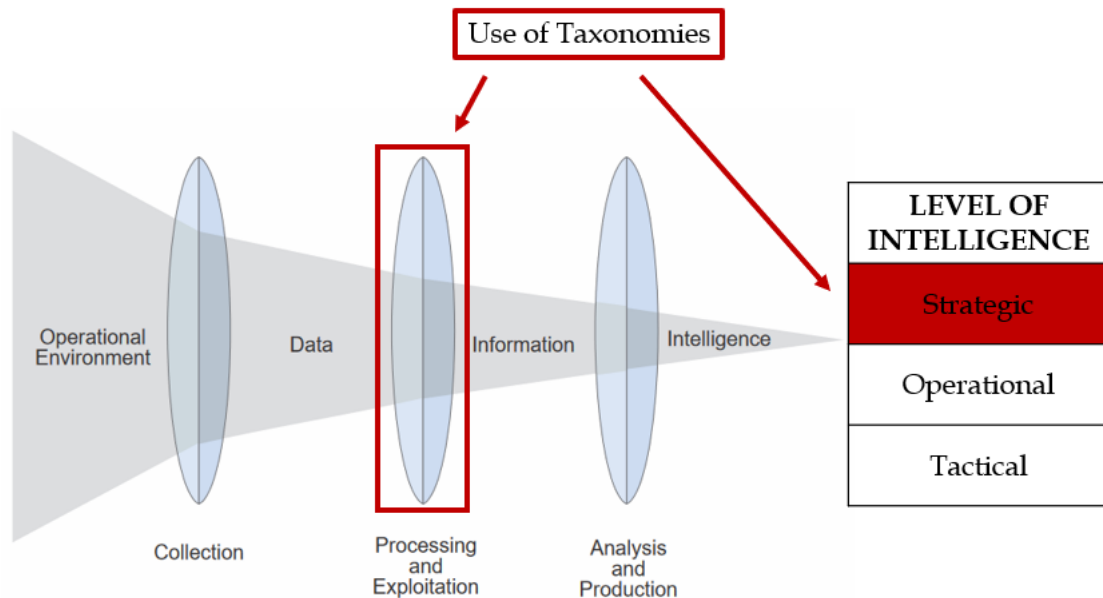


FIGURE 11 The Role and Usage of Taxonomies in Cyber Threat Intelligence

Launius (2020) states that inconsistency and lack of relevant, mutually agreed language can be seen as obstacles in communicating cyber security topics to

organisational management.¹⁷ Well-planned threat taxonomy with standard threat categories and terms is useful in strategic-level intelligence – comparing publicly available threat information to organisations’ internal threat assessments would be easier and more efficient. When cyber-attacks are analysed with a threat taxonomy it is possible to reveal trend in attack vectors and adversary methods. Quite often this has been the case, but the use of taxonomies have been inconsistent (Bahrami et al., 2019). Therefore, it is necessary to choose a clear theoretical framework for taxonomy creation and use it consistently.

In chapter 3, we have defined the environment where our thesis’s action happens, examined one of the biggest threats facing this environment, and presented different cyber-attack models to describe APT cyber-attacks for defensive purposes. In addition, cyber threat intelligence has been presented to understand, research, and in the end enable knowledge-based decisions to improve security posture. Table 8 gathers cyber concepts which will help us in understanding and investigating China related APT campaigns.

TABLE 8 Cyber Concepts of the Thesis Part 2

Concept/Term/Model	Description
Cyberspace	A humanmade digital system in which data and information is transferred via interconnected networks using electronics and electromagnetic spectrum needing the physical infrastructure, humans and digitalization for its existence.
APT Cyber-Attack	A cyber-attack consisting of multiple separate attack phases. APT is executed by an actor(s) which have the resources and skills to execute long-lasting and stealth cyber-attack operations. The tactics, techniques and procedures used in APT are often sophisticated and custom-made.
Lockheed Martin’s Cyber Kill Chain framework	Actions: Reconnaissance, Weaponization, Delivery, Exploitation, Installation, Command & Control, and Actions on Objectives. CKC framework is considered as a high-level model which is useful at understanding adversary goals and high-level processes.

(continues)

¹⁷ The cyber security field does not have a common language for adversaries, attack techniques, malwares, and threats just to name a few examples.

TABLE 8 (continues)

MITRE's ATT&CK (Adversarial Tactics, Techniques and Common Knowledge) framework	Divided into 14 stages which MITRE calls "tactics" referring to short-term tactical goals during an attack: Reconnaissance, Resource Development, Initial Access, Execution, Persistence, Privilege Escalation, Defence Evasion, Credential Access, Discovery, Lateral Movement, Collection, Command and Control, Exfiltration, and Impact. Is considered as a mid-level adversary model that ties high- and low-level components together
Cyber Threat Intelligence	Threat intelligence is threat information that has been compiled, altered, examined, summarized, or enriched to offer the necessary context for decision-making procedures. Specific definition for CTI does not exist.
Threat Information	Cyber threat information types are IOCs, TTPs, security alerts, threat intelligence reports, and tool configurations.
Taxonomy	Determining certain categories and distributing relevant information to these categories - consistent use essential.

Next, China as an actor in cyberspace is presented in a general level. However, we start the chapter by looking at the attribution issue in cyberspace and why everyone has to be cautious when referring to specific actors.

4 ATTRIBUTION ISSUE & CHINA AS AN ACTOR IN CYBERSPACE

“Instead of fighting the red dragon in the dark and without much success, it might be more useful to get to know her first. However, it is a large and complicated task due to continuing uncertainties that are not only a problem for the outsiders, but also for the Chinese themselves.” (Raud, 2016, p. 10)

In this chapter we are investigating what features determine China’s actions in cyberspace. Two issues causing distortion are worth mentioning. First, it is important to notice that we do not speak or understand the Chinese language. This of course affects our efforts. Second, as China-expert Greg Austin has brought up, the America-centric information of China’s actions in cyberspace is biased (Austin, 2015). We are aware of this. Nevertheless, our focus is to concentrate on aspects that have been interpreted as the drivers of action for China’s cyberspace activities and also some anomalies of her actions. However, before heading to special features of Chinese cyberspace activities, the issue of attribution is clarified. Chapter 4.1 works also as a disclaimer for the authors when clearly naming certain actors in cyberspace.

4.1 Cyberspace and the Issue of Attribution

This chapter helps the reader to understand why we have named our object of thesis “China related”. In the end, we cannot be 100% sure who has executed the cyber-attacks under investigation in our thesis because of the difficulty of attribution which is typical for many actions in cyberspace. However, attribution is possible, at least to a certain point, in cyberspace and the following inspection opens the complicated issue and works as a disclaimer of our thesis when it comes to naming the examined actions Chinese made. We will first look at what is attribution and its problems in general. Then we will dive into how attribution of a cyber-attack is made and what are the key elements of it. Moving forward we will look at the diamond model, which is a model designed to help analysts with the attribution. We will close this chapter by looking at different actors doing attribution and how attribution fits in current state of international law.

In the physical world, attribution is to assign credit about a work or to blame someone for a malign action (Grotto, 2020). In cyberspace, in its basic form, it means identifying the agent responsible for the action (Banks, 2017). In other words, it is the act of determining the identity or location of the attacker (Grotto, 2020). It seems like an easy task to do but in real life, cyber attribution is hard and sometimes even impossible because the Internet facilitates anonymous communication in cyberspace. It is not always clear who is attacking as the attackers can

try to mask themselves using different proxies. Banks (2017) states that even if the computer, where the attack originated from, can be identified, it is hard to know the person or group controlling the computer. He also argues that technical investigation is not enough but requires human intelligence to supplement it and that the attribution is not a simple statement of who conducted the attack but a series of judgements: was the attack just an isolated incident, who might be behind it, and are the motivations for committing the attack.

Fritz Heider, who is the father of attribution theory, has used attribution to account for the way humans reconcile perceptions and observation in their quest for understanding (Berghel, 2017). Thus, the need to understand who is to blame for a malicious action is hardwired into the human cognitive process.¹⁸ An attribution claim can mean different things for companies than to nation-states. Attribution also serves as a deterrent against possible attackers. If the target is less likely to attribute an attack, then it can be safer to attack it and not to get noticed. But there is a possibility that the attacked party might not want to attribute the attack and thus reveal what they can see in their own network or the knowledge of the existence of the attack. On the other hand, attribution might not be important at all. Grotto (2020) notes that during an attack the defending team does not necessarily need to know who is attacking them in order to defend the network. Even though “learning the identity of an adversary can help defenders set mitigation priorities and more effectively assess risks”, the attacking party might have changed their TTP from previously identified and thus the attribution and the defense priority might be misplaced.

4.1.1 Performing Attribution

When an intrusion occurs in the network, the intrusion analyst faces a bunch of questions; what happened, who did it, when did this happen, why did this happen and how did this happen? To answer these questions, the analyst must try to go through a lot of information about the incident and must know a lot of information about the threats lurking in cyberspace. And to make matters more challenging, every analyst has their own way of thinking and going through the material in hand. There is a possibility that two analysts working on the same case might end up with different contradictory conclusions. Caltagirone, Pendergast and Betz (2013) argue that there is no model or framework that provides enough information to answer if the attack is part of a coordinated campaign.

Attribution of the cyber-attack must be done quickly and consistently. Slow and inaccurate attribution has only little use for the defenders (Banks, 2017). Jaafar, Avellaneda and Alikacem (2020) argue that cyber attribution requires using a variety of datasets and analytical techniques to distill the technical evidence, which can be used to identify the malicious actor behind the attack. As stated in

¹⁸ Lehto (2022) has brought e.g., strategic decision-making to APT cyber-attack modelling to remind us that cognitive aspect is always present in planned cyber-attacks. Understanding the attacker more deeply can increase our possibilities to detect and defend future cyber-attacks.

3.3.2, there are different kinds of technical evidence, that can be divided into three different categories; atomic, behavioural and computational indicators (Hutchins et al., 2011). But not all indicators are equal. Some of them are easily modified and in the long run can be unreliable.

The pyramid of pain was created by David J. Bianco (2014) to show how much pain the defender can cause to the attacker when blocking certain kinds of indicators (see figure 12).

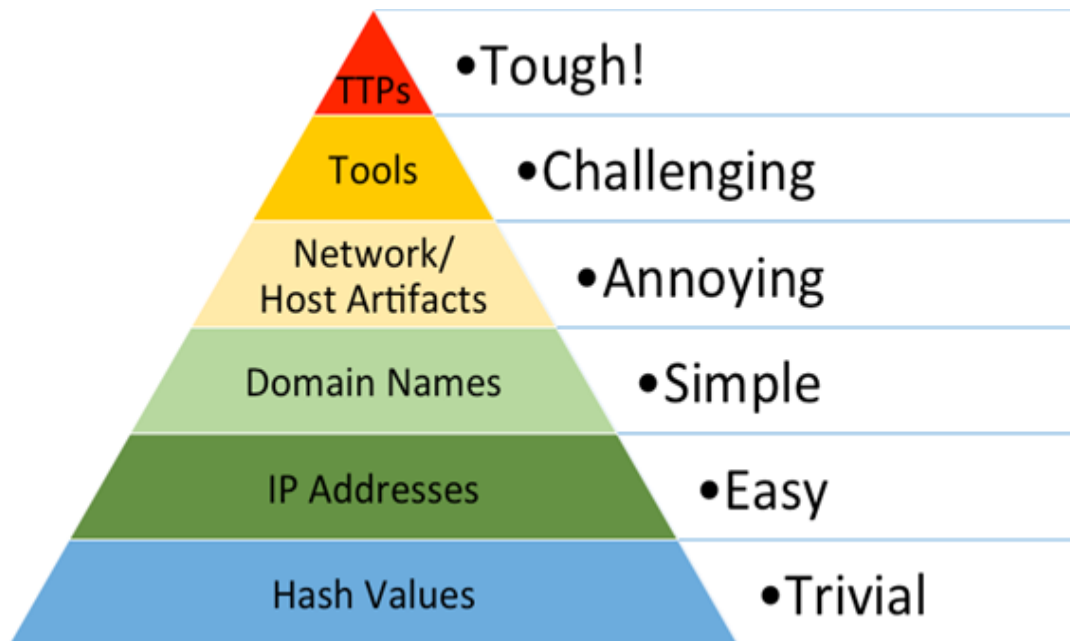


FIGURE 12 "The Pyramid of Pain" (According to Bianco, 2014, p. N/A)

Bianco's (2014) pyramid of pain can also show how reliable an indicator is: the harder it is to change an indicator, the less likely the attacker will change it. Hash values are counted from the file itself and even the smallest modification will alter the hash value. IP-addresses can change hands in the long run and an IP-address, which was detected as a part of an attack a few years earlier, may be today reassigned to a legitimate service. Even the domain names are easy to change as the registration process is made simple. On top of the pyramid is TTP, techniques, tactics, and procedures. Recognizing and blocking it would probably even stop an attack while it is happening. The pyramid of pain can also describe how hard it is to recognize an indicator.

Not all cyber-attack evidence is gathered online. According to John Carlin, most of the investigations are done offline physically examining the servers, talking to the network users, and retrieving records from the service providers (Lin, 2016). For nation-states, part of the evidence towards attribution comes from traditional intelligence like human and signals intelligence (Banks, 2017). Grotto (2020) argues that attribution has two different categories, analytical and strategic. Analytical attribution includes the forensics and all other technical evidence from the attack. It also requires all-source intelligence about threat actors and their TTPs. This procedure relies on the judgement of the analyser and analysis comes close to a strategic attribution. Strategic attribution is divided into three

different modes: private, selective, and public. What strategic attribution mode is chosen depends on the decision makers motives, which are in turn guided by political and economic factors. If private mode is chosen for the strategic attribution, the attribution and its information is not shared with anyone. In the selective mode is chosen, the attribution and its relevant information is shared to selected third parties. If a public mode is chosen, the attribution is shared with the public.

Lin (2016) states that there are three different levels of attribution: machines, human operators, and the ultimately responsible party. Machine attribution is the technical details of the attack and what computer is attacking. If the attack is done via proxy or multiple proxy computers, then the origin point of the attack chain is the attributable part. The human operator level of attribution is to discover who was carrying out the attack, who was behind the keyboard. The ultimately responsible level of attribution wants to determine who gave the orders and why. The first two levels, machine and human operators involve behaviour and the last level, ultimately responsible party, involves motivation and intent. For example, a computer from UK can be attributed at the machine level but the computer could be controlled by Pierre from France. Identifying that Pierre is controlling the computer in the UK is not easy but if it's done, that would be the human operator level. Learning that Pierre is acting under orders from the Italian mafia would be the ultimately responsible party level.

According to Caltagirone et al. (2013), the Diamond model tries to provide a model for the analyst and bring order into chaos. It tries to organize the most fundamental information of the malicious activity into a basic diamond shape (see figure 13). In order to achieve this, the model provides the user with a method to apply scientific principles to intrusion analysis. The scientific principles of measurement, testability and repeatability will allow other analysts to dive easier into the analysis and thus provide analytic effectiveness, efficiency, and accuracy. The model has been purposely kept generic and thus it can be expanded easily. It is also very flexible and captures accurately necessary information.

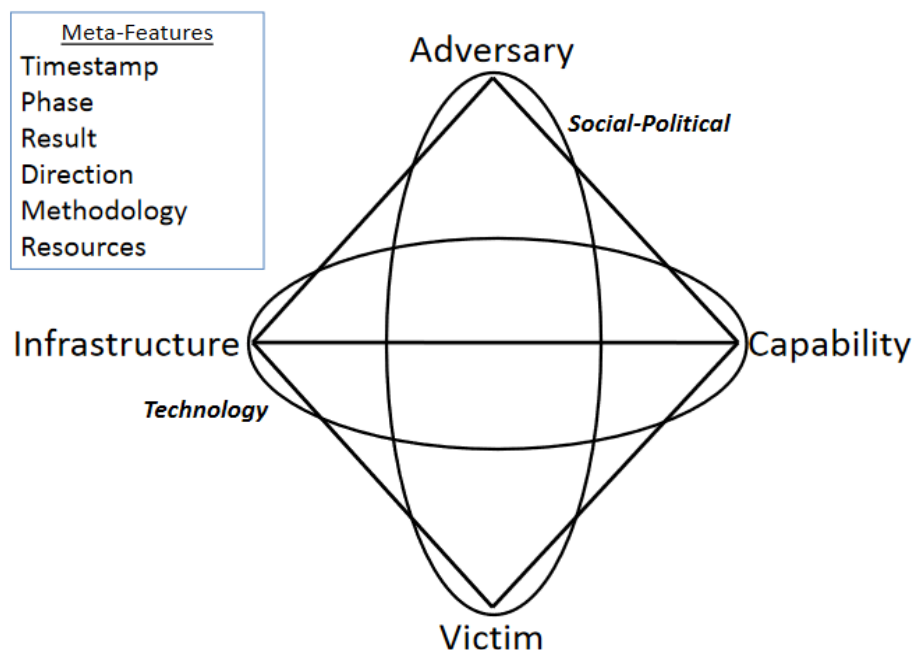


FIGURE 13 Extended Diamond Model (According to Caltagirone et al., 2013, p. 19)

Caltagirone et al. (2013) state that in its simplest form, the diamond model describes an adversary that deploys a capability over an infrastructure towards a victim. This is called an event and it's only a single step in a series that the adversary needs to do in order to reach their objective. Thus, the events can be phase-ordered into activity threads to imply the flow of an adversary operation. And once multiple activity threads have been established, similar events can be found in multiple threads, and this can be used to identify adversary campaigns. A single diamond event contains adversary, capability, infrastructure and victim as core features and timestamp, phase, result, direction, methodology and resources as meta-features. During an intrusion it is common that in the beginning most features are unknown and when more information is discovered, the missing features are filled. Adversary of the core feature means the actor who is conducting the attack. This is not limited only to outsider threats but includes all that are compromising the target system or networks. The adversary is usually elusive and is very likely to be unknown for most of the events. Capability contains the TTP of the attack, meaning that every tool and technique will fall under this feature from the simplest to the most advanced. Infrastructure features include the physical and logical communication structures that are used to deliver the attack and maintain control, like C2. Victim feature describes the target that is being exploited. It can be anything from organization to target email address or domain, only limit being that it must be necessary to the event.

The following core features of a diamond event link together. This is called the extended diamond model. The first link is adversary and victim, and they create a social-political link. This can be seen that every victim has some kind of relation to the adversary. This relation can be strong, or it can be distant or even indirect. Giles (2018) and Stubbs, Menn and Bing (2019) note an example of an indirect relation, a cyber espionage campaign called Cloud Hopper. In this attack

APT actor APT10 stole data from the US navy and multiple other organizations. They did this by hopping from a service provider to a client until they found a way to get to their real target. Even though some breached organizations did not have data that would interest APT10, they had a weak social-political link between them as the organization was in contact with APT10's real target. There is also a technology link that combines capability and infrastructure features. It represents technology that connects the two and allows them to operate and communicate.

Diamond model also provides the analyst with support to pivot (see figure 14) with the data in order to discover new related elements and thus enlarge the knowledge about the attack.

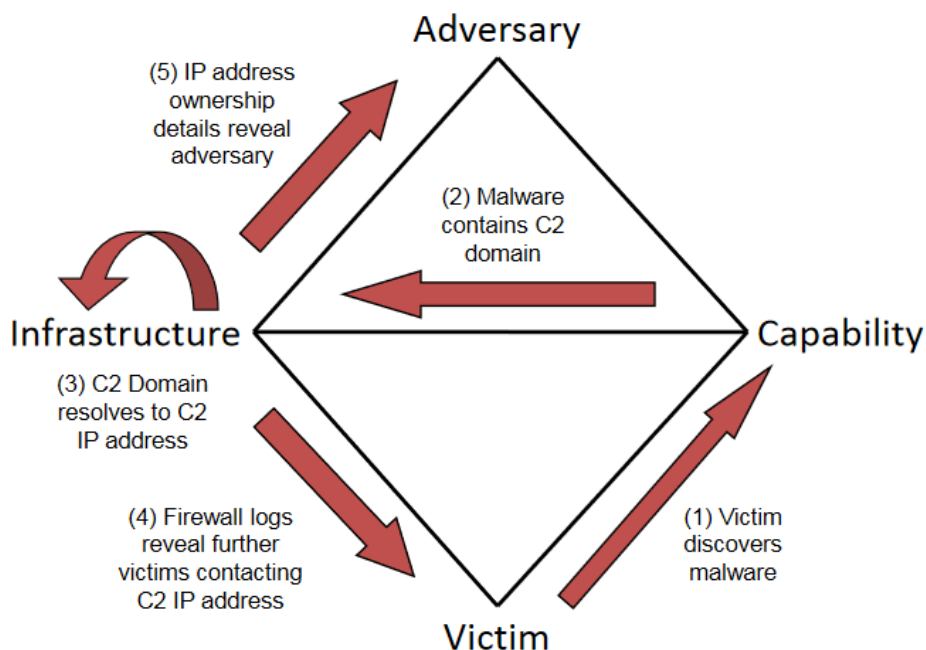


FIGURE 14 Pivoting Between the Core Features (According to Caltagirone et al., 2013, p. 27)

There are also centered approaches that concentrate on a specific feature of the diamond model. The idea is to discover activities that are connected to the feature itself. Victim-centered approach concentrates on what is happening on the victim, capability-centered investigates the malware and its related features, infrastructure-centered investigates the used malicious infrastructure and adversary-centered approach is trying to figure out the adversary and its actions.

The diamond model's relationship to attribution is very simple. As the model is used to gather and compile a large amount of data about the attack, patterns will emerge, and usage of certain tactics and techniques will be discovered. These tactics, techniques and other gathered information can be used to discover who is behind the attack. There is always some uncertainty with the attribution, but the data gathered and analysed with the diamond model is done with scientific principles and thus is as accurate as can be. If the attribution is done during the attack, it will help the defenders to anticipate the attackers' actions and mitigate or even completely block them.

4.1.2 Actors of Attribution

In the early days of the Internet, the attribution claim in the cyber realm was a state monopoly. This was broken in early 2010, when Google disclosed that it had been hacked. They even went so far as to publicly attribute the hack to China which eventually led to Google leaving the massive Chinese market (CFR, 2010). The loss of cyber attribution monopoly has turned out to be fortunate for the states. Private cyber security companies have sprung up as the world has moved more and more to cyberspace where there have been a growing number of cyber incidents. Increasing involvement of private cyber security companies rendering attribution is a positive development in the attribution landscape and can play a significant role in dissuading cyber actors from conducting attacks (Banks, 2017).

For cyber security companies, it can be easier to attribute cyber-attacks for a few reasons. The first is the lack of need to care about international politics and the second is that cyber security companies might have a better visibility to cyber incidents. Banks (2017) argue that private cyber security companies attributing a cyber-attack to a state is only speculative as the cyber security companies lack the authority and means to collect human intelligence, which is necessary for reliably attributing an attack. Still, these companies can attribute the attack and use it as a marketing tool for their own brand. As their reports are unclassified, unlike the nation actors, the attribution claim is transparent.

As there is no standardized method of naming the threat actors, complexities arise when multiple companies and researchers are claiming attribution. The attributors can use different names for the same threat actor (Lemay et al., 2018). When a company is attacked, the first order of business would be to stop the attack. After the attack and bolstering up the defenses, it is time to find out what has happened in detail and who is behind the attack. Usually, companies hire an outside cyber security company, if one is not already employed, to investigate the incident. After the incident has been researched and attribution is ready, the claim will be made if the targeted company wishes to do so. Either the hired cyber security company or the victim company can make the claim, depending on what is agreed upon in their contract.

