Multimedia Tools and Applications (2020) 79:29923–29949
https://doi.org/10.1007/s11042-020-09444-x

MCKC: a modified cyber kill chain model for cognitive APTs analysis within Enterprise multimedia network

Ankang Ju[1] · Yuanbo Guo[1] · Tao Li[1]

Received: 29 July 2019 / Revised: 7 June 2020 / Accepted: 28 July 2020 /
Published online: 13 August 2020

Abstract
The emerging cyber security threats pose many challenges to security analysts of enterprise multimedia environments when analysts attempting to analyze and reconstruct advanced persistent threats (APTs). APTs analysis activities are both time-consuming and labor-intensive. Attack modeling technology represented by kill chain can reduce the burden of manual provenience analysis. However, existing Cyber Kill Chain models represent attacks as several stages solidly, and they cannot reflect the characteristics of progressive penetration. It is difficult for security analysts to automate the correlation analysis of attack events in practical usage. In this paper, we first analyze current Cyber Kill Chain models and heterogeneous data sources for APTs detection. Then we propose MCKC (Modified Cyber Kill Chain model) that can be used for standardized correlation analysis. MCKC organizes sub-chains into a recursive structure, and different kill chain penetration processes in the same attack scenario are better connected The proposed MCKC model offers a novel approach for bi-directional attack analysis: forward analysis and backward reasoning which can facilitate threat detection effectively without relying too much on expert knowledge. The advantage of MCKC model is that it is more suitable for cognitive reasoning and APTs scenario reconstruction. Compared with existing models MCKC gives a feasible technological process for threat analysis. The result of case study shows that the modified kill chain model is effective in discovering security events and reconstructing APT attacks.

Keywords Cyber kill chain model · APT detection · Bi-directional analysis

1 Introduction

In the last decades, the dependence throughout modern societies on information and communication technology (ICT) has continued to rise [3]. Vulnerabilities in the supporting ICT

✉ Ankang Ju
jusissp@yeah.net

[1] Zhengzhou Institute of Information Science and Technology, Zhengzhou 450001, China

assets threatened the cyber activities that are performed within modern societies. Organizations need to protect their assets against a variety of threat actors that range from cyber criminals to nation states. At the more advanced and persistent end of this threat actor spectrum, actors are often described as Advanced Persistent Threats (APTs).

As can be seen from the recent APT attacks [4], the enterprise multimedia networks have become the main target of attacks. The main reason is that the value of data information within enterprise multimedia networks attract hackers to launch network attacks.

The network threat brings potential risk for the multimedia enterprise network, and it faces complex security challenges. APTs mainly aims to destroy the inter-action and synchronization of multimedia system. Multimedia enterprise network faces security risk of data theft, user information disclosure, worm infection, and data encryption ransomware. Attackers gain economic benefits by means of in-formation theft, data encryption, and ransomware. For example, Sony suffered the APT attack launched by hacker groups in 2014, the information of tens of thou-sands of employees and many unpublished movie copies were leaked. Sony's PlayStation Network service and streaming media service Qriocitywere "externally attacked", which were suspended for several days.

As for the enterprise multimedia network, it has specific characteristics [13]. First, there is a need for both internal and external network communication. Besides, it has distinctive features of specific application. It often provides services to internal networks, and internal services are not exposed to the Internet. Besides, a wide range of information and services are available on an organization's internal network that is unavailable to the public internet. So, there is a higher security and confidentiality requirements within the enterprise multimedia network. In addition, most attacks against enterprise multimedia network are APTs. These targeted and stealthy cyberattacks bring more serious security risk to the enterprise multimedia network. At the same time, spot the events related to attack from huge amount of data is needle in a haystack. How to detect anomalies from raw data of heterogeneous data sources is a hot topic in current research.

Existing intrusion detection methods, such as firewalls, intrusion detection systems, and anti-virus software, have made great progress. However, they still cannot effectively deal with APT attacks. The existence of 0-day attacks makes intrusion detection more unpredictable, and a huge number of false positives bring difficulties in attack analysis. The attack modeling method summarizes the expert knowledge [15] of attack implementation, and it makes attack investigation more convenient. Our goal is to use provenance tracking to determine causal relationships between different alarms to reconstruct the attack scenario, and to do so without relying too much on manual analysis.

Traditional detection schemes focus on one or more stages. They cannot achieve comprehensive attack detection and have a high false negative rate and false positive rate. Although there has been a lot of research and made great progress, targeted cyber-attacks are still occurring constantly. The shortcoming is that the existing method still depends on manual analysis. It cannot identify and respond quickly. Here we want to reduce human participation and make the detection process as intelligent as possible. Similar attacks should be detected according to the attack pattern. Data parsing should be as automatic as possible and be detected in an adaptive way. The intelligence of attack detection is mainly embodied in several research points.

(1) Accuracy. A large number of false alerts pose a big challenge in detection. Even if the false alarm rate is merely 1%, a large amount of data can bring a lot of alerts, which will bring enormous burdens to managers. Reducing false positive and false negative is an important content in intrusion detection.

(2) Efficiency. Fast identification of attack behavior, shortening attack detection time, and minimizing the damage caused by the attack is also important in practical application.

(3) Intelligence. Automated attack inference process. Reducing manual labor costs on irrelevant information can help human analysts put artificial judgment in the key step.

The implementation process of APT can usually be described by Cyber Kill Chain model (CKC) [9]. The Cyber Kill Chain model decomposes the APTs into a number of phases to understand the attacker's goals and modes of operation. The Cyber Kill Chain model is used in various ways as an analysis model that explains the implementation process of advanced persistent threats (APTs). Using this model, the processes of complex, advanced persistent threats within the enterprise multimedia network can be easily understood, and targeted measures can be conducted to block the kill chain at each stage. Different from IDS, the Cyber Kill Chain model provides the method of correlation analysis from the perspective of attack modeling. IDS uses anomaly detection and misuse detection technology and gives indicators of attacks. CKC and IDS focus on different objects. However, there has been limited progress in the study of cognitive attack analysis. Even variants of kill chain models still cannot describe the attack provenance tracing methods, especially in the field of attack correlation analysis. Moreover, researchers continue to argue that kill chain models are insufficient to describe threats within enterprise multimedia network.

To overcome the shortage of IDS and traditional defense methods, we study typ-ical APT attacks in APTNotes, which is a repository for various publicly-available documents and notes related to APT. We aim at revising the kill chain model to make it easier for cognitive analysis, so as to improve the efficiency of APT attack analysis.

There are two advantages to help reduce the manual analysis with the usage of our MCKC model. On the one hand, it is convenient for the analyst to recognize the attack and reduces the time that needs to be spent at the beginning of the analysis. It is convenient for the analyzer to quickly map the event into the attack chain, especially for the attack scenarios make people confused. On the other hand, it supports a bidirectional analysis model, which facilitates the process of APT attack investigation, so as to speed up the analysis process and reduce the pressure of manual attack analysis.

In this paper, we revise the Cyber Kill Chain model and generalize APT attacks into 5 detectable stages, of which the *Action* stage can be the fulfillment of attack goals or the follow-up conducted kill chains. The attack process is generalized as a recursive structure, and the structure of attack conduction is organized as a chain of the sub-chains. Based on the model, a bidirectional analysis method is proposed, which can be used to analyze APT attacks in both forward and backward directions. The method proposed in this paper explicitly maps attack events to kill chain phases and organizes them into attack scenarios. At the same time, the missing attack events can be supplemented by backward reasoning. The case study shows that it is much easier for attack provenance tracing.

The paper is organized into 4 sections. Section 2 explains the term advanced persistent threat and lists existing Cyber Kill Chain models. The technical characteristics of heterogeneous data sources for APT detection is also discussed in Section 2. Subsequently, the proposed Modified Cyber Kill Chain model (MCKC) is introduced. The section 4 gives an example of bi-directional analysis based on the MCKC model. The paper ends with concluding remarks in Section 5.