When companies are taking up cyber insurance policies to protect themselves attribution matters. McCarthy (2019) writes that during the 2017 NotPetya cyber-attack, a company called Mondelez was also one of the victims. After the attack, Mondelez filed an insurance claim with their insurer, which in turn rebuffed the claim. The insurer was claiming that as the attack was attributed by security experts and the UK government to Russian APT group, the exclusion of "hostile or warlike action in time of peace or war" by a "government or sovereign power" came in effect. The claim was later settled without disclosing the details (Martin, 2022).

If a state is attacked, attribution of the attack becomes a bit trickier. Banks (2017) note that while the state needs to use technical evidence and human intelligence, they also need to think about politics and diplomatic relations, among

others. Attributing a cyber-attack to a state needs to have credible evidence, as the attribution can have consequences. Thus, the proof of attribution is rarely definitive, states can vary from calling the attribution. The more malicious the attack, the less confidence in the attribution is needed by the state. If no action is taken against the attributed party, states do not have an international legal obligation to reveal the evidence on which the attribution was based upon.

In July 2016 WikiLeaks released a collection of emails from Democratic National Committee (DNC) (Nakashima & Harris, 2018). Later that year more emails, this time from Hillary Clinton's campaign, were released by WikiLeaks, first in October and later in November, two days before the United States Presidential elections (Banks, 2017). It was discovered that a malicious actor had hacked both the DNC and Hillary Clinton's campaign (Nakashima & Harris, 2018). Banks (2017) notes that between July and October, several cyber experts and private security firms had attributed the hack, which had been noticed by FBI already in September 2015, to Russian intelligence service linked APT group. United States first attributed the attack to Russia on October 2016 but released a detailed report in December 2016 and unclassified report on January 2017. The slowness of attributing the hacking to Russia and downplaying the first attribution might have been due to the White House fearing an escalation to cyberwar and needing the help of Russia in Syria negotiations.

4.1.3 International Law and Cyber Attribution

Banks (2017) states that there is no clear international law about cyber attribution, but the Tallinn 2.0 manual tries to transform the norms of kinetic conflict to cyberspace. In this manual, Rule 14 is that the State bears international responsibility for a cyber-related act that is attributable to the state. Rule 15 can be described as states are responsible for cyber-related acts that either they or their proxy in states command does. Proxy can be anything from their own officials to non-state actors. But if a state is only encouraging a non-state actor to undertake a malicious cyber action, it is not the state's responsibility.

Healey (2011) has created a tool called spectrum of state responsibility, which tries to help with assigning responsibility of a cyber-attack. This tool has 10 different categories from state-prohibited, where states ban cyber-attacks altogether, to the state-integrated, where the government has integrated third-party attackers to their own cyber force. All the categories are shown in table 9. Healey also suggests that there should not be too much obsession with which particular villain pressed the enter key, but nations should stop the villain altogether. Those nations that comply and put an end to cyber-attacks should be rewarded with funding, training, and technology. Those that refuse to cooperate should be dealt with full spectrum of coercive policies, from sanctions and prosecutions all the way to covert action and kinetic military force.

TABLE 9 The Spectrum of State Responsibility Explained (According to Healey, 2011, p. 62)

Category	Examples of State Actions/Involvement		
Cyberattack:	Conducting	Abetting	Ignoring
State-Prohibited	None	None	Low: Inability to secure computers, but attacks prosecuted
State-Prohibited-But-Inadequate	None	None	Low: Inability to secure computers and stop attacks
State-Ignored	None	Low: Stalling investigations and possibly tipping off attackers	High: Disregard private attacks and fail to seriously investigate
State-Encouraged	Low: Possible “off-duty” attacks by officials or military	Low to Medium: Statements to embolden or energize attackers	High: Disregard private attacks and fail to seriously investigate
State-Shaped	Low: Possible “off-duty” attacks by officials or military	Medium: Some technical and targeting support	High: Disregard private attacks and fail to seriously investigate
State-Coordinated	Low: Possible “off-duty” attacks by officials or military	Medium to High: Coordination of timing, targets, or tempo	High: Disregard private attacks and fail to seriously investigate
State-Ordered	Low: Possible “off-duty” attacks by officials or military	High: Direct command of private attackers	High: Disregard private attacks and fail to seriously investigate
State-Rogue-Conducted	Medium: Forces attacking, without authority	None: The national government is not behind the attacks and may stop them	Medium: Other agencies may disregard the rogue attacks
State-Executed	High: National forces attacking with authority	None: The only attackers belong to state organizations	None: The only attackers belong to state organizations
State-Integrated	High: National forces attacking with authority	High: Direct command of attackers; technical and targeting support	High: Disregard private attacks and fail to seriously investigate

Mueller, Krindal, Kuerbis and Badiei (2019) argue that there is a need for a clear legitimate process that uses technical attribution outcomes to attribute the attack to a responsible party. There is some international pressure, both from public and private sectors, to create a neutral global platform, where it is possible to publicly perform authoritative cyber-attributions. Nations should hold one another responsible to stop the attacks (Healy, 2011).

In this chapter, we have discussed attribution, how it is made and how to utilize the diamond model in performing attribution. It is clear that performing an

attribution to a cyber-attack is hard, even when it is possible to analyse all the technical details of the attack. These details are usually not available to the public nor academia, which will leave the reader reliant on the cyber security professionals' analysis. Thus, in our thesis we have not performed our own attribution. Instead, we are relying on attribution performed by others. As all the chosen primary research material has been attributed to China by cyber security companies, we try to understand Chinese actions in cyberspace in the following chapter.

4.2 Drivers and Anomalies of Chinese Cyberspace Activities

This chapter examines China's alleged drivers for actions in cyberspace and sheds light on why China is considered by the West as one of the most active cyber espionage countries in the world even though the practice of foreign intelligence as understood in the West has been part of the China's playbook for a relatively short period of time – Deng Xiaoping's reform and opening-up program started China's spurt in espionage efforts in the late 1970s (Inkster, 2015). We want to connect our review to APT cyber-attack modelling especially to Lehto's (2022) model presented in Chapter 3.2.1 (see figure 15) and more closely to the Early-Attack phase and Strategic decision-making to give better understanding of circumstances behind Chinese APT cyber-attacks/campaigns. The idea is to gather possible motivations/drivers for actions for Chinese decision-makers when planning APT cyber-attack based on the existing research literature. Our method is to pick thoughts and actions which can be interpreted as a possible source of motivation.

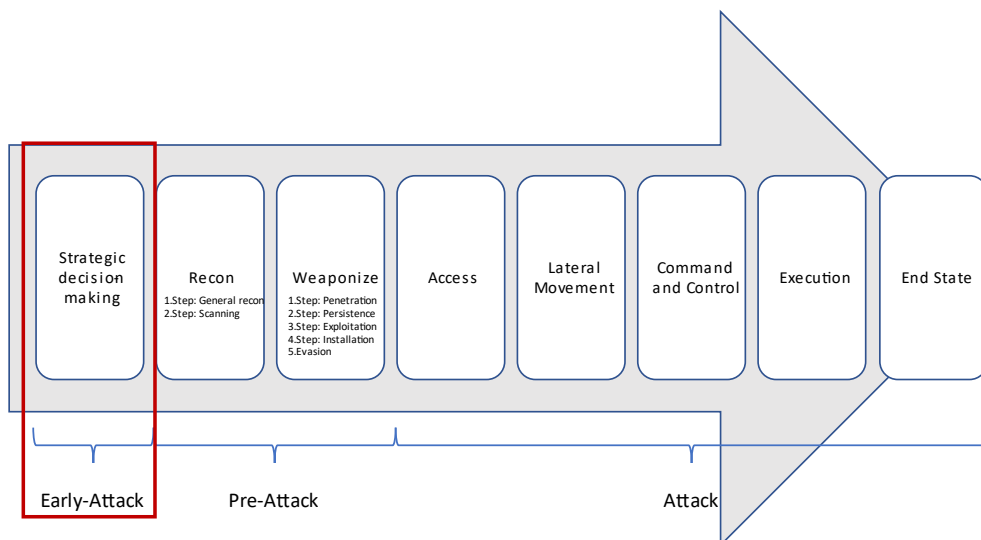


FIGURE 15 General Cyber-Attack Model (According to Lehto, 2022, p. 126)

4.2.1 Domestic Influence

When China formulates its stance, it does not happen the way we have used to in western democracies – this also affects the decisions made on actions in cyberspace. The ruling authority Chinese Communist Party (CCP) has the complete control of the nation. Political decisions and actions are made to preserve the status quo. A well-known example is China's information control machinery. According to Austin (2018) China's government spies on itself with the largest internal monitoring system in human history. CCP has two cyber security policy objectives above others: first, to assure its citizens and the world of the capabilities to monitor, shape, and dominate cyberspace political content in China. Second, to create a national cyber industrial complex which stands international competition and replaces foreign actors from large parts of China's domestic cyberspace. The internal surveillance and intelligence collection system is the crown jewel of Chinese cyberspace security, and it is only a few decades away from becoming an Orwellian-like surveillance society.

Lindsay (2014) has argued, the ruling Communist Party has not been able to transform its obsession with political security into working technical network security. Hence, cybercrime thrives inside China and creates an online underground economy. Chinese hackers target Chinese victims with a low risk of domestic police action. We, the thesis writers, argue that the internal surveillance works as a rehearsal playground for Chinese hackers to practice covert operations outside of its borders.

A few special features of China's domestic surveillance system, e.g., Great Firewall, Skynet, Great Cannon and Sharp Eyes, gives an idea what Austin means when referring to Orwellian-like surveillance society (see e.g., IISS, 2021; Austin, 2018; Cheng, 2016)¹⁹. As Cheng (2016) explains the most obvious and well-known of the Chinese cyberworld anomalies are the Great Firewall and the Great Cannon. The data that is flowing into China via fiber-optic cables from the broader global information networks is limited only to three entry points: The Beijing-Tianjin region, Shanghai, and Guangzhou. Chinese authorities are limiting and monitoring the data going through these three points and the main tool for this is the Great Firewall of China. It acts as the first line of technical defence and theoretically the firewall could shut down all the Internet traffic between China and the rest of the global Internet if necessary. In China "Baidu" plays the role of Google. It is a Chinese-run search engine and one of the most visited sites in China daily. With the use of the Great Cannon, China can redirect some of the traffic going to Baidu website towards specific Internet addresses. Together with the Great Firewall, the Great Cannon can be used as a platform for conducting

¹⁹ "Skynet" is a large video-surveillance network that contains over 200 million cameras nationwide. An extension to Skynet network is the "Sharp Eyes" that concentrates on rural areas and utilizes big data and AI for social control (IISS, 2021).

DDoS attacks against targets around the world - all with the help of regular users visiting Chinese websites.²⁰

Austin (2018) draws the obvious analogy from history when he states that modern day “great wall”, referring to the historical Great Wall of China, is a new cyber version of the wall which will keep out unwanted intruders from Chinese cyberspace. In 2017, China made the use of virtual private networks (VPN) illegal without registering VPN software with the government first. With the use of VPN, it is possible to hide your Internet traffic and make it impossible to monitor it at the same time. During the same year, the Chinese government banned anonymous access to the Internet and made Internet service providers (ISP) responsible for the verification of the real identities of account holders. China has put tremendous effort into new applications of artificial intelligence during the past few years. Facial recognition in crowded places and public transports and near instant recognition of the identity of vehicles’ licence plate holders are a few to mention.

According to Cheng (2016), Internet service providers, Internet cafes, and other instances providing access to Internet services are under constant surveillance. ISPs are forced to use Chinese software for information security and monitor their Internet traffic based on certain rules. Internet access providers must make sure that all the users use their real names when online. Many of the controlling features are written in the Chinese National Security Law. Because of the Security Law, ISPs have created different filtering systems to detect sensitive words and phrases and have hired workers to monitor online behaviour simply to keep the authorities calm. This is complemented by governments own cyber police. One tactic for the Chinese social media companies has been handing out user credits for community members who report sensitive information to administrators.

The idea of brainwashing via Western technology also lurks in the Chinese mentality. For countries with authoritative rule, one of the main fears is change – someone could provide a different world view which does not suit the ruling authorities. China sees foreign technology as a threat to its core values and this also works as a motivator to tighten the grip of its cyber security investments in the future (Doshi et al., 2021). Raud (2016) explains that China is still dependent on foreign information systems and big US tech giants are providing material and services to fulfil the rapidly growing IT demand in China. Events like the wars in Kosovo 1999 and Iraq 2003, Stuxnet 2010 and Snowden revelations 2013 have accelerated China’s suspiciousness towards foreign technologies and

²⁰ According to Kaspersky, 2023b a denial-of-service attack (DoS attack) is a cyberattack that aims to make a computer or network resource unavailable to its intended users by irreparably interrupting the operations of a host that is linked to a network. In order to overwhelm systems and prevent some or all valid requests from being fulfilled, denial of service is frequently achieved by sending an excessive number of requests to the resource or machine that is being attacked. The incoming traffic that floods the victim during a distributed denial-of-service assault (DDoS attack) comes from numerous diverse sources. Since there are numerous sources, it is necessary to employ more advanced mitigation techniques in order to prevent this kind of attack.

showed the importance of commanding the electromagnetic sphere. Chinese government's goal is to reduce dependence on foreign technologies. Thus, Chinese companies and researchers have more support and incentives to develop intellectual property. This kind of development can lead to industrial espionage because in many cases local Chinese companies lack the required skills and end up acquiring the needed information from abroad, especially if under time pressure from central government.

Quite often innovative development processes and natural development take time. Therefore, the possibility to cut corners becomes highly tempting. Gilli and Gilli (2018) call countries that steal information from more advanced countries as Imitators. Imitators can save resources, avoid making the mistakes of the innovators and use unused resources to improve existing technology. When imitation is cheaper and faster than innovation, the imitators will derive advantage.

In addition to the dominance of foreign corporations on the cyber security service market, Austin (2018) sees that it is important to notice that China's political leaders have not been able to trust what happens in cyberspace due to the absence of a significant domestic cyber security industry.²¹ For example, China's critical infrastructure remains unprotected because it does not have the capability or enough scientific study to form a strong base for the country's overall cyber security. It is important to notice that cyber security education in China is not on a sufficient level. On the other hand, Chinese researchers have made breakthroughs in certain areas of information security such as quantum computing and cryptography. Nevertheless, China has a huge deficit of cyber security skills, and it is quite unlikely that China can fill this gap in one or two decades. China has invested in the Information Security Institute in Wuhan that could provide non-degree training for 10,000 students in the future. Chinese government has reacted quite recently to the late start in informatisation of society and cyber skills deficit – it is also well-known that content security has overrun system security in the past.

4.2.2 Mentality and Motivations for Action

Chinese do not use the word "cyber" as broadly as the western countries. Cyberspace is considered as a subset of information space which is interpreted as a communication to the whole world comprising human information processing and cognitive space (Raud, 2016). Orinx and der Swielande (2019) give an illustrative example of this Chinese thinking in comparison to the traditional US thinking based on gaining raw material power; China has adapted a more holistic

²¹ Austin (2018) states that the big Chinese companies such as TenCent, Alibaba, Baidu, Huawei have thousands of information security workers but still China needs to cooperate with foreign corporations in many cyber security core technologies. In the banking sector China can be seen as one of the leading cyber security countries but on the other hand a large part of its corporate world has not taken cyber security seriously on the road to profits and competing priorities.

stance - influencing the cognitive process through cyber power. China uses the cyberspace as a dis/misinformation platform to corrupt western liberal-democratic value-base and for promoting authoritarian more nation-state-centric sovereign world order. "The Chinese general, military strategist, writer and philosopher Sun Tzu is often quoted for stating that 'the supreme art of war is to subdue the enemy without fighting'" (Kaska, Beckvard & Minárik, 2019, p.4). Cheng (2016) argues that for China it is important that the ruling and management of the Internet is in the hands of nation states. This way foreign instances cannot control who has access to the Internet and nonstate actors could not challenge Beijing's authority. The thought of state sovereignty would be expanded to cyberspace.

According to Raud (2016) one of the main concerns for China is the US dominant information technology world and how it uses its role to establish international norms and rules favourable to itself. Therefore, China entitles its actions in the cyberworld as a response to hostile US actions. China sees itself equal with the US in the digital world and this creates tensions. In general, China wants to control cyberspace through the government and military while the West prefers liberal environment, which gives space to individuals and private corporations.

Doshi et al. (2021) recognize that another issue for China is global standards. Standards are seen as the guideline and rules how to function on a global scale. Western countries have set the standards to this point but now China wants to be the actor who determines the global discourse on information technology by setting the standards that others must follow. China recognizes it missed the industrial revolution. Therefore, the country wants to be sure not to miss the ongoing technological revolution. Everything is based on gaining the competitive advantage and throughout history competitive advantage has been gained if an actor has been in the front row of revolutionary development. China sees 5G network development as an opportunity to gain the leadership in technological revolution – the one who has the discourse and rule-making power has the power to set the rules and lead in the future. In Chinese thinking, civil and military ability development go hand in hand. Therefore, 5G network development is seen as an opportunity to advance the Chinese military-civil fusion industry. Simply, the idea is that the foundation of an information force is the network. Without a proper network the information army is useless. 5G networks enable the transition to a state in which all domains: land, sea, air and space are connected to each other establishing an interconnected battlefield; advanced military tools; and battlefield communications.

Kolton (2017) sees that China has experienced a collective trauma of imperialism and the People's Liberation Army (PLA) is doing everything to prevent the same to happen in the cyber domain. Stuxnet showed Chinese leaders that even physically isolated networks are vulnerable and therefore China must be a reliable actor in cyberspace to deter aggression and protect its sovereignty.²² China has kept its cyber operations to a certain level i.e., mainly cyber espionage instead of using Stuxnet-like cyber weapons.

²² About Stuxnet see Zetter (2014).

According to Cheng (2016), digitalization and improvements in technologies have paved the way for China to use different tools to attack against foreign media organizations that cover stories related to China which are sensitive to the Chinese political leadership. At the same time, the Chinese military has learned to understand the nature of Information Age wars. The PLA is not only interested in getting all the latest equipment. On the contrary, it is interested in how to best exploit all the current equipment on its hands. This has demanded a suitable doctrine, and for the execution of this doctrine PLA has had to develop tactics, techniques and procedures. In the 1990s and early 2000 PLA apparently figured out that the most essential systems would be information related systems such as sensors, communications, and computers.

Austin (2018) sees that information- and cyber security can be seen as a subset of national security in China. When the use of the Internet in China increases and as the Internet slowly transforms more into Chinese, China can use this development to challenge Western liberal way of doing things. In general, one can say that China puts more weight on ideological threats of cyber security and the West on the other hand emphasizes technical threats. Thus, China sees cyber security as a socio-technical phenomenon which can be described as a state of mind or process for reducing risks from activities in cyberspace that are a threat to the subject's well-being. Lindsay (2015) sees that the discussion of Chinese cyber threats has two narratives: the impact of information technology on international security and the political and economic future of a rising power (see table 10).

TABLE 10 Typology of Cyber Threat Narratives (According to Lindsay, 2015, p. 12)

	Evolutionary Technology	Revolutionary Technology
Cooperative Political Environment	Open Internet Assumption: The Internet enhances the value of social and economic exchange. Threat: State censorship and surveillance violate human rights and reduce trust on the Internet. Counterargument: Prioritization of information control over technical defence exposes China to foreign and domestic cyber-attack.	Cybersecurity Norms Assumption: States must adopt common norms to protect the Internet catastrophe. Threat: Authoritarian "Internet sovereignty" norms imperil the liberal "multistakeholder" system. Counterargument: The institutional status quo is durable, and China cannot credibly commit to its proposed norms.
Competitive Political Environment	Contested Cyberspace Assumption: cyber technology improves intelligence collection methods and opportunities. Threat: Chinese cyber espionage is systematically eroding the competitiveness of Western firms. Counterargument: Absorption of stolen data is a nontrivial obstacle, and Western intelligence also exploits China.	Cyber Warfare Assumption: Cyberspace is a dangerous, asymmetric, offence dominant warfighting environment. Threat: China can paralyse U.S. military command and control and civilian infrastructure at low cost. Counterargument: China's cyber capabilities do not live up to Chinese rhetoric, and "informatisation" exposes China to attack.

As stated in IISS (2021) report, a part of China's cyber strategy has been the use of cyber operations for strategic effect abroad since the early 2000s. These operations have included industrial-scale espionage designed to acquire both commercial intellectual property and personal data. According to Bozhkov (2020), cyber operations originating from China can be divided into four intertwined categories: Industrial espionage, Domestic operations, Geopolitical and Foreign Policy operations, and Cyber activities in the military realm. China's motivations to target especially governments' networks are to control and shape narrative about regime's rule, reconnaissance through high-value networks, and information gathering for later actions against foreign governments. Diotte (2020) estimates that China acquires foreign technology through imports, foreign direct investment, industrial and cyber espionage and by establishing foreign R&D centers abroad.

The above-mentioned features in this chapter are trendy and highlighted mostly in Western literature. See table 11 for possible motivations/drivers for actions for Chinese decision-makers when planning APT cyber-attacks.

TABLE 11 Alleged Chinese Drivers/Motivations for Action in Cyberspace

Source	May Affect Strategic Decision-Making
Austin (2018)	To monitor, shape, and dominate cyberspace political content in China. Content security has overrun system security in the past.
	China sees cyber security as a socio-technical phenomenon which can be described as a state of mind or process for reducing risks from activities in cyberspace that are a threat to the subject's well-being.
	To create a national cyber industrial complex and replace foreign actors from China's domestic cyberspace.
	Chinas' political leaders have not been able to trust what happens in cyber-space due to the absence of a significant domestic cyber security industry.
Raud (2016)	China wants to control cyberspace through the government and military while the West prefers liberal environment, which gives space to individuals and private corporations.
	The importance of commanding the electromagnetic sphere
	In many cases local Chinese companies lack the required skills and end up acquiring the needed information from abroad.
	China entitles its actions in the cyberworld as a response to hostile US actions.
Orinx & der Swielande (2019)	Use of cyberspace as a dis/misinformation platform to corrupt western liberal-democratic value-base and for promoting authoritarian more nation-state-centric sovereign world order.
Doshi et al., (2021.)	5G network development is seen as an opportunity to advance the Chinese military-civil fusion industry.
	China wants to be the actor who determines the global discourse on information technology by setting the standards that others must follow.
	Officially stated goal of becoming "Cyber Great Power".

(continues)

TABLE 11 (continues)

Kolton (2017)	China has experienced a collective trauma of imperialism and the People's Liberation Army (PLA) is doing everything to prevent the same to happen in the cyber domain.
Cheng (2016)	It is important that the ruling and management of the Internet is in the hands of nation states.
	The PLA is not only interested in getting all the latest equipment. On the contrary, it is interested in how to best exploit all the current equipment on its hands. This has demanded a suitable doctrine, and for the execution of this doctrine PLA has had to develop tactics, techniques, and procedures.
Bozhkov (2020)	China's motivations to target especially governments' networks are to control and shape narrative about regime's rule, reconnaissance through high-value networks, and information gathering for later actions against foreign governments.