2 Related work

2.1 APTs within enterprise multimedia environment

An Advanced Persistent Threat (APT) [16] is a set of stealthy and continuous computer hacking processes, often orchestrated by a person or persons targeting a specific entity. Compared with traditional attack types, APT attacks have the following characteristics: (1) Attack behaviors are difficult to extract. As APT attacks usually exploiting with 0-day vulnerabilities to obtain root privileges, traditional feature-based detection methods always lag the implementation of the attacks. (2) Attackers are good at concealment. They hide network behaviors via covert channel or encrypted network communication to evade detection. Traditional intrusion detection system and security audit system are difficult to detect. (3) Another characteristic is the long duration. From initial information gathering to data theft and transmission, APT attacks often take months or even longer.

APT attacks are difficult to detect because of the application of sophisticated techniques and convert behaviors. A Fireeye report said that, the global median time from compromise to discovery is up from 99 days in 2016 to 101 days in 2017 [8], The average detection cycle of APT attacks is still more than 3 months, much longer than accomplishment of the targets. Traditional defense methods relying on real-time detecting and blocking do not work effectively when coping with APT attacks. In most cases, it still relies on manual attack analysis. APT attacks have become the main factors affecting network security. In particular, it has brought huge losses to the large organizations and enterprises, especially in governments, commercial, financial institutions.

Numerous attack modelling techniques have been proposed for cyber attack analysis such as Attack Graph, Attack Tree, dependence graph, Attack Surface, the diamond model, and Cyber Kill Chain. The Cyber Kill Chain model, as well as other attack lifecycle models, can help defenders understand the increasingly complex attacks that they are facing.

The APT attacks are composed of multiple interdependent phases that compromise hosts as a stepping stone for launching the next phase. Typically, an APT attack is defined as a seven-stage process:reconnaissance, weaponization, delivery, exploitation, installation, command and control, actions.

Existing defense methods against APT attacks include IDS, traffic analysis, sandbox, honeypot, and big data analysis. They have different characteristics in defense of APT attacks. Table 1 shows the comparisons of different defense technologies.

Heterogeneous security data from multiple sources is the key to APTs detection. The diversity of these data poses many challenges to analysts. Bryant Kill Chain represents the relationship of various data sources in indicators at each stage [6], such as firewall, IDS, IPS, anti-virus, endpoint OS logs, directory logs, NetFlow data. And anomalies generated from the original log data: probing, enumeration, host access, network delivery, host delivery, software modification, privilege escalation, privilege use, internal reconnaissance, lateral movement, data manipulation, obfuscation, external data transfer.

For a typical enterprise multimedia network (Fig. 1), there are external network, DMZ zone, and internal network. Various entities in the network produce data such as logs, alerts. The introduction of security knowledge and threat intelligence further increases the diversity of data and the complexity of data processing. Here we give an example typical enterprise multimedia network for heterogeneous security data acquisition.

Table 1 Comparisons of different defense technologies

Defense methods	Pros	Cons
IDS	High efficiency. It is the most effective detection method at present.	Unable to find unknown attacks.
Traffic analysis	It can automatically discover abnormal behavior.	Low accuracy of algorithms.
Sandbox	Able to find unknown threats.	Low efficiency, excessive consumption of local resources.
Honeypot	Active detection method for attacks and malicious behavior detection.	Poor correlation ability and high rate of false alert.
Big data analysis	Find APT attack traces from data mining and statistical analysis	Depend on the ability of data collection and analysis. Need to overcome the problem of information island.

2.2 Cyber kill chain models

Cyber Kill Chain model (CKC) [1] is a process-based model used by cybersecurity analysts to analyze APT attacks in a chained manner. Cyber Kill Chain is a structured attack model, since the hacker progresses the attack in an ordered chain according to the plan. It is used to describe some attack steps within a counter measure framework. Such a model enables the analysts to deal with smaller and easier analytical problems, it also helps the defenders to design and develop prevention measures for each stage.