For example, one could argue that China executes cyber-attacks to get an impact on the way decision-makers see the open and free Internet and the goal would be to get the international community to rewrite the rules of Internet decision making to the hands of the nation-states. One line of thought could be that China is leaving traces on purpose regarding its actions in cyberspace to deter its enemies without revealing too much of its cyber capabilities and presence in foreign systems. As chapter 4.1 has shown, as long as we cannot attribute cyber-attacks more precisely, the more we have only speculation instead of clear evidence of one's actions. However, literature gives us at least some ideas/thoughts we can refer to when thinking about the hardest part: motivation for actions behind cyber-attacks. Next, we give a snapshot of the possible Chinese actors and their actions in cyberspace.

4.2.3 Actors and Operations

Our examination of Chinese cyberspace actors is non-exhaustive and based on public information. Because in China institutions are reformulated and changed continuously, we are not even trying to solve this puzzle.²³ Instead, the idea is to get a better understanding of the possible actors capable of executing APT cyber-attacks (see appendix 2). If our information has already missed the train, it is worth remembering that the possible current/new actors are not born out of nowhere.

Because the ruling communist party surveillance apparatus is all over society, it is hard to make a difference between official and unofficial actors. China uses this posture when denying any involvement in foreign cyber espionage operations. Especially, US officials and private cyber security companies have constantly brought up how the Chinese officials are often working on the

²³ See Cheng (2017) especially on PLA's cyber and information warfare organisations.

background of certain APT groups (see e.g., McWhorter, 2021). As IISS's (2021) report states:

While China's core cyber-intelligence capabilities are therefore formidable domestically, it has also developed and extensively used cyber for overseas espionage. These intelligence efforts are often characterised in terms of their volume rather than sophistication, with Chinese intrusions featuring heavily among those detected and attributed by Western intelligence agencies and cyber-security companies. (p. 92)

US officials and cyber security companies have been the most eager to reveal hostile Chinese cyberspace actors. In 2014 United States indicted five Chinese military hackers on charges on computer hacking and this was the first time in history criminal charges were filed against known state actors for hacking (Orinx & der Swielande, 2019; see figure 16). Diotte (2020) sees that the US-China Cyber Agreement (2015) changed China's tactics from using all the possible assets in cyber espionage efforts to a more focused and professional way of doing things. This caused a decrease in the number of detected Chinese cyber-attacks but created a stealthier way of conducting cyber operations (Diotte, 2020). China's cyber espionage apparatus is a mixture of official national organizations and unofficial cyber adversary groups (see appendix 3).



FIGURE 16 Chinese Military Hackers (According to Orinx & de Swielande, 2019, p. 62)

According to Raud (2016) PLA research institution's report states that China poses specialized network warfare units that operate in the military and civilian domains executing both offensive and defensive operations. China has also so-called cyber militias formed out of hackers, IT companies, scientists, network engineers, foreign language speakers, and others with useful skills for the government's cyber operations. China is ready to take the risk that independent hacker

groups might turn against their homeland because these groups have proven their value in cyber operations.

Austin (2018) clarifies how CCP's cyber security defense systems are operated by "cyber security bureaus" that are under the Chinese Ministry of Public Security (MPS) and spread all over the country. These units have existed for many years. Since 2014, MPS has established over 1000 cyber police units which have different capabilities according to their mission e.g., Level One units are part of the 13 major website companies such as Baidu, TenCent, and Sina. Thus, all ministries and public sector agencies have cyber security units that report directly to senior officials, but the capability of these units remains unknown. The leading organization for content security is the Central Propaganda Department (CPD) which gathers all the monitoring and actions under one roof. CPD unites with parts of Cyberspace Administration of China (CAC), the Cyber Emergency Bureau, and the Coordination Bureau for Cyber security. There is also a Cyber security Administration Bureau in the Ministry of Industry and Information Technology (MIIT) whose responsibilities include industry supervision of both technical security and content together with CAC.

According to Cheng (2016, p. 185) China has three types of computer network warfare forces:

1. Specialized Network Warfare Forces, military units that implement offensive and defensive network operations.
2. Authorized Forces, non-military units that are organized with military permission, drawn from public sector e.g., Ministry of State Security and Ministry of Public Security, and other suitable government departments.
3. Civilian Forces, voluntary civilians who can conduct network operations.

As highlighted in appendix 2, the Strategic Support Force (SSF) and CAC have received special attention recently. According to IISS (2021), SSF has gathered all PLA's cyber capabilities under one umbrella. It consists of two main elements: the Space Systems Department and the Network Systems Department. PLA's cyber functions were previously scattered all over different departments and organizations, but the reform has created more unified force able to execute complex and multidimensional information operations and helps PLA to shift more efficiently from a peacetime to a wartime posture.

The previously mentioned computer network warfare forces can be seen as a home for multiple different Chinese APT actors, as some of them have been attributed to PLA (see e.g., Mandiant, 2013; Cheng, 2016, p.183). These actors, come in all shapes and sizes, but all of them seem to serve the needs of the Chinese state. Chinese APT actors tend to target those nations, that possess information that interest the Chinese government. They also target foreign companies that possess intellectual property that Chinese companies could use to get an advantage. Even though the following five examples tell the progress of Chinese hacking against US, it can be viewed as how China uses its cyber muscle against its targets. These muscles are known today as Chinese APT groups.

There is always the first time for everything and for Chinese military hacking the US, the operation, code named Titan Rain, was the first of its kind, at least for the operations that have come to public knowledge (CFR, 2021). Started in 2003, this operation continued all the way until 2007, even though it was discovered in 2005 (CFR, 2021) and there were some early indicators in November 2003 (Thornburg, 2005, Cyvera, 2016). This was not just a single breach but a string of operations against the government and private sector that are connected to the governmental entities like defence contractor Lockheed Martin (Brook, 2011). These attacks, that also targeted the government of United Kingdom, are believed to be the work of a group of hackers either part of or connected to Chinese military (Norton-Taylor, 2007). Security researchers were able to trace the attacks to the Chinese province of Guangdong and to three specific routers located there (Thornburg, 2005). Cyvera (2016) notes that even though there have been some theories that the attackers were just using the vulnerable Chinese infrastructure to launch these attacks and the Chinese had nothing to do with the Titan Rain, the attack has been attributed to hackers from the Chinese military. When the attack occurred, the concept of APT was still unknown. Still, this series of attacks can be identified as an APT.

On the January of 2010, Google (2010) announced that its networks had been breached by a sophisticated and targeted attack. Google believed that the attack was related to Chinese human rights activists and their Gmail-accounts but noted that there had been other large companies that according to their investigations had also been hacked. Soon after the Google's announcement, Adobe also came forward and informed that they had also been hacked (Zetter, 2010a). Later it was stated by the media that a total of 34 companies had been breached during the attack and the targets were from the technology, finance, and defence sectors (Zetter, 2010b). This attack was significant, as it was the first time a private sector company in the US had come forth and attributed this kind of attack to a single country. As the attack was stated as sophisticated and targeted, it means that the attacker was not some random group of script-kiddies but an APT group and as it was originating from China, meaning that the group is a part, or at least somehow connected to, the government of China. The attack was conducted with the usage of zero-day vulnerabilities to Internet Explorer, the most popular browser at the time, and to both Adobe Reader and Adobe Acrobat, the most common software to read and to modify PDF files (Zetter, 2010c). The usage of zero-day vulnerability also suggest that the attacker was a nation state actor, as zero-day vulnerabilities are very rare. Once the zero-day vulnerability has been discovered, it loses its strength and thus becomes a vulnerability that will eventually be patched.

Why the attack happened, is uncertain. Google has always been in dangerous waters with its search engine in China, having to censor its results to please the Chinese government and its regulators (Sheehan, 2018). But after the attack, Google decided to stop censoring the search results (Google, 2010; Sheehan, 2018). When the attack occurred, Google informed that some intellectual property was lost to the attackers as well as information about the human rights activists in

China (Zetter, 2010b). Later it was discovered that the attackers were looking for accounts that had lawful wiretaps on them and they managed get their hands on the database that had information about FISA, Foreign Intelligence Surveillance Act, orders (Zetter, 2013). Dave Aucsmith, a senior director from Microsoft, states that according to their investigations, the attack against Google was only counterintelligence and the hackers were an elite group and possibly state-sponsored (Corbin, 2013). Sabbagh (2010) describes another theory about the attack that is based on leaked US embassy cables from China, that were published by Wikileaks in 2010. This theory states that a senior member of Chinese communist politburo used Google to search his own name and found articles that criticized him personally. This prompted him to use the powers of the state sponsored hacking groups to launch an attack against Google. The endgame for Google was that as they refused to play by the Chinese government rules of censoring the Internet, they were forced to abandon the Chinese market, even though they managed to hold on to their Hong Kong search engine and some of their services, like Gmail and Google Maps were available in China (Sheehan, 2018).

Probably the most significant announcement, that indicates Chinese involvement in cyber-attacks against US, came from Mandiant, a US cyber security company. On this Mandiant (2013) report, released on February 2013, Mandiant attributes an APT group to a Chinese military unit and manages to provide evidence about its activity from a long period of time. The report states that APT1 is part of the Chinese military organization, specifically unit 61398, and has conducted a vast number of cyber operations against multiple different targets. The amount of target organizations mentioned on the report is 141 and the amount of data stolen can be counted at least in the hundreds of terabytes. According to the report, APT1 has spent vast amounts of time in their target network, the longest persistence was 1764 days. APT1s targets have been in conjunction with the industries that are vital to the growth of China, as detailed in the China's 12th five-year plan.

The above-mentioned report was significant that for the first time there was conclusive evidence brought to public knowledge that there are Chinese state sponsoring hacking groups that are targeting other countries. Still, the report fails to mention any target organizations by name. APT1 is the first group to carry the designation of advanced persistent threat, but it was not the last. After Mandiant released the report of APT1, there have been multitude of different APT groups in different names from different countries. But this report brought forward the information of Chinese state sponsored hacking and showed the world the APT1's TTPs, forcing them to change their way of hacking (Mandiant, 2013).

Throughout many cyber-operations conducted by China, the operation against US Office of Personnel Management (OPM) might be one of the most significant because of its aftermath. Fruhlinger (2020) notes in his article that the operation itself was started in November 2013 with an initial breach to the OPM networks. This breach was noticed on the March of 2014 and corrective actions were taken on the May of 2014. Even though OPMs security team managed to purge out the intruder, unknown to them, a second actor had managed to gain a

foothold and managed to siphon data out of OPM networks until it was noticed and blocked in April 2015. Tools used by the intruders were Chinese APT favourite PlugX and Sakula. The amount of data stolen from the OPM was around 20 million personnel records (Spetalnick & Martina 2015) and it included biometrics data from the government employees (Fruhlinger 2020). Spetalnick and Martina (2015) informed that this hack was one of the talking points, when China's leader Xi met up with then US President Obama in September 2015. Two leaders did agree that neither of the governments would knowingly support cyber theft and share the information to the state-owned enterprises or private sector. Still, this agreement left out traditional government against government cyber espionage, to which category OPM hack fell. From the first glance, it seems that the agreement did decrease the numbers of active APT groups and their operations against the US, but this might also be because the attacks have evolved to become stealthier and harder to detect (Segal, 2016). In February 2020, United States Department of Justice did attribute the OPM hack to Chinese military, when they pressed criminal charges against four Chinese military members for Equifax hack, which was linked to the OPM hack (Fruhlinger, 2020).

The commercial aircraft market has long been a duopoly between the US-based Boeing and its European rival Airbus. Even though there are some smaller manufacturers, these two western companies dominate the market, even in China, which is growing to be the largest aviation market in the world. A new Chinese manufacturer called Comac intends to become an alternative to the western manufacturers. This is part of the China's most recent Five-year Plan, where one of the priority industries is aerospace and aviation equipment (CrowdStrike, 2019). But to create an airplane and its parts takes time, knowledge, and technological know-how. A cyber security company CrowdStrike (2019) released a report in 2019 detailing about activities of an APT group called Turbine Panda. The report states that the APT group Turbine Panda conducted cyber operations between 2010 and 2015 targeting several companies that manufacture aviation components, especially to the Comac's new aircraft C919. These operations were mainly conducted by compromising one aviation manufacturer and from there spreading out to other targets. Against an aerospace manufacturer called Safran, Turbine Panda utilized an insider to install their Sakula-malware to Safran's networks. According to the report the insider was recruited Jiangsu Bureau of the MSS (JSSD), part of the Chinese Ministry of State Security. The report also ties JSSD to Turbine Panda. The aim of these cyber operations was to gather technological information so that Chinese state-owned enterprise like Comac could accelerate its progress of developing an alternative to western aircraft. The report estimates that the technological information gathered during the Turbine Pandas operations have shaved of several years and billions of dollars of Chinese made jet engines development time.

These operations still had their consequences. CrowdStrike (2019) reports that an MSS Officer, that had been recruiting Chinese born insiders to work at the target organizations, was arrested by the US officials along with his recruits. The arrest of Yu Pingan, who was the developer of Sakula malware, by the US

officials in 2017 (Fruhlinger, 2020; CrowdStrike, 2019) led MSS to issue a ban for the Chinese Security researchers not to attend Security conferences overseas in a fear that they might be arrested (CrowdStrike, 2019). Table 12 sums up Chinese APT operations' investigations.

TABLE 12 Details of Chinese APT Operations

Operation	Year	Performed by	Target	Motivation
Titan Rain	2003	A group of hackers either connected or a part of Chinese military	US and UK Governments and private sector connected to them	Espionage
Operation Aurora	2010	Hackers originating from China	Google and 33 other private US companies	Espionage
Mandiant APT1	2013	A first public report indicating a connection between cyber hackers and the Chinese Military	Multiple targets	Espionage
OPM attack	2013	Chinese Military	US office of personnel management and Equifax.	Data theft and espionage
Turbine Panda	2015	Turbine Panda in orders from JSSD	Multiple aerospace manufacturers	Data theft

In chapter 4, we have discussed the hardest issue in a cyber-attack, the attribution. Even though there are tools to perform attribution, it is very hard to perform a 100% attribution claim. This chapter also took a closer look to China and investigated its motivation and actions in cyberspace. Table 13 illustrates part of the spectrum of the States Responsibility presented in table 9, where the states conducting cyber-attacks are high. As presented in this chapter, China allegedly utilizes its national cyber force and conducts cyber-attacks with authority.

TABLE 13 A Part of the Spectrum of State Responsibility (According to Healey, 2011, p. 62)

Category	Examples of State Actions/Involvement		
Cyberattack:	Conducting	Abetting	Ignoring
State-Executed	High: National forces attacking with authority	None: The only attackers belong to state organizations	None: The only attackers belong to state organizations
State-Integrated	High: National forces attacking with authority	High: Direct command of attackers; technical and targeting support	High: Disregard private attacks and fail to seriously investigate

The difference between state-executed and state-integrated as presented by Healey (2011) is that on state-executed, the state is behind everything, while on state-integrated, state is utilizing also third-party attackers by coordinating and directing them. Cheng (2016) notes that China has individual attackers that are not part of the military, but their actions are coordinated and even directed. This would position China's responsibility as state-integrated, where according to Healey (2011):

The national government integrates third-party attackers and government cyberforces, with common direction and coordination. Orders and coordination may be formal or informal, but the government is in control of selecting targets, timing, and tempo. The attackers are de facto agents of the state. (p.61)

5 RESEARCH METHODOLOGY

As Myers (1997) has stated: “A research method is a strategy of inquiry which moves from the underlying philosophical assumptions to research design and data collection. The choice of research method influences the way in which the researcher collects data.” (Myers, 1997, p. n/a). This chapter follows Myers’s logic and entangles our thesis to scientific research field by elaborating the philosophical and analytical decisions we have taken to answer our research question. First, we elaborate the philosophical basis of our thesis and attach our thesis to qualitative research approach and more specifically to theory-driven content analysis. Second, we present how the study of China related APT actors’ TTPs has been executed, and third, we reflect on the reliability and validity of our thesis.

5.1 Scientific Research Basis

Among the triad of philosophical perspectives: positivist, interpretive, and critical, our thesis’s epistemological stance can be seen as interpretive. According to interpretive epistemology access to reality happens only through social constructions such as language, consciousness, and shared meanings. (Myers, 1997.)²⁴ However, when taking the interpretive research position, we do not exclude the ideas other epistemologies have about knowledge and how it can be acquired. As Töttö (2004) points out, all research is actually just scratching the surface – research can never achieve the phenomenon in its entirety and in all its depth. However, thorough research, i.e., well-planned and implemented research designs, repeating studies, and approaching the studied phenomenon from several perspectives, can reach diverse information and thus increase understanding of both the cause-and-effect relationships related to the phenomenon and the nature of the phenomenon. We argue that if a researcher focuses too strictly on some predetermined scientific form of seeing things, it inevitably affects the analysis, and some important observations may be left out.

We position our study as qualitative but do not want to stick to this position too strictly either: our study also includes elements of quantification. Qualitative research includes numerous different traditions, approaches, and data collection and analysis methods for studying people and their lives, so it is not a research approach to any particular discipline or just one way of studying (Hirsjärvi, Remes & Sajavaara, 2004). Qualitative research focuses on meanings that manifest themselves in a wide variety of ways (Varto, 1992). According to Myers (1997) the roots of qualitative research rest on social sciences and qualitative research methods were developed for the study of social and cultural phenomena.

²⁴ See Myers (1997) for epistemological questions in information systems research.

However, qualitative research has been also part of information systems research already for a while.

Saaranen-Kauppinen and Puusniekka (2006) note that sometimes quantitative analysis can also be useful in qualitative research. Counting qualitative data is sometimes easier to grasp than just reading it over and over again. Calculations sometimes also provide assurance that the conclusions presented as research results are not based only on feeling. Qualitative analysis can be supported by, for example, calculating the number of codes or elements belonging to different themes. Calculation systematizes the analysis, but calculations are naturally also constructs of the researcher. The research task and the researcher's interests affect the method of calculation and the choice of objects of calculation.

Alastalo and Vuori (2023) describe that one of the main features that determine if a study is qualitative in nature is the use of qualitative data such as interviews, documents, and records often presented as text. Documents can be divided roughly into personal and institutional, as well as public and non-public. Institutional documents are created in the everyday work of an institution or organization. Some of the institutional documents are public and some are not. For example, when we visit a doctor, we collect medical records. When we are interrogated by the police, interrogation protocols and police reports are written about what we say and do. These are not public documents and their availability for research use is strictly regulated. However, the institutions also produce a lot of public documents, such as plans, reports, investigation, and strategies. Compiling systematic material from public documents is not always simple, as locating documents related to the studied phenomenon can be difficult and require contacting their producers or painstaking archival work. Sometimes documents can be compiled from the websites of organizations. Website updates are problematic in this case because documents may be lost and therefore unreachable in the future. That is why online documents should be saved to ourselves.

As brought up in Chapter 2, usually the cyber security industry has a monopoly of APT primary source information. Therefore, often researchers have to rely on secondary data as a pivotal source of information. When using ready-made data, the research process can progress either from one's own problems to searching for ready-made material or vice versa: some ready-made material can spark imagination and make the researcher develop new research problems, or suitable ready-made material for analysis can be found for one's own interests (Saaranen-Kauppinen & Puusniekka 2006). Not all ready-made data are suitable for research as such, but they must be modified or limited to a suitable format according to research interests (Hirsjärvi et al., 2004).

Tuomi and Sarajärvi (2018) state that in content analysis data is examined in detail looking for similarities and differences and summarizing. Content analysis is textual analysis, which examines materials that are already in text form or have been converted into text. The texts studied can be almost anything. Content analysis aims to form a condensed description of the studied phenomenon, which links the results to the broader context of the phenomenon and other research results related to the topic. Content analysis can refer to both qualitative

content analysis and quantitative analysis of content, and both of these can be utilized when analyzing the same material. Content analysis can be continued by producing, for example, quantitative results from verbally described material. In qualitative content analysis of research data, the data is first fragmented into small parts, conceptualized and finally rearranged into a new kind of entity. There are three ways of doing content analysis: material-oriented, theory-driven, and theory-guiding. If the analysis and classification of the research material is based on existing theory or the research material itself brings up classification schemes makes the difference between the above-mentioned ways of doing the research. Research is theory-driven when the analysis of research material is based on existing theory or model. Often the idea is to test the model or theory in a new context.

Theory-driven analysis can be seen as a traditional model of analysis in natural science research and is based on existing theory, model or thinking of an authority. In research utilizing theory-driven analysis the theoretical model is described and the concepts under examination are defined in theoretical framework of the study. In theory-driven analysis, something already known defines the phenomenon under investigation and is based on deductive reasoning – from broad generalizations to specific observations (Tuomi & Sarajärvi, 2018). Vuori (2023) notes that qualitative content analysis is based on the researcher's coding, where the researcher identifies, and names content elements found in the material. In coding guided by theory, the researcher chooses, based on his theoretical understanding, which aspects of the data are of interest to him. In the case of texts, these passages can be very different in scope, from single phrases to long fragments.

5.2 Theoretical Framework

Our thesis can be seen as theory-driven because we had chosen existing theories, CKC- and MITRE ATT&CK frameworks, for basis of our examination. To best of our knowledge, combination of CKC- and MITRE ATTA&CK frameworks had not been used in Chinese APT research, and also other APT-studies utilizing these two frameworks were scarce (see Chapter 2).

On a higher-level, we have utilized Kuusisto and Kuusisto (2015) thinking and modelling when adapting content analysis to our theoretical framework forming the theory-driven model of analysis. Kuusisto and Kuusisto (2015) see the cyber world as a complex adaptive system which can be examined reasonably by forming simple rules from complex systems (see figure 17).²⁵ It is possible to

²⁵ Kuusisto & Kuusisto (2015) have used the term “Cyber World” instead of the term “Cyberspace” we are using in this study (see Chapter 3.1.1).

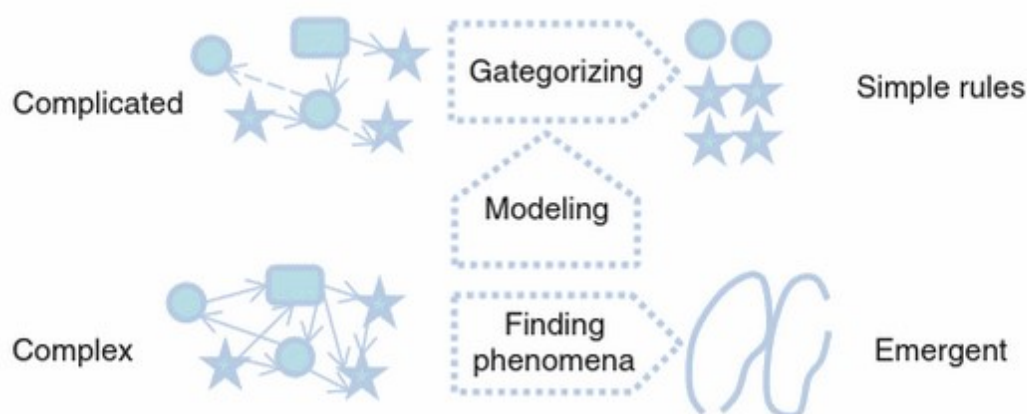


FIGURE 17 The Content Analysis of the Cyber World (Adapted from Kuusisto & Kuusisto, 2015, p. 37)

form simple rules when understanding and analysis of phenomena of the complex systems increases. During the process complicated models are used to categorize and analyse information. With the support of the identified phenomena the complicated models can be further simplified by abstracting and categorizing their contents. The results of content analysis are seen as patterns presenting emergent phenomena of the cyber world. These patterns can be helpful in a more thorough analysis of the cyber world and end up improving future security planning and implementation.