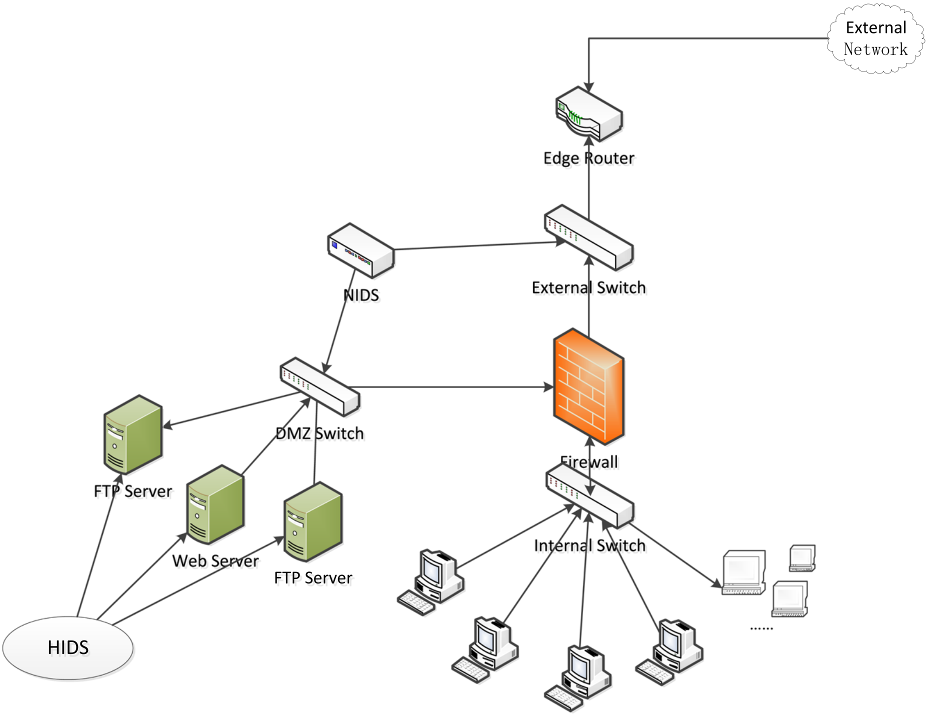

Fig. 1 Typical enterprise multimedia network

Cyber Kill Chain summarized a model of offensive actions of a cyber attack. Studying Cyber Kill Chain is utmost important for system defenders, as it helps them learn what an attacker is plotting. Threat modeling methods are required to perform a structured analysis of APTs, and the Cyber Kill Chain put forward by Lockheed Martin is frequently regarded as the industry standard model for defending against APTs. The Cyber Kill Chain model has been used by many organization and authorities for many years.

2.2.1 Lockheed Martin kill chain model

Inspired by phase-based model presented by military establishment, which models military action with stages find, fix, track, target, engage, and assess, Lockheed Martin proposed intrusion Kill Chain model in 2011 [9]. As shown in Fig. 1, all steps of an adversary and its targets are described by a series of events. More concretely, Lockheed Martin intrusion Kill Chain is defined as Reconnaissance, Weaponization, Delivery, Exploitation, Installation, Command and Control (C2), and Actions on Objectives. Every kill chain stage represents a weak indicator of compromise (IoC), a signature which is used to detect a potentially compromised system.

Most existing intrusion detection method (such as IDS, firewall, anti-virus software) will provide knowledge about a single intrusion phase, but they do not illustrate the relationship between phases. Lockheed Martin Kill Chain is proposed after thorough analysis of successful compromises. It helps to identify and prevent intrusions. If the intrusion is stopped at any step it will break the Kill Chain and prevent the intruder from completing their objective. However, defense the intrusion requires more cost and complexity.

Lockheed Martin Kill Chain is a model describing intrusion attempt. It analyzes the principle and process of an attack. Lockheed Martin Kill Chain model help to analyze the intrusion in a systematization way. This Kill Chain analysis method is a guidance to understand APT attacks and help to reconstruct intrusion. Lockheed Martin Kill Chain model provide a basis for subsequent models. However, it's just a conceptual guide and insufficient in attack detection. It only acts as a postmortem verifying explanation in practical analysis.