Edgar and Manz (2017) have quite similar logic as Kuusisto and Kuusisto (2015) of taking a grip of complicated phenomena by dividing it into pieces, when stating that the goal of scientific exploration in cyber security is to gain knowledge to enable quantification of security and help us predict what tools and practices are better for prevention of cyber attackers (see also, Myers, 1997).

Following Kuusisto and Kuusisto's analogy, our theory-driven content analysis (combination of CKC- and MITRE ATT&CK frameworks) is adapted into Kuusisto and Kuusisto (2015) "The content analysis of the cyber world"-figure - forming simple rules from complex system. The usage of two different frameworks does not break the idea presented in figure 18, where complex systems become simple rules, but they actually complete each other as highlighted previously.

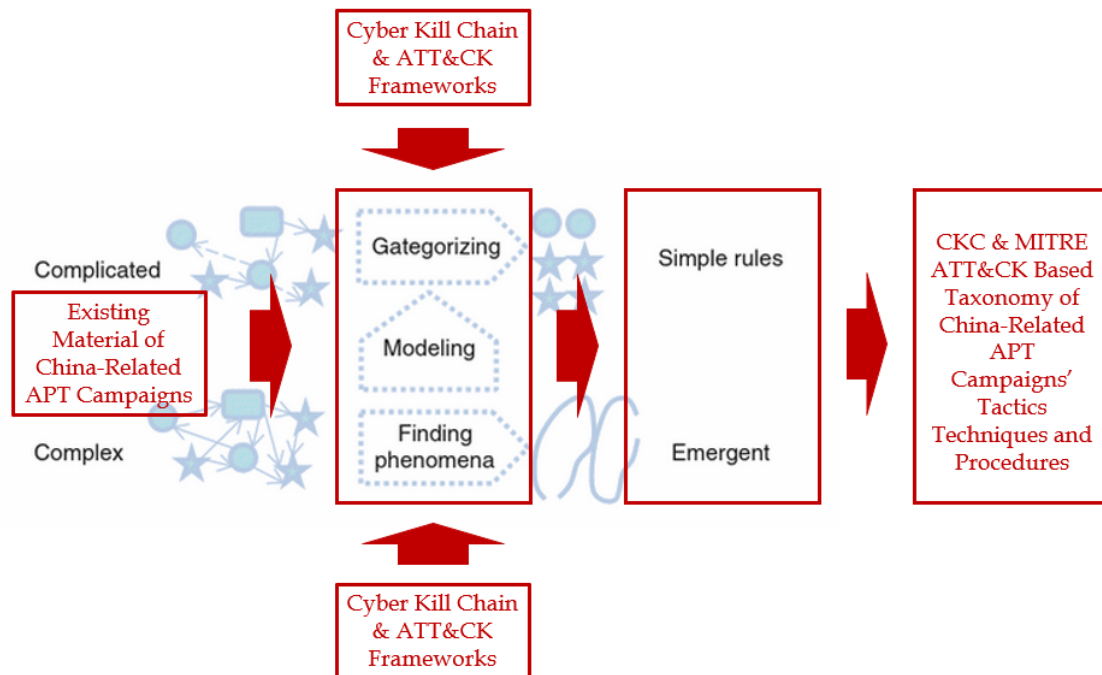


FIGURE 18 Theory-Driven Content Analysis of China Related APT Actors' Tactics, Techniques and Procedures

In the chapter 3.3.2, the definition for TTP was presented as how a suspected advanced persistent group, or a threat actor, behaves. Both NIST and MITRE note that the tactic part is the highest level, technique or techniques are a more detailed description and procedure is the lowest level, the most detailed description of a behaviour (Strom et al., 2020; NIST, 2021). In this thesis, CKC-driven content analysis will be used as a tool to extract procedures from the selected research material. Procedures will be grouped using the CKC framework as a primary data grouping classification. It is worth noting that while CKC is a framework, it is a high-level framework. Hence, it only covers the tactic aspect of TTP, leaving out the technique part. To address the CKC inability to handle techniques, the technique categorizing from MITRE ATT&CK Framework was utilized. The results of CKC-driven content analysis were furthermore tagged with ATT&CK Framework's techniques, resulting in more complete vision of the attackers TTPs. Thus, the usage of both CKC and MITRE ATT&CK frameworks complete each other in this thesis. In a nutshell, a procedure is the center of our attention i.e., text phrase taken from a report handling Chinese APT group's actions during an APT campaign. A tactic is the goal an actor wants to achieve in the target system, and this is based on CKC-framework's attack phases. Technique is what kind of techniques an actor utilizes to achieve these tactical goals in the target system. The actor executes certain techniques based on MITRE ATT&CK-framework, single or multiple, in order to achieve the tactical goal (see figure 19).

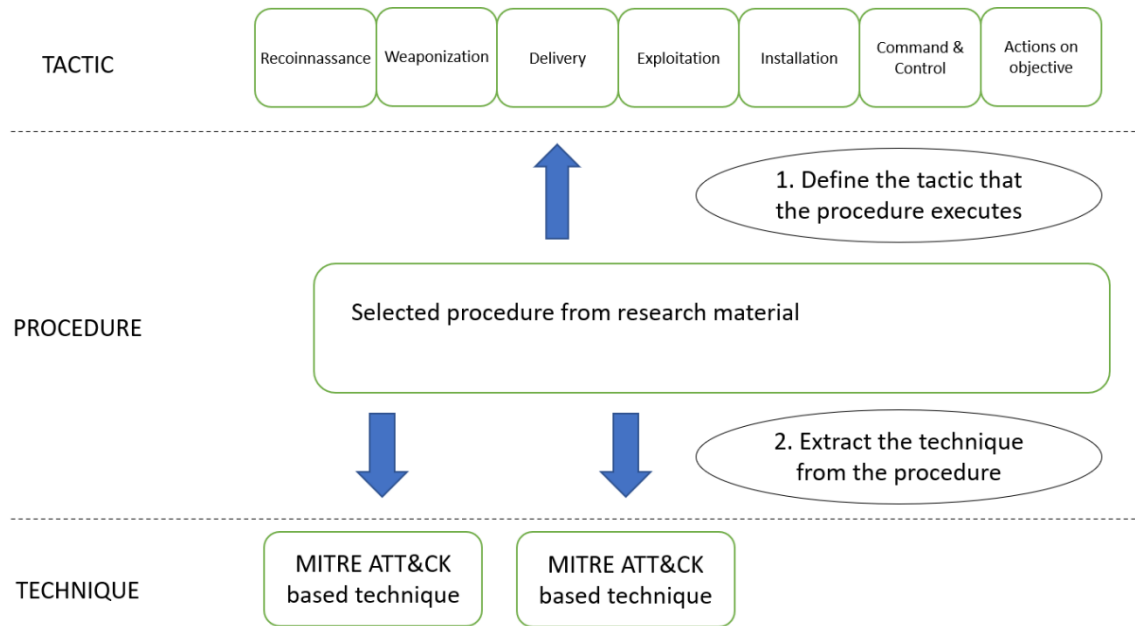


FIGURE 19 TTP Mapping

5.3 Execution of the Study

The research process of the thesis is presented in this chapter. First, the process of gathering research material is presented. Second, the execution of CKC & MITRE ATT&CK driven content analysis is explained. The process follows Tuomi and Sarajärvi (2018) steps of conducting content analysis (modified from Tuomi & Sarajärvi, 2018, chapter 4.1):

1. Decide what is interesting, especially in your research material, and make a strong decision.
2. Go through the research material, separate, and mark the things included in your interest.
3. Everything else is not included in your study.
4. Collect the marked things together and separate them from the rest of the research material.
5. Classify, make themes, and categorize the research material.
6. Write a summary.

5.3.1 Research Material

The information gathered by Thailand Computer Emergency Response Team (ThaiCERT) was chosen as the source of primary research material of this thesis

(see figure 20).²⁶ There are few reasons why information collected by ThaiCERT was utilized. First, ThaiCERT has an online portal of cyber threat groups which is updated frequently. Second, the material collected to the portal comes from multiple recognizable sources: MISP Threat Actors galaxy, MITRE ATT&CK Framework, Malpedia and AlienVault Open Threat Exchange in addition to ThaiCERT's own Cyber Threat Intelligence archive and extensive searches on the Internet. Third, the portal has a search function which enables different inquiries.²⁷ Even though APT actors can change their TTPs at any moment, evaluating the long trend is important to gain a better overall understanding. This of course puts a challenge for the use of research material. It helps if the data source is a living document having the historical and possible fresh, updated information of APT actors available. ThaiCERT's portal fulfilled the above-mentioned criteria.

Altogether the portal had listed 418 threat groups including 132 groups attributed to China in May 2022. ThaiCERT's portal's search function was used with two entries "Source country" "China" and "Free text search" "2021".

← ↻ 🏠 🔒 <https://apt.etchda.or.th/cgi-bin/aptsearch.cgi> 🔍

ETDA สำนักงานพัฒนาธุรกรรมทางอิเล็กทรอนิกส์
Electronic Transactions Development Agency

Groups Tools Search Statistics

Home > Search

Threat Group Cards: A Threat Actor Encyclopedia

Database search

Actor

Source country

Victim country or Worldwide

Victim sector

Motivation

Free text search (can use '*' and '?' wildcards)

FIGURE 20 ThaiCERT Portal

These entries were chosen based on the learned function that the portal lists all China related threat groups' information containing the string "2021". Thus, two criteria were fulfilled: the threat groups were attributed to China and relatively

²⁶ For information about ThaiCERT see e.g., <https://apt.etchda.or.th/cgi-bin/aptgroups.cgi>. ThaiCERT is part of the Electronic Transactions Development Agency (ETDA) that operates under supervision of Ministry of Information and Communication Technology in Thailand.

²⁷ <https://apt.etchda.or.th/cgi-bin/aptsearch.cgi>

recent information, in our case during the year 2021, was retrieved. The result was 36 listed China related threat groups out of which 34 was recognized as APTs, 1 other, and 1 unknown by ThaiCERT respectively (see figure 21)

All groups from China and whose card contains the string *2021*



Changed	Name	Country	Observed
APT groups			
	APT 31, Judgment Panda, Zirconium		2016-Feb 2022 🔥
	APT 41		2012-Aug 2021 ⚙️
	Bronze Butler, Tick, RedSaidNight, Stalker Panda		2010-Apr 2021 ⚙️
	Calypto		2016-Aug 2021
	Chimera		2018-Oct 2019
	Earth Lusca		2019
	Emissary Panda, APT 27, LuckyMouse, Bronze Union		2010-Mar 2021
	FunnyCream		2018
	Gelsemium		2014-Jan 2021
	GhostEmperor		2020
	Hafnium		2021-Dec 2021 ⚙️
	IndigoZebra		2014
	IronHusky		2017-Aug 2021
	Ke3chang, Wxen Panda, APT 15, GREY, Playful Dragon		2010-May 2020
	Leviathan, APT 40, TEMP.Periscope		2013-Jul 2021 ⚙️
	Mikrooson		2017-Mar 2021
⚙️	Mustang Panda, Bronze President		2014-Mar 2022 🔥
	Operation EmailThief, TEMP_Heretic		2021
	Operation PseudoManuscript		2021
	PKPLUG		2016-Mar 2021
	Polson Carp, Evil Eye		2018-Mar 2021 ⚙️
	RedDelta		2020-Feb 2022 🔥
⚙️	RedFootrot		2014-Aug 2021
	SharpPanda		2018
	Stone Panda, APT 10, menuPass		2006-Feb 2022 🔥 ⚙️
	TA413		2019-Jan 2021
	TA428		2013-May 2021
	TAG-22		2021
	TAG-28		2021
	TAG-38		2021
	TaskMasters		2010-May 2021
	Tonto Team, HartBeat, Karma Panda		2009-Mar 2021
	UNC215		2019
	Winnt! Group, Blackfly, Wicked Panda		2010-Mar 2021
Other groups			
	Rogue, Iron Group		2018-Apr 2021
Unknown groups			
	Phantom Panda		


36 groups listed (34 APT, 1 other, 1 unknown)

FIGURE 21 Result List of Chinese APT Groups

The results were further narrowed to those groups which had been active during the year 2021 or later because even though the string “2021” was used, the search engine gave results of groups without activity during the year 2021. The information of a single threat group is presented in a separate webpage as a “Threat Group Card” (see figure 22).

Threat Group Cards: A Threat Actor Encyclopedia

⇒ **APT group: Hafnium**

Names	Hafnium (Microsoft) Silk Typhoon (Microsoft)
Country	 China
Sponsor	State-sponsored, Ministry of State Security
Motivation	Information theft and espionage
First seen	2021
Description	<p>(Microsoft) HAFNIUM primarily targets entities in the United States across a number of industry sectors, including infectious disease researchers, law firms, higher education institutions, defense contractors, policy think tanks, and NGOs.</p> <p>HAFNIUM has previously compromised victims by exploiting vulnerabilities in internet-facing servers, and has used legitimate open-source frameworks, like Covenant, for command and control. Once they've gained access to a victim network, HAFNIUM typically exfiltrates data to file sharing sites like MEGA.</p> <p>In campaigns unrelated to these vulnerabilities, Microsoft has observed HAFNIUM interacting with victim Office 365 tenants. While they are often unsuccessful in compromising customer accounts, this reconnaissance activity helps the adversary identify more details about their targets' environments.</p> <p>HAFNIUM operates primarily from leased virtual private servers (VPS) in the United States.</p> <p>(Recorded Future) Coalition officials pinned the attacks on groups tracked as APT 31, Judgment Panda, Zirconium and Leviathan, APT 40, TEMP.Periscope by cybersecurity experts, according to a press release from the UK National Cyber Security Centre. Supporting statements were also issued by NATO, the UK government, the European Union Council, Australia, Japan, Canada, Latvia, Lithuania, Estonia, Slovenia, Finland, and Denmark.</p>
Observed	Countries: Worldwide.
Tools used	4 MS Exchange 0-days.
Operations performed	Dec 2021 Log4Shell attacks expand to nation-state groups from China, Iran, North Korea, and Turkey < https://therecord.media/log4shell-attacks-expand-to-nation-state-groups-from-china-iran-north-korea-and-turkey/ >
Counter operations	Jul 2021 The United States, Joined by Allies and Partners, Attributes Malicious Cyber Activity and Irresponsible State Behavior to the People's Republic of China < https://www.whitehouse.gov/briefing-room/statements-releases/2021/07/19/the-united-states-joined-by-allies-and-partners-attributes-malicious-cyber-activity-and-irresponsible-state-behavior-to-the-peoples-republic-of-china/ >
Information	< https://www.microsoft.com/security/blog/2021/03/02/hafnium-targeting-exchange-servers/ > < https://therecord.media/white-house-formally-blames-chinas-ministry-of-state-security-for-microsoft-exchange-hack/ > < https://www.dropbox.com/s/dwiygk49pos4vqx/Whitepaper%204%20MS%20Exchange%200-days.pdf?dl=0 >
MITRE ATT&CK	< https://attack.mitre.org/groups/G0125/ >

Last change to this card: 26 April 2023

FIGURE 22 APT Actor's Threat Group Card

The Threat Group Card contains different information of a threat group and links to additional information. One of the listed information is “Observed” “Countries” meaning the countries in which APT group has been active. A decision was made to further narrow the research material into those Chinese APT groups which had operated at least in one European country. Sources of information on the Threat Group Cards are divided under two headlines: “Operations performed” and “Information”. After further examination of the links and their

content, a decision was made to use links from the Information section with the exception that if the information provided by the links under the Information headline was not sufficient, the research material of a particular APT group was fulfilled from the operations performed section's links.

The gathered information from the links "Information" and "Operations performed" were used as a source of our thesis's research material. The result of data gathering was 11 Chinese APT groups, 41 links to information from public sources, which were mainly produced by cyber security practitioners, and 673 pages of text (see Appendix 1). The results (documents and records of Chinese APT groups) were compiled to a single PDF file for further examination.²⁸

5.3.2 CKC & MITRE ATT&CK Driven Content Analysis

A spreadsheet was created for the first round of coding of the research material according to CKC-framework's seven attack phases: Reconnaissance, Weaponization, Delivery, Exploitation, Installation, Command & Control and Actions on Objectives. More specifically, the spreadsheet was divided into seven different sheets and each sheet was named after one attack phase. In addition, each separate seven sheets were divided into two different columns: "Information" and "Source".

A decision was made that one of the authors coded the first six (APT10, APT27, APT31, APT41, Hafnium, Mustang Panda) Chinese APT groups' campaigns and the other author the remaining five (RedDelta, TA413, Temp_Heretic, TontoTeam, Winnti) into two separate spreadsheets utilizing the same structure as described in figure 23. The research material was examined, and different statements/sentences/word queues were mapped to different, according to CKC cyber-attack phases. For example, a sentence: "Initial attack methods include phishing to gain entry into target organization networks." was mapped under the "Delivery" phase expressing how a weapon was transferred i.e., delivered to targeted environment as highlighted in CKC-framework. Whereas a statement: "When the malicious RTF attachment named "Covdi.rtf" is executed, it exploits a Microsoft Equation Editor vulnerability and installs an embedded malicious RTF object in the form of a Windows meta-file (WMF) to the file directory %\AppData\Local\Temp\wd4sx.wmf." gave answers to the "Exploitation" and "Installation" phases.

After all the material was coded, a workshop was arranged to gather all the coded material to a single spreadsheet with the same structure as seen in figure 23. When all the chosen and coded material was in a one spreadsheet, the authors decided to analyse the material separately. The idea was to see what kind of topics raise up without communication in between the analysis rounds. Later,

²⁸ The authors have made a copy of all the used research material in case the information is deleted from the public Internet. The copy can be delivered upon request.

another workshop was scheduled in which both authors presented their own analysis to each other.

Information	Source
CTU researchers assess it is highly likely that the BRONZE UNION threat group gathers defense, security, and political intelligence from organizations around the world.	BRONZE UNION Cyberespionage Persists Despite Disclosures Secureworks
BRONZE UNION activity on multiple U.S.-based defense manufacturer networks included the threat actors seeking information associated with aerospace technologies, combat processes, and naval defense systems.	BRONZE UNION Cyberespionage Persists Despite Disclosures Secureworks
continued focus on information that would be of interest to individuals or groups living in a country that has a significant manufacturing base and a strategic interest in U.S. military capabilities	BRONZE UNION Cyberespionage Persists Despite Disclosures Secureworks
In 2016, CTU researchers observed the group using native system functionality to disable logging processes and delete logs within a compromised environment. The threat actors used the command-line tool to unlock and disable the default logging component on the server (system.webServer/httplogging) and then delete existing logs from the system (see Figure 4)." CTU researchers assess it is highly likely that the BRONZE UNION threat group gathers defense, security, and political intelligence from organizations around the world.	Threat Group-3390 Targets Organizations for Cyberespionage Secureworks
The threat actors have used the Baidu search engine, which is only available in Chinese, to conduct reconnaissance activities.	Threat Group-3390 Targets Organizations for Cyberespionage Secureworks
CTU researchers have observed the threat group obtaining information about specific U.S. defense projects that would be desirable to those operating within a country with a manufacturing base, an interest in U.S. military capability, or both.	Threat Group-3390 Targets Organizations for Cyberespionage Secureworks

Navigation: Reconnaissance | Weaponization | Delivery | Exploitation | Installation | Command & Control | Actions on Objectives | Total Values ...

FIGURE 23 Procedure Categorization

After all the material had been gone through and the data had been categorized based on CKC's attack phases, the researchers built a new spreadsheet. In this spreadsheet every single procedure was counted and gathered to display how the CKC phases were presented in the reports. To support qualitative analysis, the number of codes or elements belonging to different tactics can be calculated. Hence, from this summary the researchers were able to see which reports had information that corresponds with any of the CKC phases. Figure 24 illustrates what kind of tactics are present in a single report. Thus, the researchers were able to point out which phase is most utilized among chosen reports (research material).

Source Material	Pages	Tactics							Total appearances
		Reconnaissance	Weaponization	Delivery	Exploitation	Installation	C2	Actions on object	
ESET_Winnti	27	0	7	0	1	6	3	0	17

FIGURE 24 Tactics Summary

The next step was to utilize MITRE's ATT&CK framework to categorize different techniques used in these cyber-attacks. Researchers labelled every procedure gathered during the previous step with techniques that correlated the best with the procedure. A single procedure can be seen utilizing multiple different techniques, thus there can be multiple technique codes for a single procedure. The procedure "The group extensively uses long-running strategic web compromises (SWCs), and relies on whitelists to deliver payloads to select victims." contains two techniques, MITRE code-T1189 (Drive-by compromise) and MITRE code-T1199 (Trusted Relationship). Even though ATT&CK framework also utilizes

subcategories for some techniques, the researchers made a conscious decision not to use them as it would have gone too specific and made the results too complicated. Figure 25 illustrates how techniques and tactics were extracted from a single procedure.

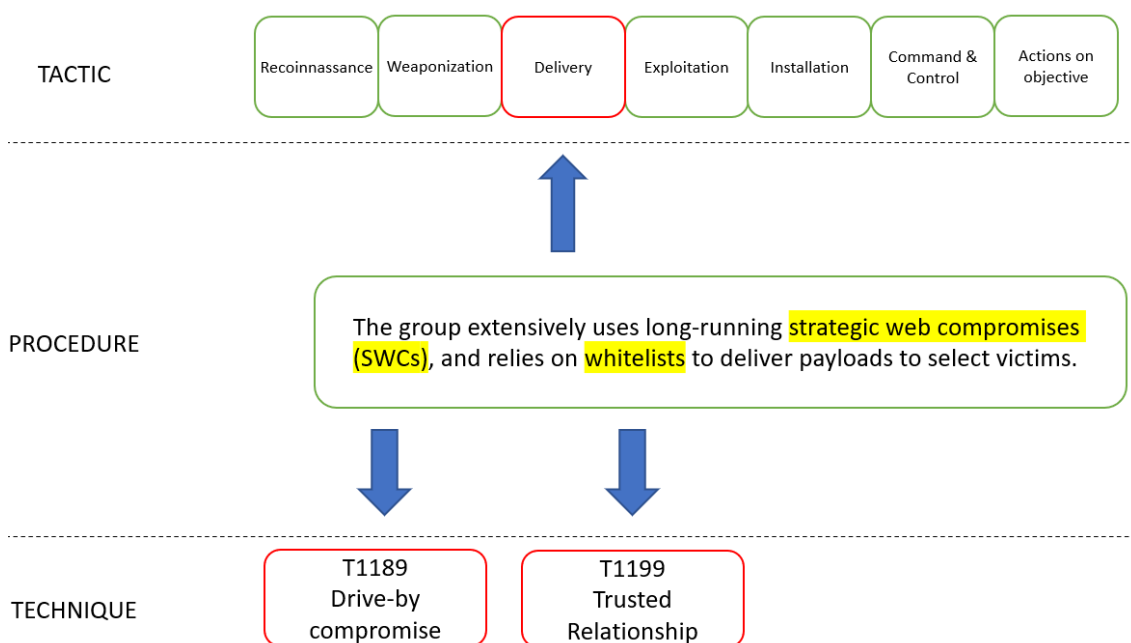


FIGURE 25 Tactic & Technique Categorization

After the technique categorization, the information was inserted into the same spreadsheet used in the tactic categorization (see figure 26). There were also a few procedures where the researchers did not find matching technique from the ATT&CK framework.

Information	Source	Techniques used	
The group extensively uses long-running strategic web compromises (SWCs), and relies on whitelists to deliver payloads to select victims.	Threat Group-3390 Targets Organizations for Cyberespionage Secureworks	T1189	T1199

FIGURE 26 Technique Categorization

When all the procedures were analysed with the best corresponding techniques, all the data were gathered into one summary spreadsheet. Every single technique label was then counted from the matching CKC phase and displayed in the spreadsheet. The results were then inspected by the researchers and the most used technique in a CKC phase was highlighted, making it easier to find. The percentage amount of each technique on a certain phase was also counted and displayed, making it easier to find and compare results. Figure 27 shows a part of the techniques categorization results and a highlighted area of the most used technique in the context of a specific CKC phase.

Technique	Technique name		Reconnaissance		Weaponization		Delivery	
T1591	Gather Victim Org Information		7	37%	0	0%	0	0%
T1589	Gather Victim Identity Information		1	5%	0	0%	0	0%

FIGURE 27 Summary Spreadsheet

5.3.3 Reliability and Validity

This thesis combines qualitative and quantitative research even though the baseline of the thesis is qualitative because the authors have ended up using textual analysis which is enriched with qualitative coding. According to Saaranen-Kauppinen and Puusniekka (2006) reliability assessment is a key part of scientific research, as certain norms and values have been set for the research, which it should strive for. Reliability and validity have traditionally been central concepts in reliability issues when assessing the reliability of quantitative research measurements. Perceptions of the appropriateness of the concepts of reliability and validity in the evaluation of qualitative research vary. The validity and reliability of qualitative research cannot be evaluated in exactly the same way as quantitative research (Eskola & Suoranta 1998, 208-222; Mäkelä 1990). In short, validity is about whether the research is valid; has it been done thoroughly, are the results obtained and the conclusions drawn "correct" (Saaranen-Kauppinen & Puusniekka 2006). Kirk and Miller (1986) identify three points for assessing the reliability of qualitative research: Assessment of the reliability of a specific method (quixotic reliability); under what circumstances a method is reliable and consistent. Diachronic reliability refers to the stability of measurements or observations over time. Consistency in results (synchronic reliability) obtained simultaneously by different instruments.

As stated on chapter 2, the cyber security industry has a monopoly on the materials concerning APT groups and thus the academia cannot independently verify it. Therefore, academia must rely on the material presented by the cyber security industry and accept the report as it is along with its preformed attribution. According to Alastalo and Vuori (2023), almost without exception, documents have been created for a completely different purpose than research. Therefore, the analysis must pay special attention to the context in which they originate and try to understand how it affects how people and things are portrayed in documents. Institutional documents are often governed by strict rules and conventions about what they contain and the style in which they are produced.

How can the reports be trusted then? The ThaiCERT portal was chosen as the source of research material for its usability and because Thailand's Computer Emergency Response Team, that handles cyber security breaches on a national scale had gathered the information. The Researchers decided that this would give validity for the research material gathering. In many cases private cyber security companies' reports are laid out and used as marketing material for companies' marketing purposes. Thus, the writers must only guess what information cyber security companies want to reveal and what information is left out due to reasons such as the companies do not want to reveal their sources of information and what their own tools are capable of spotting in the end. As we have stated in

Chapter 4, attribution is an issue in cyber security and in some of the analysed reports attribution methods are left out and the reader must rely on that the attackers are indeed Chinese. Also, the amount of revealed information in these reports vary. In some reports the used TTPs were described more thoroughly when in some reports TTPs were only described on a higher level.

Although there are conflicting views on the possibility of using traditional concepts (validity, reliability) in qualitative research, there is still a general consensus on the means of improving reliability: for example, when analysing texts, it would be good to try to make well-founded and open-ended categorizations and coding (Saaranen-Kauppinen & Puusniekka 2006). In this thesis deductive qualitative content analysis was used based on the CKC-framework which can be seen as the guiding theory. The CKC division was then enriched with MITRE's more specific techniques mapping. Both frameworks are generally accepted in the cyber security field, often referred, and used in different ways and occasions. Even though during the coding phase of the thesis the authors had mixed feelings using the predetermined cyber-attack phases, CKC-framework was used consistently the whole time. The coding process also revealed how difficult it is to decide beforehand how certain cyber actors operate.

While coding, the researchers decided to split the material in half and just code their own half. This hurts the reliability of the thesis, as the material is now inconsistently coded. It leans to one researcher making own decisions about the coding and the other researcher his own. If the coded material is combined, it has been coded using two different opinions. This goes against Potter and Levine-Donnersteins idea, that when using multiple coders, there should be at least an overlap of the coded material to improve the reliability of the study (Potter & Leviene-Donnerstein, 1999). Another way would have been for both researchers to code the whole material and then combine their coding and eliminate the duplicates. Still the researchers made a conscious decision to split up the material due to the sheer size of the it.

We have fallen to the trap of modifying taxonomy to our own use and based on our own research when putting together CKC and ATT&CK frameworks. We have also used our own interpretation of tactics, techniques and procedures but the interpretation is well-defined and used consistently. There are other interpretations, but these are also presented as a source of our own perception. According to Potter and Levine-Donnerstein (1999), validity is received in two steps. The first is a reliable coding scheme that guides research material coding, and the coders stick with the theory guiding the process. The second is self-assessment by the coders against some standard i.e., if the decisions made throughout the coding are reflected based on the chosen standard, then the coding is valued as creating proper data.

The secondary coding was done by only one of the authors. Because the second coding round lacks the other author, it can be stated that the coding is an opinion of a single author. This hurts the reliability as the other author may have had different opinions on the coding.

When utilizing theory-driven content analysis predetermined frameworks were used and it was clear what to look for from the research material. If inductive reasoning would have been used, the results might have been different due to the reason that in this case the research material would have revealed us better what the sampling has to say instead of predetermined division of things. The authors have tried to take care of this controversy by adding additional chapter handling on additional findings from the research material even though still based on theories decided to use beforehand.

In general, the repeatability of our examination could be an issue due to changing information and future availability of the used research material from the public Internet. The authors have taken care of this by composing a separate pdf file consisting of all the research material. The use of CKC and MITRE ATT&CK frameworks enables repeatability of the used research material with well-known attack modelling. On the other hand, the mapping of the research material is ambiguous and may end up in different kind of results. Hence, it would be interesting to see the results of some other with the same method and research material. The results of this thesis are heavily dependent on coding of the research material. However, certain obvious results of the thesis might indicate that if other would repeat this study, the results might be close to the ones the authors have received.

6 RESULTS

The research question of the thesis is: What are Chinese APT actors' tactics, techniques, and procedures? By analyzing 41 different documents relating to Chinese APT groups, we were able to discover characteristics of Chinese APT cyber-attacks using the CKC and MITRE ATT&CK frameworks. In this chapter, the results are presented. In our theoretical model the tactics and techniques are based on procedures which describe how the tactics and techniques were used. CKC cyber-attack phases are seen as tactics and the specific techniques used in each tactical phase are analysed with MITRE's attack framework's technique labelling. The results are presented following CKC frameworks cyber-attack phase categorization. In the end, some additional findings are brought up, which arose from the analysis. In the results, after a technique is mentioned, a matching MITRE code is presented in brackets. This code is from the MITRE ATT&CK framework's techniques. Code T1591 would refer to Gather Victim Org Information-technique. Also in the results, there are procedure examples to show how techniques were used in the reports. The sources of these examples can be found from appendix 1, where every used report is referenced. After every procedure, there is a mention in brackets for example APT27/1. This refers to the report about APT27 and the numerical order of the report used to extract procedures.

6.1 Overall CKC-phases

The material that the researchers went through and coded contained 41 different reports from different authors. There were three reports that did not contain any procedures relating to different CKC-phases. The results of all the phases are presented on table 14. Out of the 38 articles that did contain information about the CKC-phases, 33 had one or more procedures that were categorized under the Weaponization phase. This phase was the most present on the material.

TABLE 14 Appearances of CKC-phases in the Research Material

CKC-phase	Unique appearances of a CKC phase	Percentage
Reconnaissance	10	24,4%
Weaponization	33	80,5%
Delivery	27	65,9%
Exploitation	25	61%
Installation	27	65,9%
Command & Control	26	63,4%
Actions on objectives	16	39%

6.1.1 Reconnaissance

Table 15 presents the ten most used techniques during the Reconnaissance phase. The most used technique (36,8%) in this phase was Gathering Victim Organization Information (T1591), in which the attacker utilizes open and other sources to gather information about the target as described in the following procedure:

CTU researchers have observed the threat group obtaining information about specific U.S. defense projects that would be desirable to those operating within a country with a manufacturing base, an interest in U.S. military capability, or both. (APT27/3)

This information can be utilized to get a better understanding of the target and help the attacker to make decisions on the following steps. Other techniques that rose from the general mass were Search Open Websites/Domains (T1593) and Phishing for Information (T1598). The former being almost invisible to the defenders, though some hints might surface like seen in the following procedure:

The threat actors have used the Baidu search engine, which is only available in Chinese, to conduct reconnaissance activities. (APT27/3)

Phishing for information (T1598) is more noticeable than Search Open Websites/Domains (T1593), but still might not pop out from the basic mass if obfuscated correctly. The usage of this technique is seen in the following procedure:

The initial phase was aimed at reconnaissance and involved emails designed to simply track if a target received and opened the messages. (TEMP_Heretic/1)

TABLE 15 Ten Most Used Reconnaissance Techniques

Reconnaissance				
Order	Appearances	MITRE Code	Technique Name	Percentage
1	7	T1591	Gather Victim Org Information	36,8%
2	3	T1593	Search Open Websites/Domains	15,8%
2	3	T1598	Phishing for Information	15,8%
3	1	T1589	Gather Victim Identity Information	5,3%
3	1	T1090	Proxy	5,3%
3	1	T1590	Gather Victim Network Information	5,3%
3	1	T1588	Obtain Capabilities	5,3%
3	1	T1070	Indicator Removal on Host	5,3%
3	1	T1190	Exploit Public-Facing Application	5,3%

6.1.2 Weaponization

Table 16 presents the ten most used techniques during the Weaponization phase. The most used technique (24,6%), that the researchers found from the weaponization phase, was Develop Capabilities (T1587). This would mean that the APT groups themselves possess the capabilities of developing their own custom malware and if needed, retooling them if the tool becomes too detectable. This is apparent from the following procedures:

Another favorite Winnti technique is theft of certificates for code signing. Compromised certificates are used to sign malicious files intended for future attacks. (APT41/5)

and

The Mustang Panda APT actor uses PlugX with minor changes, in an attempt to evade detection. (Mustang Panda/4)

The second most utilized technique (17,3%) that was noticed was Obtain Capabilities (T1588). This technique shows that the APT groups use tools that are publicly available. One of the tools that the researchers noted is a popular penetration testing toolkit called Cobalt Strike. A procedure describing this:

APT41's tactics, including their use of malicious documents, exploits and Cobalt Strike. (APT41/3)

The ATP groups can also copy or steal exploits from other groups like mentioned in the following procedure:

A far more probable scenario is that APT31 copied the exploit from Equation Group. (APT31/3)

and

Stealing private keys or compromising an organization's infrastructure to access and steal digital certificates abuses trust relationships between firms and certificate authorities. Malicious files signed with valid digital certificates can circumvent automated scanning/blocking solutions and bypass Windows group policies which restrict unsigned code from running. (APT41/7)

The third most utilized technique (14,6%) according to the results was Server software components (T1505). This technique is the single most utilized technique that was observed in this study. It can be explained as web shells and different backdoors fall under this technique. Researchers found procedures like:

Tiny Shell has been used by multiple threat actors since several years now and it is not surprising to see APT31 using it. (APT31/5)

pointing out that different APT groups weaponize web shells for their usage.

TABLE 16 Ten Most Used Weaponization Techniques

Weaponization				
Order	Appearances	MITRE Code	Technique Name	Percentage
1	64	T1587	Develop Capabilities	24,6%
2	45	T1588	Obtain Capabilities	17,3%
3	38	T1505	Server Software Component	14,6%
4	12	T1204	User Execution	4,6%
5	10	T1608	Stage Capabilities	3,9%
5	10	T1027	Obfuscated Files or Information	3,9%
7	7	T1036	Masquerading	2,7%
8	6	T1583	Acquire Infrastructure	2,3%
9	5	T1059	Command and Scripting Interpreter	1,9%
9	5	T1566	Phishing	1,9%

6.1.3 Delivery

Table 17 presents the ten most used techniques during the Delivery phase. The most common technique (40,6%) that the researchers noticed was Phishing (T1566). This technique contains both phishing and spear phishing. With phishing, an email is sent to the target with either a malicious attachment or a malicious link and the aim is to get the user to either open the attachment or click on the link. Spear phishing is similar than phishing but with exception that its customized to the target. The usage of this technique can be seen with the following procedures found from the material:

APT41 has often used phishing emails with malicious attachments as an initial infection vector. (APT41/3)

and

the ZIP file is likely to have been delivered via a spearphishing email. (RedDelta/1)

The second most common technique (16,4%) found from this phase was Drive-by Compromise (T1189). Drive-by Compromise is also known as strategic web compromise (SWC) or with a more familiar name of Watering Hole Attack. This technique relies on the target navigating themselves to either a compromised website or a website that was specifically set up for a certain attack. Target unwillingly downloads the exploit from the website like stated in the procedure:

The attackers used a phishing website masquerading as the Huawei company career page to target people working in the telecommunications industry. (RedDelta/3)

or redirects to another compromised website:

it was delivered via a redirect from a strategic web compromise (SWC). (APT27/1)

TABLE 17 Ten Most Used Delivery Techniques

Delivery				
Order	Appearances	MITRE Code	Technique Name	Percentage
1	52	T1566	Phishing	40,6%
2	21	T1189	Drive-by Compromise	16,4%
3	7	T1195	Supply Chain Compromise	5,5%
4	6	T1036	Masquerading	4,7%
5	5	T1587	Develop Capabilities	3,9%
5	5	T1505	Server Software Component	3,9%
5	5	T1090	Command and Script Interpreter	3,9%
8	4	T1133	External Remote Services	3,1%
8	4	T1091	Replication Through Removable Media	3,1%
10	3	T1204	User Execution	2,3%

6.1.4 Exploitation

Table 18 presents the ten most used techniques during the Exploitation phase. The most common technique (15,6%) used during this phase was Command and Scripting Interpreter (T1059). This technique leverages different commands and script interpreters to achieve a desired outcome for the attacker. This was identified from the procedures:

The JavaScript code used to facilitate mail theft has to be customized per version of Zimbra, as the attacker needs to request a page containing a CSRF-Token in order to make subsequent requests to steal mail data. (TEMP_Heretic/1)

and

The .lnk file uses an embedded VBScript component to retrieve a decoy PDF file and a PowerShell script from the adversary-controlled web page. (Mustang Panda/1)

The second most common technique (13,8%) was Exploitation for Client Execution (T1203). In Exploitation for Client Execution, the actor uses different vulnerabilities on used software to execute desired code. An example procedure of this is:

They sent a malicious RTF document to the targets with an exploit targeting the CVE-2018-0798 (Microsoft's Equation Editor vulnerability). The purpose of the shellcode was not to execute the malware (as it is usual) but simply to drop it in the %APPDATA%\microsoft\word\startup\ repository with the .wll extension. (Tonto Team /2)

The third most common technique (10,1%) that the researchers found was User Execution (T1204). This technique might be seen to overlap with the previous step, Delivery. The delivery step can be spearphishing, but the exploitation

requires the target to click on something. This is highlighted on the following procedure:

“The biggest difference between the second and third lures is that one uses a self-extracting archive named “India records highest ever single day covid_19 recoveries.pdf.exe,” and the other uses a ZIP file named “India records highest ever single day COVID-19 recoveries.zip.”” (APT41/3)

TABLE 18 Ten Most Used Exploitation Techniques

Exploitation				
Order	Appearances	MITRE Code	Technique Name	Percentage
1	17	T1059	Command and Scripting Interpreter	15,6%
2	15	T1203	Exploitation for Client Execution	13,8%
3	11	T1204	User Execution	10,1%
4	8	T1078	Valid Accounts	7,3%
5	7	T1587	Develop Capabilities	6,4%
6	6	T1190	Exploit Public-Facing Application	5,5%
7	5	T1505	Server Software Component	4,6%
8	4	T1219	Remote Access Software	3,7%
9	3	T1014	Rootkit	2,8%
10	2	T1070	Indicator Removal on Host	1,8%

6.1.5 Installation

Table 19 presents the ten most used techniques during the Installation phase. The most common technique (26%) used during this phase was Server Software Component (T1505). Like stated earlier in the Weaponizing phase, this technique incorporates web shells, backdoors and RATs. Researchers found procedures like:

They then identify the Exchange server and attempt to install the OwaAuth web shell. (APT27/3)

and

HAFNIUM operators deployed web shells on the compromised server. (Hafnium/1)

Hijack Execution Flow(T1574) technique was the second most used technique (10,7%) that the researchers found in the Installation phase. Hijack Execution Flow contains techniques like DLL Side-loading or DLL Hijacking as apparent in following procedures:

PlugX executes DLL hijacking with benign applications such as ESET antivirus, AdobeUpdate etc. (Mustang Panda/4)

and

the main purpose of this stage of the malware is to perform the DLL sideloading step in order to execute the PlugX variant. (RedDelta/1)

TABLE 19 Ten Most Used Installation Techniques

Installation				
Order	Appearances	MITRE Code	Technique name	Percentage
1	44	T1505	Server Software Component	26,0%
2	18	T1574	Hijack Execution Flow	10,7%
3	9	T1140	Deobfuscate/Decode Files or Information	5,3%
4	8	T1036	Masquerading	4,7%
5	7	T1053	Scheduled Task/Job	4,1%
6	6	T1547	Boot or Logon Autostart Execution	3,6%
6	6	T1071	Application Layer Protocol	3,6%
8	5	T1106	Native API	3,0%
8	5	T1059	Command and Scripting Interpreter	3,0%
8	5	T1543	Create or Modify System Process	3,0%

6.1.6 Command and Control

Table 20 presents the ten most used techniques during the Command-and-Control phase. The most used technique (26,6%) during this phase was Application Layer Protocol (T1071). This technique consists of different web protocols like HTTP or DNS and the aim is to blend the traffic to the existing traffic and thus not arouse any suspicion on the defender side. Utilizing common ports, even though the used protocol might not correspond with the port in question. Example procedures for this are:

separate C&C domains were assigned to each targeted company. Virtually all the C&C domains were arranged as follows: a second-level domain was created without a DNS A-record, i.e., there was no IP address assigned to it. (Winnti Group/4)

and

To traverse the firewall, C2 traffic for most TG-3390 tools occurs over ports 53, 80, and 443. The PlugX malware can be configured to use HTTP, DNS, raw TCP, or UDP to avoid network-based detection. (APT27/3)

Another technique (12,0%) that was utilized during this phase was Server Software Component (T1505). This technique was already present in earlier phases,

and it's also present in this stage as the web shells and backdoors. Example procedures for this technique are:

when they detected PlugX malware command and control (C&C) servers, operated by the Mustang Panda group, communicating with hosts inside the networks of the Indonesian government. (Mustang Panda/2)

and

The Cobalt Strike Beacon implant beacons to the command-and-control (C2) IP address, which is used to remotely control the implant. (Mustang Panda/1)

In these previously mentioned procedures were tools like PlugX and Cobalt strike that fall into the category of T1505.

The third most popular technique (9,5%) during this phase was Web Service (T1102). This technique utilizes present web services in the attackers C2 communications. These web services might be commonly known, for example GitHub or Google Docs, but the usage is malicious. Known web services are used to blend the C2 traffic to the normal traffic and allow the attacker to stay hidden. Utilizing legitimate sites can be used to bypass network defenses as seen in the procedure:

bypassing network defenses by employing only legitimate websites and services to host their implants (GitHub) and interact with them once executed on the victims' workstation (use of DropBox API). (APT31/5)

Another noticeable procedure was:

A similar technique has been used by Winnti in the past: according to Trend Micro, an encoded C2 address was stored in GitHub repositories in 2017. (APT41/5)

where it is stated that a legitimate site held C2 addresses, and this could have allowed the attackers to modify their infrastructure more easily.

TABLE 20 Ten Most Used Command & Control Techniques

Command & Control				
Order	Appearances	MITRE Code	Technique Name	Percentage
1	42	T1071	Application Layer Protocol	26,6%
2	19	T1505	Server Software Component	12,0%
3	15	T1102	Web Service	9,5%
4	11	T1573	Encrypted Channel	7,0%
5	8	T1583	Acquire Infrastructure	5,1%
5	8	T1568	Dynamic Resolution	5,1%
7	6	T1036	Masquerading	3,8%
7	6	T1095	Non-Application Layer Protocol	3,8%
9	5	T1090	Proxy	3,2%
10	4	T1588	Obtain Capabilities	2,5%

6.1.7 Actions on Objective

Table 21 presents the ten most used techniques during the Actions on Objective phase. Most used technique (10,9%) during this phase was Server Software Components (T1505). This technique consists of different tools and their usage in Actions on Objective phase. A good example is procedure:

In other intrusions, data was exfiltrated using the PlugX remote access tool. (APT27/3)

where a tool that is used to exfiltrate data is mentioned. Another example is procedure:

BRONZE PRESIDENT has demonstrated intent to steal data from organizations using tools such as Cobalt Strike, PlugX, ORat, and RCSession. (Mustang Panda/6)

where multiple different tools, that are used in this procedure, are mentioned.

The second most utilized technique (8,1%) during this phase was Data Staged (T1074). While the attackers can gather target data from multiple systems, if the exfiltration or C2 connections would come from all those systems, defenders might notice the attack easier. Thus, it is better for the attackers to gather the stolen data to a one central staging point where it can be exfiltrated stealthier. The staging can be in either a local or a remote location. Attackers can also compress the data for easier exfiltration. Good example procedures of this technique are:

From our telemetry data, we found a different way of stealing documents. It abuses RAR to search document files and compress them locally in order to send it to its CNC server. (Mustang Panda/5)

and

CTU researchers have observed TG-3390 actors staging RAR archives, renamed with a.zip file extension, on externally accessible web servers. The adversaries then issue HTTP GET requests, sometimes with the User-Agent MINIXL, to exfiltrate the archive parts from the victim's network. (APT27/3)

The third most utilized technique (7,2%) during this phase was Account Discovery (T1087). This technique enables the attackers to gather account information for future actions as seen in procedure:

BRONZE UNION uses various tools for credential theft. In one incident, the threat actor used the Wrapikatz tool (w.exe) with a usage statement that retrieves various passwords and Windows credentials from memory and compiles them in w.txt: (APT27/2)

TABLE 21 Ten Most Used Actions on Objectives Techniques

Actions on Objectives				
Order	Appearances	MITRE Code	Technique Name	Percentage
1	12	T1505	Server Software Component	10,8%
2	9	T1074	Data Staged	8,1%
3	8	T1087	Account Discovery	7,2%
4	5	T1105	Ingress Tool Transfer	4,5%
5	5	T1083	File and Directory Discovery	4,5%
6	5	T1218	System Binary Proxy Execution	4,5%
7	5	T1030	Data Transfer Size Limits	4,5%
8	4	T1036	Masquerading	3,6%
9	4	T1555	Credentials from Password Stores	3,6%
10	4	T1018	Remote System Discovery	3,6%

6.2 Additional Findings

Analyzing the techniques that were used, the researchers were able to identify that there are some techniques that are more popular than others inside the same CKC-phase. Figure 28 shows that in the CKC phases of Delivery, Installation and Command & Control, there is a significant gap between the most used and the second most used technique.