Lockheed Martin Kill Chain also summarized the action matrix of APT attack detection [9], it integrated the existing basic security audit mechanism into each stage of the kill chain. Supplementary version of the action matrix is listed in Table 2.

Table 2 The action matrix of APT attack detection

Phase	Detect	Deny	Disrupt	Degrate	Deceive	Destroy
Reconnaissance	Web Analysis	Firewall ACL	encryption	access control	Honeypot	
Weaponization	NIDS	NIPS			Honeypot	
Delivery	Vigilant User	Proxy filter	In-line AV	Queuing	Honeypot	
Exploitation	HIDS	Patch	DEP	permission restriction	Honeypot	
Installation	HIDS	"chroot" jail	AV	access control	Honeypot	
C2	NIDS	Firewall ACL	NIPS	Tarpit	DNS redirect	
Actions on Objectives	Audit log, Trusted computing	Trusted computing, encryption		Quality of Service	Honeypot	

2.2.2 Mandiant attack lifecycle model

Another famous Kill Chain model is Mandiant attack lifecycle [12]. It focuses on internal network activities, it defines the entire attack lifecycle as: initial reconnaissance, initial compromise, establish foothold, escalate privileges, internal reconnaissance, move laterally, maintain presence, complete mission. Mandiant attack lifecycle considers escalate privileges, internal reconnaissance, move laterally, maintain presence as an internal loop process. Internal loop process together with initial intrusion, as well as final mission complete makes up a whole attack lifecycle.

The biggest difference between Lockheed Martin Kill Chain and Mandiant Attack Lifecycle is, the former didn't show lateral movement process, while the latter takes internal movement into account and gives a procedural description. The combination of these two models is a reasonable description of APT attacks, and the subsequent models are improved on this foundation.

Mandiant Attack Lifecycle considers APT attack as an external intrusion. Firstly, an APT attack will build foothold within internal network and wait for the opportunity of reaching the attack target. it doesn't translate each indicator into group, and still brings heavy burden to security analysts (Figs. 2, 3, 4, 5, 6, 7 and 8).

2.2.3 Diamond model

Different from Kill Chain models, the Diamond Model of Intrusion Analysis provides a formalized way to characterize network intrusions [7]. The Diamond Model gets its name from the fundamental data structure it uses to describe intrusion events.

Diamond model describes that an adversary deploys a capability over the infrastructure against a victim. The meta-features are: timestamp (both start and end), phase, result, direction, methodology and resources. The meta-features are used to order events within an activity thread.

Security analysts use the Diamond model's vertices to discover and detect events [1]. These vertices are connected by edges which illustrate the natural relationships between the features. By pivoting within vertices and across edges, analysts reveal more information about an adversary or the adversary's operations, and can discover new infrastructure, capabilities, and victims. The meta-features shown in Fig. 3 are used to capture critical knowledge, when possible, about times of intrusion (both beginning and end), phase, result, direction, methodology, and resources. This kind of model is useful and important when the organization are dealing with more advanced attacker [1]. This model has been chose to be included in this study because of the simplicity of the implementation of the models.

2.2.4 MITRE ATT&CK model

Adversarial Tactics, Techniques, and Common Knowledge (ATT&CK™) for Enterprise is an adversary model and framework for describing the actions an adversary may take to compromise and operate within an enterprise network.