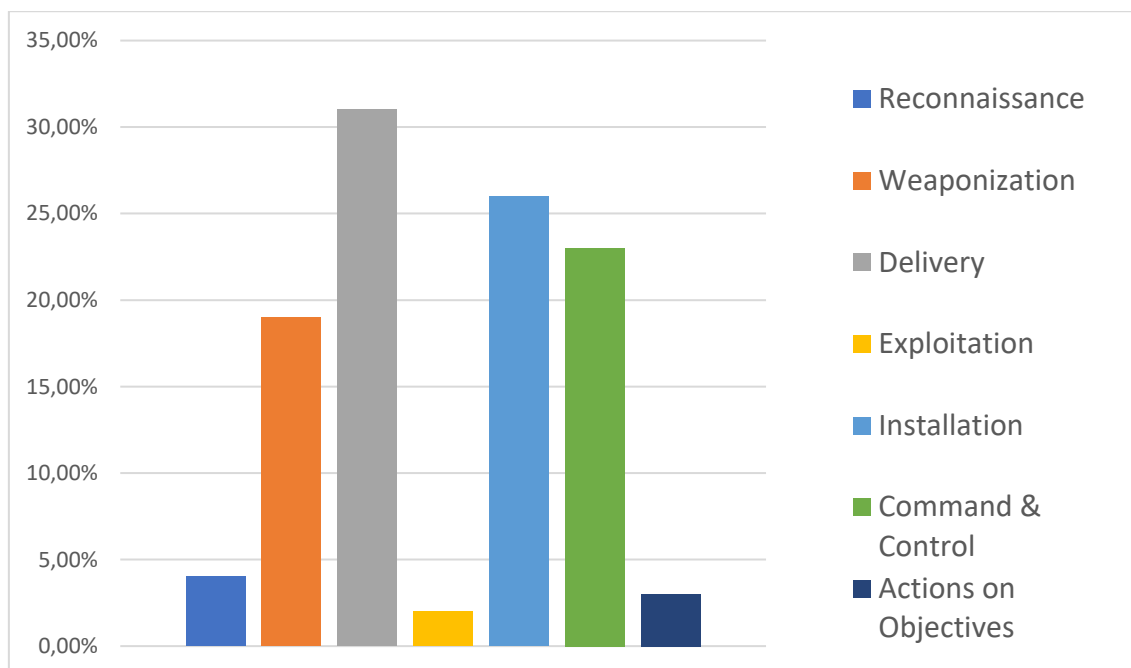


FIGURE 28 The Difference Between the Most and the Second-most Used Techniques in CKC Phases

There is also a gap in the Weaponization phase, but it is not as large as in the previously mentioned other three CKC phases. Also, as the Weaponization phase is mainly done outside the target's visibility, it is very hard to counter and not as significant as the other phase.

There is also a significant difference between the materials used in how well there was information about the attacks that could be categorized into a CKC-phase. From figure 29, it is possible to see that some of the materials had information from all the CKC-phases, while others had only from some and a few did not have anything that could be categorized to a CKC-phase.

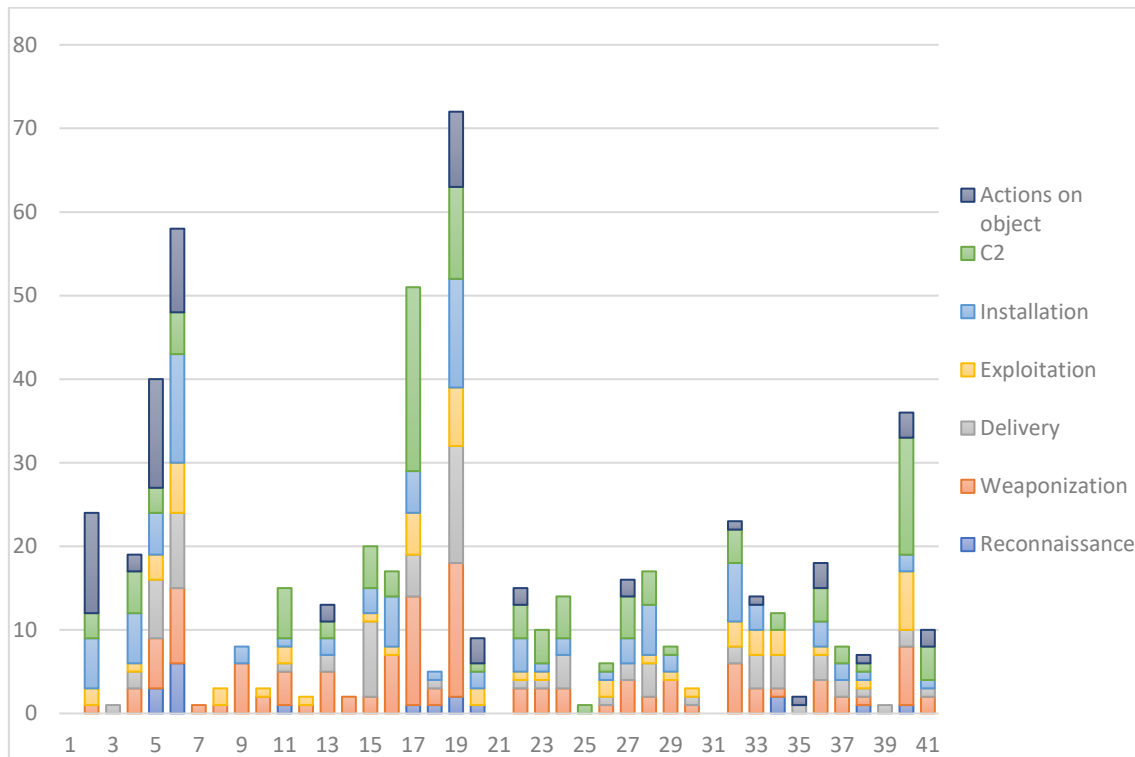


FIGURE 29 Number of Procedures Retrieved per Report Categorized by CKC-Phases

It's also worth noting that some of the materials featured multiple procedures that could be categorized into a single phase. Figure 30 points out that there were only five out of 41 reports, which were found to include all the CKC-phases.

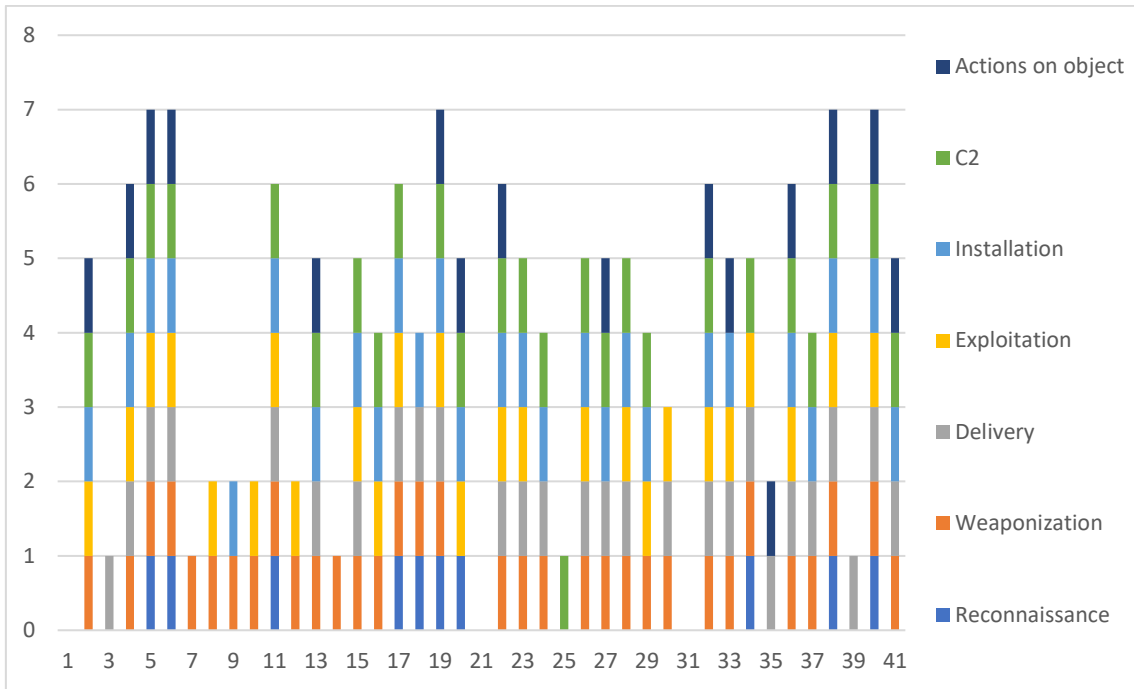


FIGURE 30 CKC-Phases Present in the Research Material

7 ANALYSIS AND DISCUSSION

The research question of this thesis was: what are Chinese APT actors' tactics, techniques and procedures? We have ended up analysing 41 different documents relating to Chinese APT groups with CKC and MITRE ATT&CK frameworks to give us answers to this question. In this chapter, the results are analysed, and further research is discussed. Our analysis follows the seven phases of CKC framework starting from Reconnaissance and ending up in Actions on Objectives.

Reconnaissance is the first phase of the CKC framework, and it is where the attack itself begins. This phase is the hardest for the defenders to get information on. There are multiple ways for the attacker to perform this phase passively, and thus only 24,4% of the research materials had procedures that were a part of this phase. The attackers can utilize time in their favour, conducting reconnaissance by interacting with the target environment, for example scanning, and after a long period of time conducting the actual attack (Ussath et al., 2016). The small number of procedures on this phase also reflects that there was only a small amount of different techniques present and only Gather Victim Organization Information (T1591) stood out. This would suggest that either the main information gathering is happening outside the defender's visibility, or the defender cannot distinguish APT-actors reconnaissance from the generic scanning and phishing activity which is happening constantly on the Internet. Also, it is possible that the APT groups themselves are so advanced, as they should be, to be able to obfuscate their reconnaissance activities. It is also possible that the reconnaissance of an APT group is constant, and it is looking for opportunities to exploit in the future if needed.

The second phase of CKC is Weaponization, which also suffers from the same issue as the previous Reconnaissance phase. This phase can be done offline and out of visibility of the defenders. Still this phase was the most present on the research materials, 80,5%. This can be explained as the materials that contained information about the weaponization usually analysed the malware used on the attacks. And this analysis was done after the attack, not during it. Thus, the information from the analysis will benefit against future attacks. But if the attackers are constantly evolving their tools like stated by Lemay et al. (2018) where they noted that the APT actors might be using common developers and new tools are developed, what good is it for the defenders? Sometimes the ATP-actors reuse their tools in later attacks, which can be identified by their hashes, or they might reuse code (Lemay et al., 2018). Also, defenders can learn from the tools and thus understand its behaviour and try to detect it.

For the Delivery phase of CKC, which is the third one, it is not surprising that the most utilized technique was Phishing (T1566). Humans can be seen as the weakest link in cyber security and correctly crafted spear phishing message might be able to trick even the most trained and security focused user. A well performed reconnaissance will eventually present spear phishing targets for the attackers, and it might not even need to be a direct attack. As not all the entities

in cyberspace have similar protection, a spear phishing attack via a known and reliable partner can be an option. The results from this phase support Bahrami et al. (2019) argument that majority of attacks during this phase are carried out using spear phishing. It is worth noting that this thesis did not differentiate between phishing and spear phishing, nor did it analyse the content of a phishing message, only the usage of a technique itself. Chen et al. (2014) had noted that Chinese APT actors tend to prefer spear phishing and watering hole attacks, also known by Drive-by Compromise (T1189) as their delivery methods. These two techniques were the majority of the observed delivery techniques in this thesis as well. The most surprising observation was that the usage of Supply Chain Compromise (T1195) technique was relatively low. After the NotPetya incident, which used supply chain to infect computers around the world, the researchers thought it might be used more, but it seems that this technique is not so popular among the Chinese APT actors.

The Exploitation phase is the fourth phase of CKC, and it was decently presented in the material, but it was surprising that only 61% contained at least one exploitation procedure. There was no single most used technique in this phase, as the three most popular techniques had 16%, 14% and 10% share of the procedures. The surprising fact was that the most popular used technique, Command and Scripting Interpreter (T1059), did not have more substantial usage. The attacker would like to hide its presence as long as possible and thus it would be preferable to not need user action, which could alert the user that something is wrong. So, utilizing automation would be the perfect solution. The second most utilized technique during this phase relies on users not updating their software. Exploitation for Client Execution utilizes vulnerabilities in different client programs. These vulnerabilities surface as they are found and are patched as soon as possible, though it is vendor dependent. The vulnerabilities can have different impacts on the target, but the aim of its usage is always the same, to complete the exploitation phase.

After a successful exploitation, the attacker moves to the next CKC phase, the Installation. This phase was also decently covered in the material, 65,9% of the research material contained procedures belonging to this phase. The most common technique was unsurprisingly Server Software Component (T1505), which contained sub techniques like web shells, backdoors and RATs. It's clear that the attacker is keen to acquire a foothold in the target system and what better way to do it than to utilize different programs that give them access to it later. During this phase it is important for the attackers, like in every other phase, not to get notice, but after this phase, they will have a presence in the target system and a possibility to move laterally to other systems. So, it is clear that the attacker wants to obfuscate their actions during this phase and thus techniques like masquerading, scheduled task/job and native API were used. The more normal the attackers' actions seem on the target systems, the better for the attacker. This is also what the second most used technique during this stage, Hijack Execution Flow, does, hide the actions of the attackers to benign applications.

When the attacker has managed to complete the installation phase, it's time to move on to the Command-and-Control phase, where the malicious software contacts attackers' infrastructure to receive further instructions on what actions should be done. This generates network traffic and creates the need for the attacker to hide their actions. Thus, the most logical thing is to blend into the normal network traffic, and this is what the technique Application Layer Protocol (T1071) represents. If an infected client starts to utilize unfamiliar or seldom used protocols, this can alert the defender security and require a closer inspection of the infected client. The utilization of this technique did not come as a surprise but what was surprising was that there was only a nine percent utilization of Web Services (T1102) technique. This technique uses legitimate web services for malicious actions, thus further obfuscating the attackers' actions. Legitimate web services may have their own security checks, but on some web services the security level might be lower. The researchers thought that this kind of obfuscation would be more popular.

The last CKC phase is Actions on Objectives, where the attacker applies actions to achieve its goals. This phase was only present in 39% of the research materials, which was surprising. This might be that the authors of the material or the victims have decided not to publish what kind of actions the attacker has done to their systems. The actions are done with different tools during this phase, which is apparent with the most utilized technique of Server Software Components (T1505). The tools used can be installed either on previous phases or on this phase. Also, the attacker seems to stage captured data to a single point to allow data transfer from a single client rather than multiple clients, which could raise defenders' suspicions.

On their study, Ussath et al. (2016) argue that there are only three main phases for an APT campaign: Initial Compromise, Lateral Movement and Command & Control. While CKC phases are named differently, our thesis found similar results. On three different phases, Delivery, Installation and Command & Control, the difference of the most utilized and the second most utilized technique was significant, over 20%. This would imply that certain techniques are more popular on certain CKC phases and thus analyzing these phases, the characteristics of an APT campaign can be identified.

Figure 31 summarizes the results of this study. According to the results, first Chinese APT actors rely on gathering the victim organizations information, then develop capabilities to attack and deliver the weapon to the target system by utilizing phishing, usually spear phishing. Once the weapon is delivered, it will utilize command and scripting interpreter to exploit the target system. After the exploitation, the attack will continue with installation of a web shell, backdoor or something similar, and afterwards the weapon will contact the C2 network utilizing application layer protocols. Finally, the cyber-attack will be concluded using different remote access tools to exfiltrate data or to expand the attack.

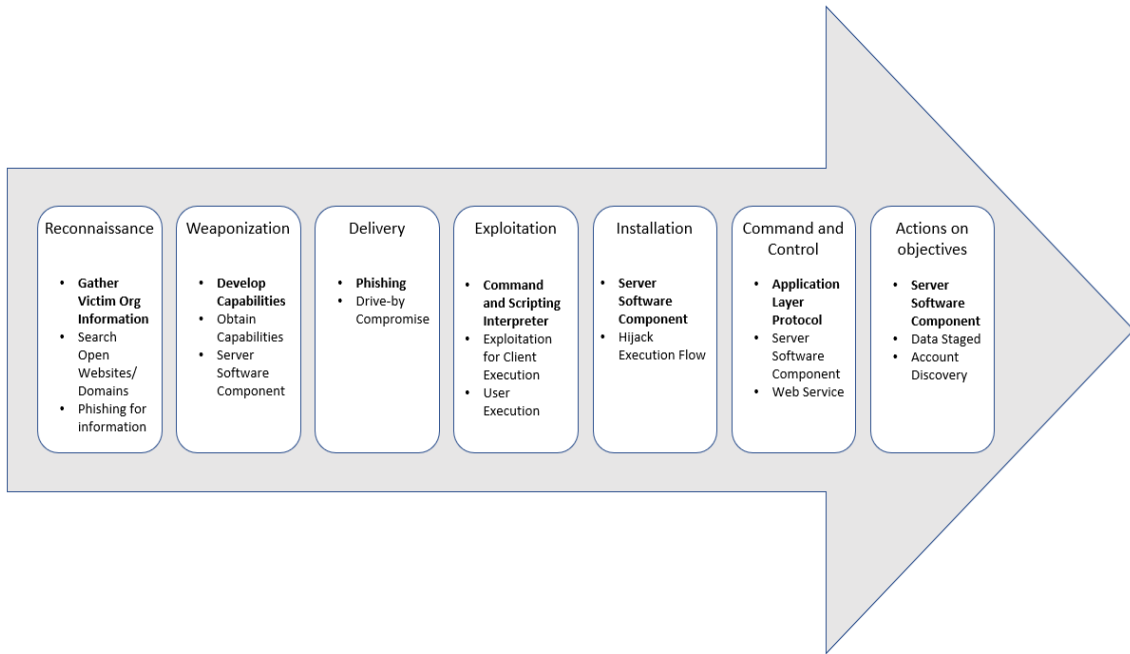


FIGURE 31 Chinese APT Actors' TTPs

8 CONCLUSION

The objective of this thesis was to find out how Chinese APT actors execute APT cyber-attacks. More specifically, the research question of the thesis concentrated on finding out what are Chinese APT actors' tactics, techniques, and procedures? During the research a decision was made to utilize ThaiCert's web portal as the source of research material. Finally, 41 China related APT reports were chosen based on certain criteria as the research material of this thesis. These reports handled altogether 11 different China related APT groups' cyber-attacks. Qualitative research was chosen as the primary research method and more specifically theory-driven content analysis was utilised. Qualitative analysis was enriched with quantitative elements especially in the coding part of the research. As often in qualitative research, when the understanding of the topic under investigation increased, the research itself was reformulated based on the increased understanding of the research issue on hand.

The decision to investigate tactics, techniques, and procedures mainly determined the theoretical frameworks used in this thesis. Quickly it was noticed that the term TTP has different definitions and therefore the writers also interpreted TTPs in a unique way. First, the research material was coded with Cyber Kill Chain Framework's seven phases: Reconnaissance, Weaponization, Delivery, Exploitation, Installation, Command and Control, and Actions on Objectives. The coded texts and phrases were treated as *procedures* of an APT cyber-attack which describe how tactics and techniques were used. The seven CKC phases were treated as the *tactics* i.e., objectives the attacker wants to achieve in certain phase of the cyber-attack. Second, the procedures were further analysed with MITRE's ATT&CK framework's cyber-attack techniques forming the *techniques* part of our TTP inspection. To our understanding, the above-mentioned frameworks are not interpreted as in this thesis and this might cause some confusion and affect to validity and reliability of this thesis depending on the interpretation and its validity by e.g., other researchers.

The most used TTPs of China related APT actors were recognized from the research material. The results were presented following CKC framework's seven cyber-attack phases. The most used Reconnaissance techniques were Gather Victim Org Information (T1591), Search Open Websites/Domains (T1593), and Phishing for Information (T1598). The Weaponization was performed utilizing Develop Capabilities (T1587), Obtain Capabilities (T1588), and Server Software Component (T1505) techniques. Phishing (T1566) and Drive by Compromise (T1189) were most popular techniques during the Delivery phase. The Exploitation was performed with techniques Command and Scripting Interpreter (T1059), Exploitation for Client Execution (T1203), and User Execution (T1204). During the Installation phase the most used techniques were Server Software Component (T1505), and Hijack Execution Flow (T1574). Application Layer Protocol (T1071), Server Software Component (T1505), and Web Service (T1102) were the most utilized techniques in the Command-and-Control phase. Finally, the

Actions on Objectives phase was executed with Server Software Component (T1505), Data Staged (T1704), and Account Discovery (T1087) techniques. The results are in line with the previous research by highlighting spear-phishing as the most used delivery method for China related APT actors, the use of HTTP(s) protocol as the C2 channel, and the used attacking tools shared and developed within the China related APT groups.

For further research it might be interesting to examine more thoroughly Lehto's (2022) idea of strategic decision-making phase, which is conducted before the first CKC-phase. There can be multiple different motivations for Chinese APT actors to conduct cyber-attacks and these motivations were already touched in chapter 4.2 of this thesis. Also, it would be interesting to test our theoretical model with different research material e.g., APT groups attributed to other countries than China. During the coding phase of our research the writers noticed that the use of CKC framework was a bit ambivalent. Therefore, utilizing only MITRE's ATT&CK framework would be sufficient and might also be clearer for the reader.

There are certain restrictions for the TTP research for independent individuals e.g., often the research material is in the public Internet composed by private cyber security companies as their marketing material, and the attribution part is left out or presented with insufficient information. However, to our knowledge China related APT actors' TTPs had not been examined the same way as in this thesis. The writers see that the results increase the knowledge and awareness of the modus operandi of Chinese cyber-attackers' actions in cyberspace.

REFERENCES

- Ahmad, A., Webb, J., Desouza, K. C., & Boorman, J. (2019). Strategically-motivated advanced persistent threat: Definition, process, tactics and a disinformation model of counterattack. *Computers & Security*, 86, 402-418. <https://doi.org/10.1016/j.cose.2019.07.001>
- AI Amin, M. A. R., Shetty, S., Njilla, L., Tosh, D. K., & Kamhoua, C. (2021). Hidden Markov Model and Cyber Deception for the Prevention of Adversarial Lateral Movement. *IEEE Access*, 9, 49662-49682. <https://doi.org/10.1109/ACCESS.2021.3069105>
- Alastalo, M. & Vuori, J. (2023). *Dokumentit*. Teoksessa J. Vuori (toim.), Laadullisen tutkimuksen verkkokäsikirja. Tampere: Yhteiskuntatieteellinen tietoaarkisto. Retrieved from <https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus/kvali/laadullisen-tutkimuksen-aineistot/dokumentit/>.
- Allen, J., Yang, Z., Landen, M., Bhat, R., Grover, H., Chang, A., Yang, J., Perdisci, R., & Lee, W. (2020). Mnemosyne: An Effective and Efficient Postmortem Watering Hole Attack Investigation System. *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security*, 787-802. <https://doi.org/10.1145/3372297.3423355>
- Alshamrani, A., Myneni, S., Chowdhary, A., & Huang, D. (2019). A survey on advanced persistent threats: Techniques, solutions, challenges, and research opportunities. *IEEE Communications Surveys & Tutorials*, 21(2), 1851-1877. <https://doi.org/10.1109/COMST.2019.2891891>
- Assante, M. J., & Lee, R. M. (2015). *The industrial control system cyber kill chain*. SANS Institute InfoSec Reading Room. Retrieved from https://scadahacker.com/library/Documents/White_Papers/SANS%20-%20ICS%20Cyber%20Kill%20Chain.pdf
- Austin, G. (2015). *Cyber Security: All China's Fault*. Retrieved from <https://www.theglobalist.com/cyber-diplomacy-us-china-problem/>.
- Australian Government. (2020). *Australia's Cyber Security Strategy 2020*. Retrieved from <https://www.homeaffairs.gov.au/cyber-security-sub-site/files/cyber-security-strategy-2020.pdf>
- Bahrami, P. N., Dehghantanha, A., Dargahi, T., Parizi, R. M., Choo, K. K. R., & Javadi, H. H. (2019). Cyber Kill Chain-Based Taxonomy of Advanced Persistent Threat Actors: Analogy of Tactics, Techniques, and Procedures.

Journal of Information Processing Systems, 15(4), 865-889.
<http://dx.doi.org/10.3745/JIPS.03.0126>

- Baker, K. (2023). What is cyber threat intelligence. Retrieved from <https://www.crowdstrike.com/cybersecurity-101/threat-intelligence/>.
- Banks, W. (2017). State responsibility and attribution of cyber intrusions after Tallinn 2.0. *Texas Law Review*, 95(7), 1487-1513.
- Berghel, H. (2017). On the Problem of (Cyber) Attribution. *Computer*, 50(3), 84-89. <https://doi.ieeecomputersociety.org/10.1109/MC.2017.74>
- Bianco, D. (2014). *The pyramid of pain*. Retrieved from <https://detect-respond.blogspot.com/2013/03/the-pyramid-of-pain.html?view=classic>
- Blanco, C., Lasheras, J., Valencia-García, R., Fernández-Medina, E., Toval, A., & Piattini, M. (2008). A systematic review and comparison of security ontologies. *2008 Third International Conference on Availability, Reliability and Security*, 813-820. IEEE. <https://doi.org/10.1109/ARES.2008.33>
- Bhatnagar, D., Som, S., & Khatri, S. K. (2019). Advance Persistent Threat and Cyber Spying - The Big Picture, Its Tools, Attack Vectors and Countermeasures. *2019 Amity International Conference on Artificial Intelligence (AICAI)*, 828-839. IEEE. <https://doi.org/10.1109/AICAI.2019.8701329>
- Binde, B. E., McRee, R., & O'Connor, T. J. (2011). Accessing outbound traffic to uncover advanced persistent threat. *SANS Technology Institute*. <http://dx.doi.org/10.13140/RG.2.2.16401.07520>
- Bozhkov, N. (2020). *China's Cyber Diplomacy: A Primer*. EU Cyber Direct. Retrieved from <https://eucd.s3.eu-central-1.amazonaws.com/eucd/assets/-GX150Cl/bozhkov-digital-dialogue-final.pdf>
- Brook, C. (2011). *Titan Rain*. Threat post. Retrieved December 7, 2021, from <https://threatpost.com/titan-rain/91835/>.
- Bunda, J. (2020). *APT28 : tapaustutkimus Venäjään yhdistettyjen kyberoperaatioiden kehittymisestä vuosina 2007-2016* (Pro gradu-thesis). University of Jyväskylä. Retrieved from <https://jyx.jyu.fi/handle/123456789/67845>.
- Caltagirone, S., Pendergast, A., & Betz, C. (2013). *The diamond model of intrusion analysis*. Retrieved from <https://apps.dtic.mil/sti/pdfs/ADA586960.pdf>
- Cambridge Dictionary. (2023a). *Intelligence*. Cambridge Dictionary. Retrieved from <https://dictionary.cambridge.org/dictionary/english/intelligence>.

- Cambridge Dictionary. (2023b). *Taxonomy*. Cambridge Dictionary. Retrieved from <https://dictionary.cambridge.org/dictionary/english/taxonomy>.
- Central Intelligence Agency (2021). *China*. The world factbook. Retrieved November 23, 2021, from <https://www.cia.gov/the-world-factbook/countries/china/>.
- Chapman, I., Leblanc, S., Partington, A. (2011). Taxonomy of cyber attacks and simulation of their effects. *MMS '11: Proceedings of the 2011 Military Modeling & Simulation Symposium*, 73-80. Boston: Society for Computer Simulation International. <https://dl.acm.org/doi/10.5555/2048558.2048569>
- Chen, J., Su, C., Yeh, K. H., & Yung, M. (2018). Special issue on advanced persistent threat. *Future generation computer system*, 79(1), 243-246. <https://doi.org/10.1016/j.future.2017.11.005>
- Chen, P., Desmet, L., & Huygens, C. (2014). A study on advanced persistent threats. In De Decker, B., Zúquete, A. (eds), *Lecture Notes in Computer Science : Vol. 8735. Communications and Multimedia Security* (pp. 63-72). Springer. https://doi.org/10.1007/978-3-662-44885-4_5
- Cheng, D. (2016). *Cyber Dragon: Inside China's Information Warfare and Cyber Operations : Inside China's Information Warfare and Cyber Operations*. Praeger.
- Cho, S., Han, I., Jeong, H., Kim, J., Koo, S., Oh, H., & Park, M. (2018). Cyber Kill Chain Based Threat Taxonomy and its Application on Cyber Common Operational Picture. In *2018 International Conference On Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA)* (pp. 1-8). IEEE. <https://doi.org/10.1109/CyberSA.2018.8551383>
- Cimpanu, C. (2021). *Finland pins Parliament hack on Chinese hacking group APT31*. Retrieved from <https://therecord.media/finland-pins-parliament-hack-on-chinese-hacking-group-apt31/>
- Cole, E. (2012). *Advanced persistent threat: Understanding the danger and how to protect your organization*. Syngress Publishing.
- Committee on National Security Systems. (2017). *Committee on National Security Systems (CNSS) Glossary*. Retrieved from <https://rmf.org/wp-content/uploads/2017/10/CNSSI-4009.pdf>
- Corbin, K. (2013). 'Aurora' cyber attackers were really running counter-intelligence. Retrieved December 9, 2021, from

<https://www.cio.com/article/2386547/-aurora--cyber-attackers-were-really-running-counter-intelligence.html>.

Council on Foreign Relations. (CFR). (2010). *Operation Aurora*. Retrieved from <https://www.cfr.org/cyber-operations/operation-aurora>.

Council on foreign relations. (CFR). (2021). *Titan Rain*. Retrieved from <https://www.cfr.org/cyber-operations/titan-rain>.

CrowdStrike. (2019). *Intelligence report : Huge fan of your work : How TURBINE PANDA and China's top spies enabled Beijing to cut corners on the C919 passenger jet*. Retrieved from <https://passle-net.s3.amazonaws.com/Passle/5c752afb989b6e0f5cda12f4/MediaLibrary/Document/2019-10-18-10-42-26-646-huge-fan-of-your-work-intelligence-report.pdf>

Cyvera. (2016). *Remembering operation titan rain*. Retrieved December 7, 2021, from <https://cyware.com/news/remembering-operation-titan-rain-c54ad3e4>

Dargahi, T., Dehghantanha, A., Bahrami, P. N., Conti, M., Bianchi, G., & Benedetto, L. (2019). A Cyber-Kill-Chain Based Taxonomy of Crypto-Ransomware Features. *Journal of Computer Virology and Hacking Techniques*, 15(4), 277-305. <https://doi.org/10.1007/s11416-019-00338-7>

Derbyshire, R., Green, B., Prince, D., Mauthe, A., & Hutchison, D. (2018). An Analysis of Cyber Security Attack Taxonomies. *2018 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)* (pp. 153-161). IEEE. <https://doi.org/10.1109/EuroSPW.2018.00028>

Diotte, P. (2020). *The Big Four and Cyber Espionage: How China, Russia, Iran and North Korea Spy Online*. Retrieved from <http://www.journal.forces.gc.ca/Vol20/No4/page32-eng.asp>

Doshi, R., de la Bruyère, E., Picarsic, N., & Ferguson, J. (2021). *China as a 'cyber great power': Beijing's two voices in telecommunications*. Retrieved from <https://www.brookings.edu/research/china-as-a-cyber-great-power-beijings-two-voices-in-telecommunications/>

Edgar, T., W. & Manz, D., O. (2017). *Research Methods for Cyber Security*. Syngress.

ENISA. (2018). *ENISA Threat landscape report 2017*. Retrieved from <https://www.enisa.europa.eu/publications/enisa-threat-landscape-report-2017/@@download/fullReport>

Eskola, J. & Suoranta, J. (1998). *Johdatus laadulliseen tutkimukseen*. Tampere: Vastapaino.

- Finnish Security and Intelligence Service (SUPO). (2021). *National Security Overview 2021*. Retrieved from <https://supo.fi/documents/38197657/39761266/National+Security+Overview+2021.pdf/a772aa98-30bc-1bcc-62c0-9d77dda2733b/National+Security+Overview+2021.pdf?t=1649405324653>
- Finnish Security Committee (FSC). (2018). *Vocabulary of Comprehensive Security*. Retrieved from https://turvallisuuskomitea.fi/wp-content/uploads/2018/02/Kokonaisturvallisuuden_sanasto.pdf
- Fruhlinger, J. (2020). *The OPM hack explained: Bad security practices meet China's Captain America*. Retrieved from <https://www.csoonline.com/article/3318238/the-opm-hack-explained-bad-security-practices-meet-chinas-captain-america.html>
- Ghafir, I., & Prenosil, V. (2016). Proposed Approach for Targeted Attacks Detection. In Sulaiman, H., Othman, M., Othman, M., Rahim, Y., Pee, N. (eds), *Lecture Notes in Electrical Engineering: Vol. 362. Advanced Computer and Communication Engineering Technology* (pp. 73-80). Springer. https://doi.org/10.1007/978-3-319-24584-3_7
- Giles, M. (2018). *Chinese Hackers allegedly stole data of more than 100,00 US Navy personnel*. Retrieved from <https://www.technologyreview.com/2018/12/20/239760/chinese-hackers-allegedly-stole-data-of-more-than-100000-us-navy-personnel/>
- Gilli, A., & Gilli, M. (2018). Why China has not caught up yet: military-technological superiority and the limits of imitation, reverse engineering, and cyber espionage. *International security*, 43(3), 141-189. https://doi.org/10.1162/isec_a_00337
- Google. (2010). *A new approach to China*. Retrieved December 8, 2021, from <https://googleblog.blogspot.com/2010/01/new-approach-to-china.html>
- Greenberg, A. (2018). *The untold story of NotPetya, the most devastating cyberattack in history*. Retrieved from <https://www.wired.com/story/notpetya-cyberattack-ukraine-russia-code-crashed-the-world/>
- Grooby, S., Dargahi, T., & Dehghantanha, A. (2019). Protecting IoT and ICS platforms against advanced persistent threat actors: analysis of APT1, silent chollima and molerats. In Dehghantanha, A., Choo, KK. (eds), *Handbook of big data and IoT security* (pp. 225-255). Springer, Cham. <https://doi.org/10.1007/978-3-030-10543-3>

- Grotto, A. (2020). Deconstructing Cyber Attribution: A Proposed Framework and Lexicon. *IEEE Security & Privacy*, 18(1), 12-20. <https://doi.org/10.1109/MSEC.2019.2938134>
- Healey, J. (2011). The spectrum of national responsibility for cyberattacks. *The Brown Journal of World Affairs*, 18, 57-70. <https://doi.org/10.26300/6zn1-c269>
- Heuer, R. J. (1999). *Psychology of intelligence analysis*. Central Intelligence Agency.
- Hirsjärvi, S., Remes, P. & Sajavaara, P. (2004). *Tutki ja kirjoita*. (10. uud. painos). Helsinki: Tammi.
- Hoglund, G. (2009). *Advanced persistent threat, what APT means to your enterprise*. Retrieved from <https://pdfs.semanticscholar.org/d0a0/47c6b19fc3645973f8f300b507886b54196a.pdf>
- HM Government. (2016). *National Cyber Security Strategy 2016-2021*. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/567242/national_cyber_security_strategy_2016.pdf
- Hutchins, E. M., Cloppert, M. J., & Amin, R. M. (2011). Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains. *Proceedings of the 6th International Conference on Information Warfare and Security* (pp. 113-125). Washington, DC: George Washington University.
- Inkster, N. (2015). China's Cyber power. *Adelphi Series*, 55(465), 7-150.
- Jaafar, F., Avellanade, F., & Alikacem, E. (2020). Demystifying the Cyber Attribution: An Exploratory Study. *2020 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech)*. (35-40). Calgary, AB: IEEE. <http://dx.doi.org/10.1109/DASC-PiCom-CBDCom-CyberSciTech49142.2020.00022>
- Jeun, I., Lee, Y., & Won, D. (2012). A practical study on advanced persistent threats. In Th. Kim, et al., *Communications in Computer and Information Science: Vol. 339. Computer applications for security, control and system engineering* (pp. 144-152). Springer. https://doi.org/10.1007/978-3-642-35264-5_21
- Johnson, C., Badger, M., Waltermire, D., Snyder, J., & Skorupka, C. (2018) *Guide to Cyber Threat Information Sharing*. Retrieved March 19, 2023 from

<https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-150.pdf>

Joint Publication 2-0 (JP 2-0). (2013). *Joint Intelligence*. Retrieved from https://irp.fas.org/doddir/dod/jp2_0.pdf

Kaska, K., Beckvard, H., & Minarik, T. (2019). Huawei, 5G and China as a security threat. *NATO Cooperative Cyber Defence Center for Excellence (CCDCOE)*, 28.

Kaspersky. (2023a). *What Is an Advanced Persistent Threat (APT)?*. Retrieved from <https://www.kaspersky.com/resource-center/definitions/advanced-persistent-threats>

Kaspersky. (2023b). *What is a DDoS Attack? -DDos Meaning*. Retrieved from <https://usa.kaspersky.com/resource-center/threats/ddos-attacks>

Kim, H., Kwon, H., & Kim, K. K. (2019). Modified Cyber Kill Chain Model for Multimedia Service Environments. *Multimedia Tools and Applications*, 78(3), 3153-3170. <http://dx.doi.org/10.1007/s11042-018-5897-5>.

Kime, B. P. (2016). *Threat Intelligence: Planning and Direction*. Retrieved from <https://www.sans.org/white-papers/36857/>.

Kirk, J. & Miller, M., L. (1986) *Reliability and Validity in Qualitative Research*. SAGE.

Kiwia, D., Dehghantanha, A., Choo, K. K. R., & Slaughter, J. (2018). A Cyber Kill Chain Based Taxonomy of Banking Trojans for Evolutionary Computational Intelligence. *Journal of computational science*, 27, 394-409. <https://doi.org/10.1016/j.jocs.2017.10.020>

Kolton, M. (2017). Interpreting China's Pursuit of Cyber Sovereignty and its Views on Cyber Deterrence. *The Cyber Defense Review*, 2(1), 119-154. <https://www.jstor.org/stable/26267405>

Kuusisto, T., & Kuusisto, R. (2015). Cyber World as a Social System. In Lehto, M., & Neittaanmäki, P. (Eds.). *Cyber security: Analytics, Technology and Automation* (31-43). Springer International Publishing AG. https://doi.org/10.1007/978-3-319-18302-2_2

Laari, T., Flyktman, J., Härmä, K., Timonen, J. & Tuovinen, J. (2019). *#Kyberpuolustus [#Cyberdefense]*. Helsinki. Maanpuolustuskorkeakoulu.

Laufer, E. (2021). *A look at the MITRE ATT&CK Content Pack from Cortex XSOAR Marketplace*. Retrieved March 19, 2023, from <https://www.paloaltonetworks.com/blog/security-operations/MITRE-attck-for-cortex-xsoar/>

- Launius, S. (2020). Evaluation of comprehensive taxonomies for information technology threats. *Journal of the cyber security & information systems information analysis center*, 7(4), 4-17. Retrieved from <https://csiac.org/articles/evaluation-of-comprehensive-taxonomies-for-information-technology-threats/>
- Lee, R. M., & Brown, R. (2021). *FOR578: Cyber Threat Intelligence*. Retrieved November 17, 2021, from <https://www.sans.org/cyber-security-courses/cyber-threat-intelligence/>.
- Lehto, M. (2021). *Digitaalisen kybermaailman ilmiöitä ja määrittelyjä* (Version 15.0). Jyväskylän Yliopisto. Retrieved from https://www.jyu.fi/it/fi/hae-opiskelemaan/hakukohteet/kybma/kybermaailma_v15-0.pdf.
- Lehto, M. (2022). APT cyber-attack modelling – building a general model. In R. P. Griffin, U. Tatarand, & B. Yankson (Eds.), *ICCWS 2022 : Proceedings of the 17th International Conference on Cyber Warfare and Security 17* (pp. 121-129). Academic Conferences International Ltd. <https://doi.org/10.34190/ic-cws.17.1.36>
- Lemay, A., Calvet, J., Menet, F., & Fernandez, J. M. (2018). Survey of publicly available reports on advanced persistent threat actors. *Computers & Security*, 72, 26-59. <https://doi.org/10.1016/j.cose.2017.08.005>
- Lin, H. (2016). Attribution of malicious cyber incidents: From soup to nuts. *Journal of International Affairs*, 70(1), 75-137. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2835719
- Lindsay, J. R. (2014). The impact of China on cyber security: Fiction and friction. *International Security*, 39(3), 7-47. https://doi.org/10.1162/ISEC_a_00189
- Lockheed Martin (2023). *The Cyber Kill Chain®*. Retrieved January 1, 2023, from <https://www.lockheedmartin.com/en-us/capabilities/cyber/cyber-kill-chain.html#>
- Mandiant. (2010). *M-Trends. The Advanced Persistent Threat*. Retrieved from <https://www.slideshare.net/FireEyeInc/mtrends-2010-the-advanced-persistent-threat>
- Mandiant. (2013). *APT1, Exposing One of China's Cyber Espionage Unit*. Retrieved April 6, 2023, from <https://www.mandiant.com/sites/default/files/2021-09/mandiant-apt1-report.pdf>

- Mandiant. (2023). *Advanced Persistent Threats (APTs)*. Retrieved from <https://www.mandiant.com/resources/insights/apt-groups>
- Markstedter, M. (2020). (n.t.). Retrieved November 17, 2021, from <https://azerialabs.com/advanced-persistent-threat>
- Marsiano, R. (2022). *Joint forces – MS Sentinel and the MITRE framework*. Retrieved March 19, 2023, from <https://techcommunity.microsoft.com/t5/microsoft-sentinel-blog/joint-forces-ms-sentinel-and-the-MITRE-framework/bap/3191589>
- Martin, A. (2022). *Mondelez and Zurich reach settlement in NotPetya cyberattack insurance suit*. Retrieved September 24, 2023, from <https://therecord.media/mondelez-and-zurich-reach-settlement-in-notpetya-cyberattack-insurance-suit>
- Mavroeidis, V., & Bromander, S. (2017). Cyber threat intelligence model: an evaluation of taxonomies, sharing standards, and ontologies within cyber threat intelligence. *2017 European Intelligence and Security Informatics Conference (EISIC)* (pp. 91-98). IEEE. <https://doi.org/10.1109/EISIC.2017.20>
- Mazerik, R. (2014). *Remote access tool*. Retrieved November 13, 2021, from <https://resources.infosecinstitute.com/topic/remote-access-tool/>
- McCarthy, K. (2019) *Cyber-insurance shock: Zurich refuses to foot NotPetya ransomware clean-up bill – and claims it's 'an act of war'*. Retrieved from https://www.theregister.com/2019/01/11/notpetya_insurance_claim/
- McMillan, R. (2013). *Definition: Threat Intelligence*. Retrieved from <https://www.gartner.com/en/documents/2487216>
- McWhorter, D. (2021). *APT1: Exposing One of China's Cyber Espionage Units*. Retrieved November 24, 2021, from <https://www.mandiant.com/resources/apt1-exposing-one-of-chinas-cyber-espionage-units>
- Merriam-Webster. (2021). *Cyber*. Retrieved from <https://www.merriam-webster.com/dictionary/cyber>
- Merriam-Webster. (2023). *Taxonomy*. Retrieved from <https://www.merriam-webster.com/dictionary/taxonomy>
- MITRE ATT&CK. (2023a). *ATT&CK Matrix for Enterprise*. Retrieved March 19, 2023, from <https://attack.MITRE.org/>

- MITRE ATT&CK. (2023b). *Groups*. Retrieved August 8, 2023, from <https://attack.mitre.org/groups/>
- MITRE Engenuity (2023). *ATT&CK Evaluations: Using evaluations*. Retrieved March 19, 2023, from [https://MITRE-engenuity.org/cyber security/ attack-evaluations/using-attack-evaluations/](https://MITRE-engenuity.org/cyber-security/attack-evaluations/using-attack-evaluations/)
- Mohsin, M., & Anwar, Z. (2016). Where to kill the cyber kill-chain: An ontology-driven framework for iot security analytics. *2016 International Conference on Frontiers of Information Technology (FIT)* (pp. 23-28). IEEE. <https://doi.org/10.1109/FIT.2016.013>
- Mueller, M., Grindal, K., Kuerbis, B., & Badiei, F. (2019). Cyber Attribution: Can a New Institution Achieve Transnational Credibility? *The Cyber Defense Review*, 4(1), 107-122. Retrieved from <https://www.jstor.org/stable/26623070>
- Myers, M. D. (1997). Qualitative Research in Information Systems. *MIS Quarterly*, 21, 241-242. <http://dx.doi.org/10.2307/249422>
- Mäkelä, K. (1997). *Kvalitatiivisen aineiston analyysi ja tulkinta*. Helsinki: Gaudeamus.
- Nakashima, E., & Harris, S. (2018). *How Russia hacked the DNC and passed its emails to WikiLeaks*. Retrieved April 16, 2023, from https://www.washingtonpost.com/world/national-security/how-the-russians-hacked-the-dnc-and-passed-its-emails-to-wikileaks/2018/07/13/af19a828-86c3-11e8-8553-a3ce89036c78_story.html
- New Zealand Government. (2019). *New Zealand's Cyber Security strategy 2019*. Retrieved from <https://www.dpmc.govt.nz/sites/default/files/2019-07/Cyber%20Security%20Strategy.pdf>
- NIST. (2011). *Special Publication 800-39, Managing information security risk*. Retrieved from <https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecial-publication800-39.pdf>
- NIST. (2012). *Special Publication 800-30 Revision 1, Guide for conducting risk assessments*. Retrieved from <https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-30r1.pdf>
- NIST. (2016). *Special Publication 800-150, Guide to Cyber Threat Information Sharing*. Retrieved from <https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-150.pdf>

- NIST. (2021). *tactics, techniques and procedures (TTP)*. Retrieved from https://csrc.nist.gov/glossary/term/tactics_techniques_and_procedures
- NIST. (2023). *Cyber Threat Intelligence*. Retrieved from https://csrc.nist.gov/glossary/term/cyber_threat_intelligence
- NIST Glossary. (2023). *Glossary*. Retrieved from <https://csrc.nist.gov/glossary>
- Norton-Taylor, R. (2007). *Titan Rain – how Chinese hackers targeted whitehall*. Retrieved December 7, 2021, from <https://www.theguardian.com/technology/2007/sep/04/news.internet>
- OASIS Open. (2023). *Introduction to STIX*. Retrieved September 24, 2023, from <https://oasis-open.github.io/cti-documentation/stix/intro>
- Open Web Application Security Project (OWASP). (2021). *Server Side Request Forgery*. Retrieved November 13, 2021, from https://owasp.org/www-community/attacks/Server_Side_Request_Forgery
- Orinx, K., & de Swielande, T. S. (2019). A Chinese Fox against an American Hedgehog in Cyberspace. *Military Review*, 100(5). Retrieved from https://dial.uclouvain.be/pr/boreal/object/boreal%3A219034/datastream/PDF_01/view
- Pande, D. N., & Voditel, P. S. (2017). Spear phishing: Diagnosing attack paradigm. In *2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)* (pp. 2720-2724). IEEE. <http://dx.doi.org/10.1109/WiSPNET.2017.8300257>
- Pols, P. (2017). *The Unified Kill Chain*. Retrieved from <https://www.unified-killchain.com/assets/The-Unified-Kill-Chain.pdf>
- Potter, W. J., & Levine-Donnerstein, D. (1999). Rethinking validity and reliability in content analysis. *Journal of Applied Communication Research*, 27(3), 258–284. <https://doi.org/10.1080/00909889909365539>
- Quintero-Bonilla, S., & Martín del Rey, A. (2020). A new proposal on the advanced persistent threat: a survey. *Applied Sciences*, 10(11), 3874. <http://dx.doi.org/10.3390/app10113874>
- Raud, M. (2016). *China and Cyber : Attitude, Strategies and Organization*. Retrieved from https://ccdcoe.org/uploads/2018/10/CS_organisation_CHINA_092016_FINAL.pdf

- Saaranen-Kauppinen, A. & Puusniekka, A. (2006) *KvaliMOTV – Menetelmäopetuksen tietovaranto*. Tampere: Yhteiskuntatieteellinen tietoaarkisto. Retrieved from <https://www.fsd.tuni.fi/menetelmaopetus/>
- Sabbagh, D. (2010). *WikiLeaks cables blame Chinese government for Google hacking*. Retrieved December 14, 2021, from <https://www.theguardian.com/technology/2010/dec/04/wikileaks-cables-google-china-hacking>
- Schmidt, R., Rattray, G. J., & Fogle, C. J. (2008). *Methods and apparatus for developing cyber defense processes and a cadre of expertise*. Retrieved from <https://patents.google.com/patent/US20080167920A1/en>
- Segal, A. (2016). *The U.S.-China Cyber Espionage Deal One Year Later*. Retrieved from <https://www.cfr.org/blog/us-china-cyber-espionage-deal-one-year-later>
- Sexton, J., Storlie, C., & Neil, J. (2015). Attack chain detection. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 8(5-6), 353-363. <https://doi.org/10.1002/sam.11296>
- Shaikh, K. A., Bhat, A. K., & Moharir, M. (2017). A survey on SSL packet structure. *2017 2nd International Conference on Computational Systems and Information Technology for Sustainable Solution (CSITSS)* (pp. 1-5). IEEE. <https://doi.org/10.1109/CSITSS.2017.8447634>
- Sheehan, M. (2018). *How Google took on China – and lost*. Retrieved April 17, 2023, from <https://www.technologyreview.com/2018/12/19/138307/how-google-took-on-china-and-lost/>
- Sood, A. K., & Enbody, R. J. (2012). Targeted cyberattacks: a superset of advanced persistent threats. *IEEE security & privacy*, 11(1), 54-61. <http://dx.doi.org/10.1109/MSP.2012.90>
- Spetalnick, M., & Martina, M. (2015). *Obama announces ‘understanding’ with China’s Xi on cyber theft but remains wary*. Retrieved from <https://www.reuters.com/article/us-usa-china-idUSKCN0RO2HQ20150926>
- Strom, B. E., Applebaum, A., Miller, D. P., Nickels, K. C., Pennington, A. G., & Thomas, C. B. (2020). *MITRE ATT&CK®: Design and Philosophy*. Retrieved from <https://www.mitre.org/sites/default/files/2021-11/prs-19-01075-28-mitre-attack-design-and-philosophy.pdf>
- Stubbs, J., Menn, J., & Bing, C. (2019). *Inside the West’s failed fight against China’s “Cloud Hopper” hackers*. Retrieved from <https://www.reuters.com/investigates/special-report/china-cyber-cloudhopper>

- Swisscom. (2019). *Targeted Attacks Cyber Security Report 2019*. Retrieved from https://documents.swisscom.com/product/filestore/lib/7657c513-a231-4725-9d04-eeb343c164e1/Swisscom_Security_Report_2019_EN.pdf
- Tankard, C. (2011). Advanced persistent threats and how to monitor and deter them. *Network security*, 2011(8), 16-19. [https://doi.org/10.1016/S1353-4858\(11\)70086-1](https://doi.org/10.1016/S1353-4858(11)70086-1)
- ThaiCERT. (2023). *Threat Group Cards: A Threat Actor Encyclopedia*. Retrieved from <https://apt.etcha.or.th/cgi-bin/aptgroups.cgi>
- The International Institute for Strategic Studies (IISS). (2021). *Cyber Capabilities and National Power: A Net Assesment*. Retrieved from https://www.iiss.org/globalassets/media-library---content--migration/files/research-papers/cyber-power-report/cyber-capabilities-and-national-power---a-net-assessment____.pdf
- Thornburg, N. (2005). *The invasion of Chinese cyberspies*. Retrieved February 7, 2021, from <http://content.time.com/time/subscriber/article/0,33009,1098961,00.html>
- Tuomi, J. & Sarajärvi, A. (2018) *Laadullinen tutkimus ja sisällönanalyysi*. Helsinki: Tammi.
- Turvallisuuskomitea. (2019). *Suomen Kyberturvallisuusstrategia 2019*. Retrieved from https://turvallisuuskomitea.fi/wp-content/uploads/2019/10/Kyberturvallisuusstrategia_A4_SUOMI_WEB_300919.pdf
- Töttö, P. (2004). *Syvällistä ja pinnallista. teoria, empiria ja kausaalisuus sosiaalitutkimuksessa*. Tampere: Vastapaino.
- Ussath, M., Jaeger, D., Cheng, F., & Meinel, C. (2016). Advanced persistent threats: Behind the scenes. In *2016 Annual Conference on Information Science and Systems (CISS)* (pp. 181-186). IEEE. <https://doi.org/10.1109/CISS.2016.7460498>
- van der Watt, R., & Slay, J. (2021). Modification of the Lockheed Martin Cyber Kill Chain (LMCKC) for Cyber Security Breaches Concerning Low Earth Orbit (LEO) Satellites. In *ICCVS 2021 16th International Conference on Cyber Warfare and Security* (p. 473). Academic Conferences Limited.
- Varto, J. (1992). *Laadullisen tutkimuksen metodologia*. Helsinki: Kirjayhtymä.
- Vatanen, V. (2020). *Venäjän lähialueillaan toteuttamien kyberoperaatioiden analysointi* (Pro gradu -thesis). University of Jyväskylä. Retrieved from

<https://jyx.jyu.fi/bitstream/handle/123456789/69032/URN%3ANBN%3Afi%3Aju-202005183286.pdf?sequence=1>

- Virvilis, N., & Gritzalis, D. (2013). The big four-what we did wrong in advanced persistent threat detection?. In *2013 international conference on availability, reliability and security* (pp. 248-254). IEEE. <https://doi.org/10.1109/ARES.2013.32>
- Vukalović, J., & Delija, D. (2015). Advanced persistent threats-detection and defense. In *2015 38th international convention on information and communication technology, electronics and microelectronics (MIPRO)* (pp. 1324-1330). IEEE. <https://doi.org/10.1109/MIPRO.2015.7160480>
- Vuori, J. (2023). *Laadullisen tutkimuksen verkkokäsikirja*. Tampere: Yhteiskuntatieteellinen tietoaarkisto. Retrieved from <https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus/>
- Walters, P. (2012). *The Risks of Using Portable Devices*. Retrieved November 13, 2021, from <https://www.cisa.gov/sites/default/files/publications/RisksOfPortableDevices.pdf>
- Web Application Security Consortium (WASC). (2010). *WASC Threat Classification*. Retrieved from http://projects.webappsec.org/f/WASC-TC-v2_0.pdf
- White House. (2018). *National Cyber Strategy of the United States of America*. Retrieved from <https://trumpwhitehouse.archives.gov/wp-content/uploads/2018/09/National-Cyber-Strategy.pdf>
- Xiang, Z., Guo, D., & Li, Q. (2020). Detecting mobile advanced persistent threats based on large-scale DNS logs. *Computers & Security*, 96, 101933. <https://doi.org/10.1016/j.cose.2020.101933>
- Yadav, T., & Rao, A. M. (2015). Technical aspects of cyber kill chain. In Abawajy, J., Mukherjea, S., Thampi, S., Ruiz-Martínez, A. (eds), *Communications in Computer and Information Science: Vol. 536. International Symposium on Security in Computing and Communication* (pp. 438-452). Springer. http://dx.doi.org/10.1007/978-3-319-22915-7_40
- Zeng, W., & Germanos, V. (2019). *Modelling hybrid cyber kill chain*. Retrieved from <https://ceur-ws.org/Vol-2424/paper10.pdf>
- Zetter, K. (2010a). *Google to stop censoring search results in China after hack attack*. Retrieved December 8, 2021, from <https://www.wired.com/2010/01/google-censorship-china/>

- Zetter, K. (2010b). *Google attack was ultra sophisticated, new details show*. Retrieved December 8, 2021, from <https://www.wired.com/2010/01/operation-aurora/>
- Zetter, K. (2010c). *Hack of Google, Adobe conducted through zero-day IE flaw*. Retrieved December 8, 2021, from <https://www.wired.com/2010/01/hack-of-adob/>
- Zetter, K. (2013). *Hackers who breached Google in 2010 accessed company's surveillance database*. Retrieved December 9, 2021, from <https://www.wired.com/2013/05/google-surveillance-database/>
- Zetter, K. (2014). *Countdown to Zero Day: Stuxnet and the launch of the world's first digital weapon*. New York: Broadway books.
- Zhang, R., Huo, Y., Liu, J., & Weng, F. (2017). Constructing APT attack scenarios based on intrusion kill chain and fuzzy clustering. *Security and Communication Networks*, 2017. <https://doi.org/10.1155/2017/7536381>

APPENDIX 1 PRIMARY RESEARCH MATERIAL

APT 10 (62 pages)

1. <https://intrusiontruth.wordpress.com/2018/08/15/apt10-was-managed-by-the-tianjin-bureau-of-the-chinese-ministry-of-state-security/>
2. https://adeo.com.tr/wp-content/uploads/2020/02/APT10_Report.pdf
3. <https://www.virusbulletin.com/virusbulletin/2020/03/vb2019-paper-defeating-apt10-compiler-level-obfuscations/> or <https://web.archive.org/web/20220122233010/https://blogs.vmware.com/security/2019/02/defeating-compiler-level-obfuscations-used-in-apt10-malware.html>

APT 27 (56 pages)

1. <https://www.secureworks.com/research/a-peek-into-bronze-unions-toolbox>
2. <https://www.secureworks.com/research/bronze-union>
3. <https://www.secureworks.com/research/threat-group-3390-targets-organizations-for-cyberespionage>

APT31 (87 pages)

1. <https://www.mandiant.com/resources/definitive-dossier-of-devilish-debug-details-part-one-pdb-paths-malware>
2. <https://threatpost.com/microsoft-offers-analysis-of-zero-day-being-exploited-by-zirconium-group/124600/>
3. <https://research.checkpoint.com/2021/the-story-of-jian/>
4. <https://blog.confiant.com/uncovering-2017s-largest-malvertising-operation-b84cd38d6b85>
5. <https://blog.sekoia.io/walking-on-apt31-infrastructure-footprints/>
6. <https://blog.confiant.com/zirconium-was-one-step-ahead-of-chromes-redirect-blocker-with-0-day-2d61802efd0d>

APT41 (133 pages)

1. <https://arstechnica.com/information-technology/2018/05/researchers-link-a-decade-of-potent-hacks-to-chinese-intelligence-group/>
2. <https://www.infosecurity-magazine.com/news/chinas-apt41-manages-library/>
3. <https://blogs.blackberry.com/en/2021/10/drawing-a-dragon-connecting-the-dots-to-find-apt41>
4. https://www.welivesecurity.com/wp-content/uploads/2019/10/ESET_Winnti.pdf

5. <https://www.ptsecurity.com/ww-en/analytics/pt-esc-threat-intelligence/higaisa-or-winnti-apt-41-backdoors-old-and-new/>
6. https://www.kaspersky.com/about/press-releases/2019_operation-shadowhammer-new-supply-chain-attack
7. <https://content.fireeye.com/apt-41/rpt-apt41/>

Hafnium (17 pages)

1. <https://www.microsoft.com/security/blog/2021/03/02/hafnium-targeting-exchange-servers/>
2. <https://therecord.media/white-house-formally-blames-chinas-ministry-of-state-security-for-microsoft-exchange-hack/>

Mustang Panda (70 pages)

1. <https://www.crowdstrike.com/blog/meet-crowdstrikes-adversary-of-the-month-for-june-mustang-panda/>
2. <https://therecord.media/indonesian-intelligence-agency-compromised-in-suspected-chinese-hack/>
3. <https://www.anomali.com/blog/covid-19-themes-are-being-utilized-by-threat-actors-of-varying-sophistication>
4. <https://www.avira.com/en/blog/new-wave-of-plugx-targets-hong-kong>
5. <https://www.anomali.com/blog/china-based-apt-mustang-panda-targets-minority-groups-public-and-private-sector-organizations>
6. <https://www.secureworks.com/research/bronze-president-targets-ngos>

RedDelta (41 pages)

1. <https://go.recordedfuture.com/hubfs/reports/cta-2020-0728.pdf>
2. <https://go.recordedfuture.com/hubfs/reports/cta-2020-0915.pdf>
3. <https://www.mcafee.com/blogs/other-blogs/mcafee-labs/operation-dianxun-cyberespionage-campaign-targeting-telecommunication-companies/>
4. <https://www.proofpoint.com/us/blog/threat-insight/ta416-goes-ground-and-returns-golang-plugx-malware-loader>

TA413 (25 pages)

1. <https://www.proofpoint.com/us/blog/threat-insight/chinese-apt-ta413-resumes-targeting-tibet-following-covid-19-themed-economic>
2. <https://www.proofpoint.com/us/blog/threat-insight/ta413-leverages-new-friarfox-browser-extension-target-gmail-accounts-global>

TEMP_Heretic (15 pages)

1. <https://www.volexity.com/blog/2022/02/03/operation-emailthief-active-exploitation-of-zero-day-xss-vulnerability-in-zimbra/>

Tonto Team (82 pages)

1. <https://www.fireeye.com/content/dam/fireeye-www/summit/cds-2019/presentations/cds19-executive-s08-achievement-unlocked.pdf>
2. <https://blog.talosintelligence.com/2020/03/bisonal-10-years-of-play.html>
3. https://www.trendmicro.de/cloud-content/us/pdfs/security-intelligence/white-papers/wp_the-heartbeat-apt-campaign.pdf tai
toimiva
https://web.archive.org/web/20220309025602/https://www.trendmicro.de/cloud-content/us/pdfs/security-intelligence/white-papers/wp_the-heartbeat-apt-campaign.pdf

Winnti Group (85 pages)

1. <https://web.archive.org/web/20180721031609/https://401trg.com/burning-umbrella/>
2. https://www.trendmicro.com/en_us/research/17/d/pigs-malware-examining-possible-member-winnti-group.html#
3. <https://www.blackberry.com/us/en/pdfviewer?file=/content/dam/blackberry-com/asset/enterprise/pdf/direct/report-bb-decade-of-the-rats.pdf>
4. <https://securelist.com/winnti-more-than-just-a-game/37029/>

APPENDIX 2 CHINESE CYBERSPACE ACTORS

Gathered from Raud (2016), Boszkov (2020) and IISS (2021).

Actor	Abbreviation	Affiliated	Information
Politburo Standing Committee, State Council and Central Military Commission			The highest-level of decision-makers
Central Discipline and Inspection Commission of the CCP			Collects intelligence on leading members of CCP
Ministry of Foreign Affairs	MFA		Cyber diplomacy, rulemaking for cyberspace
Ministry of Industry and Information Technology	MIIT	Operates under the State Council	Manages China's telecommunications, ICT and network infra
National Computer Network Emergency Response Technical Team/Coordination Centre of China	CNCERT	Subordinate of CAC	Deals with improving China's overall cybersecurity posture
State Administration for Science, Technology and Industry for National Defence	SASTIND	Operates under MIIT	
Ministry of Public Security	MPS	Operates under the State Council	Responsible for implementing Internet traffic control in China together with MIIT, concentrates on content, critical infra protection together with CAC
Ministry of State Security	MSS	Operates under the State Council	Foreign and domestic intelligence, to ensure domestic stability
State Encryption Bureau		Operates under the State Council	
State Secrets Bureau		Operates under the State Council	
Chinese Institute of Contemporary International Relations		Operates under MSS	
Chinese Academy of Engineering		Operates under MSS	

Chinese Academy of Science		Operates under MSS	
Tsinghua University			
Peking University			
Academy of Military Science		Operates under PLA	
PLA Information Engineering University		Operates under PLA	
Cyber Security Association of China			
Central Commission for Cybersecurity and Informatization	CCCI	Operates under the direction of President XI	Centralised decision-making body on cyberspace and ICT affairs
State Information Leading Group	SILG	Operates under the State Council	
State Network and Information Coordination Small Group	SNISCSG	Operates under the State Council	
Central Leading Small Group for Internet Security and Informatisation	CLSGISI	Operates under the direction of President XI	
Cyber Administration of China	CAC		Shapes tenets of China's cybersecurity policy, secretariat of CCCI
Cybersecurity Association of China	CSAC		
State Council Information Office	SCIO		
NA29:D41ational Information Security Standardization Technical Committee	TC260		Enforce Cybersecurity laws to lower levels
National Computer Network and Information Security Management Centre	NCNISMIC	Associated with CNCERT, subordinate of CAC	Technical responsibility for the deployment and maintenance of China's censorship system (The Great Firewall of China)
Informatisation Department			
Strategic Support Force	SSF		Oversees China's development of offensive intrusion capabilities, cyber defence, military intelligence and cyber exploitation reconnaissance, 3/PLA and

			4/PLA functions are centered to SSF
Cyber Strategic Intelligence Research Centre		Provides support for SSF	
3rd General Staff Department	3/PLA		
Science and Technology Intelligence Bureau	STIB	Managed together with 3/PLA	
Science and Technology Equipment Bureau	STEB	Managed together with 3/PLA	
56th Research Institute		Managed by 3/PLA, STIB and STEB. Responsible for computing, sensor technology and cryptography	
57th Research Institute		Managed by 3/PLA, STIB and STEB. Responsible for computing, sensor technology and cryptography	
58th Research Institute		Managed by 3/PLA, STIB and STEB. Responsible for computing, sensor technology and cryptography	
National Research Centre for Information Security Technology		Managed by or is affiliated with 3/PLA	
Information Security Research Institute		Managed by or is affiliated with 3/PLA	
PLA Communications Security Bureau		Managed by or is affiliated with 3/PLA	
12 Operational Bureaus		Managed by or is affiliated with 3/PLA, the most important role in 3/PLA structure	
2nd Bureau		Unit 61398	
12th Bureau		Unit 61486	
Military Region Technical Reconnaissance Bureaus	TRB		
4th General Staff Department	4/PLA		
54th Research Institute		Operated by 4/PLA	
PLA Electronic Engineering Academy		Operated by 4/PLA	

Red Hacker Alliance		Tolerated by the government	Several hundred thousand members
Cyber Militias		Part of a enhanced civil-military cooperation	Over eight million citizens
Qihoo 360		Under supervision of the Central Commission for Integrated Military and Civilian Development	Builds up China's cyber defence systems for military-related Internet services and enhances PLA's cyber threat awareness
Baidu, Alibaba, Tencent, JD.com	BATJ		Chinese companies with a vital role in management and control of public information spaces

APPENDIX 3 CHINA'S CYBER ESPIONAGE APPARATUS

Retrieved from Bozhkov, 2020, p. 7.

China's cyber espionage apparatus

Suspected China-nexus groups and their possible strategic intent

Circles refer to advanced persistent threat groups (APT) identified, by the original industry sources, as having a nexus to China. Based on their campaign activities, those actors' operational focus and targeting patterns are driven by multifold collection requirements in support of bolstering domestic manufacturers' expansion, domestic control, long-term state-led economic objectives, and current geopolitical events.

Note that neither the threat actors nor their targeting rationales are monolithic, and rigidly aligned to state interests. Rather, this threat landscape is evolving and dynamic, operationally, strategically and institutionally.

This list is non-exhaustive, and solely based on open-source reporting. Note that may be overlaps between different actors, clusters and activities, and that institutional affiliations are exclusively based on evidence of criminal indictments by the US Department of Justice.



Economic espionage

Directed collection

- Bilateral trade and foreign policy negotiations
- State-subsidised strategic economic plans and 'core' technologies
- self-reliance and indigenous innovation
- "secure and controllable" supply chains
- Business negotiations involving domestic manufacturers
- Infrastructure investment projects and the Belt and Road Initiative (BRI)



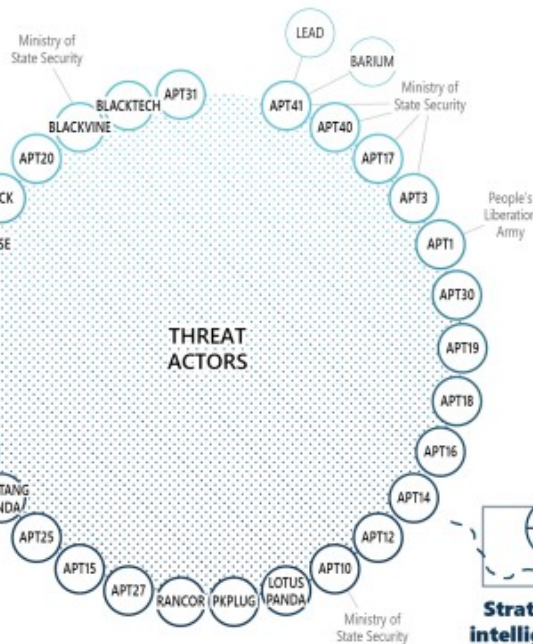
National security espionage

Strategic access

Counterintelligence and targeted surveillance
Office of Personnel Management (OPM)
Anthem Inc.
Marriott International
Equifax

Upstream data entries
Telecoms
Managed Service Providers (MSPs)
Internet Service Providers (ISPs)
data centers
law firms

Internal politics, 'social stability' and CCP narrative control
Hong Kong
Taiwan
Macau
Xinjiang Province
Tibet



Strategic intelligence collection

Positional access

- Military intelligence objectives related to the South China Sea
- Development of indigenous (dual-use) maritime, aerospace, defense technologies
- Regional power projection