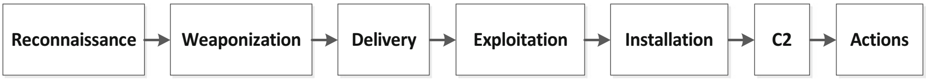

Fig. 2 Lockheed Martin Kill Chain Model

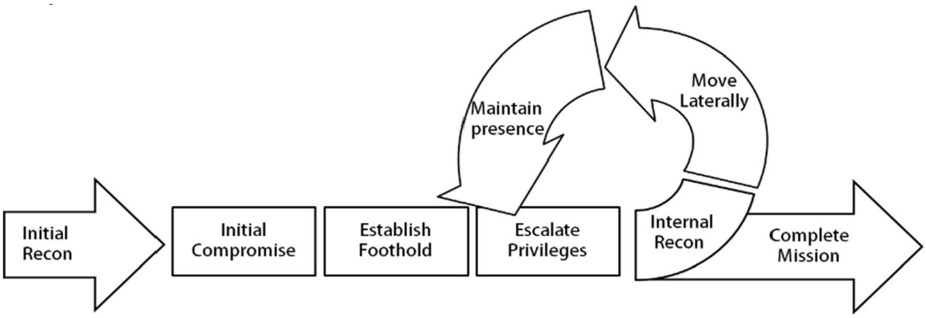

Fig. 3 Mandiant Attack Lifecycle Model

MITRE ATT&CK Model [14] is a more detailed analysis model, it describe 11 tactic categories: initial access, execution, persistence, privilege escalation, defense evasion, credential access, discovery, lateral movement, collection, exfiltration, command and control. The application of this model in practical analysis is more convenient, and it provides detailed analysis of attack implementation. However, the problem of this model is that the model is too complex, and the details of attack process expressed by it are too cumbersome, which increases the complexity of attack analysis to a certain extent.

2.2.5 Malone kill chain model

As mentioned before, Lockheed Martin Cyber Kill Chain model is excellent for describing external attacks, but doesn't exactly work for insider threats. Besides, it is incomplete and can lead to over-focusing on perimeter security, to the detriment of internal security controls.

In RSAC 2016, Sean T Malone (https://www.blackhat.com/docs/us-16/materials/us-16-Malone-Using-An-Expanded-Cyber-Kill-Chain-Model-To-Increase-Attack-Resiliency.pdf) presented an expanded model including the Internal Kill Chain and the Target Manipulation Kill Chain. The primary limiting factor of the traditional Cyber Kill Chain is that it ends with Stage 7: Actions on Objectives, conveying that once the adversary reaches this stage and has access to a system on the internal network, the defending victim has already lost. In reality,

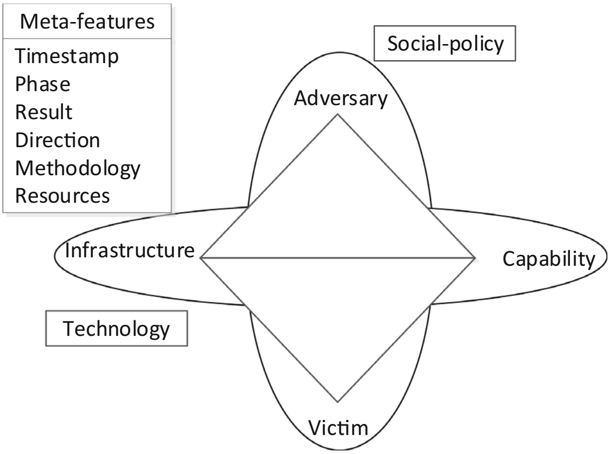

Fig. 4 Diamond Model

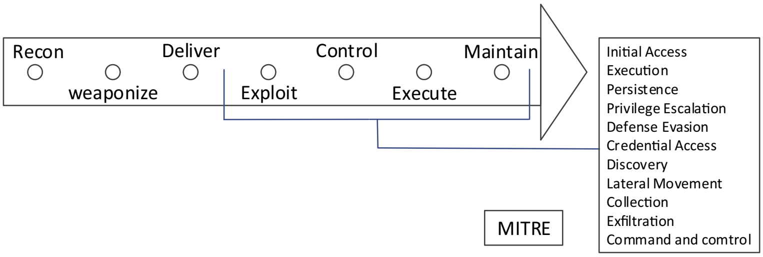

Fig. 5 MITRE ATT&CK Model

there should be multiple layers of security zones on the internal network, to protect the most critical assets.

The adversary often has to move through numerous additional phases in order to access and manipulate specific systems to achieve his objective. By increasing the time and effort required to move through these stages, we decrease the likelihood of the adversary causing material damage to the enterprise.

2.2.6 The unified kill chain model

In December 2017, Paul Pols (https://www.csacademy.nl/images/scripties/2018/Paul-Pols%2D%2D-The-Unified-Kill-Chain.pdf) proposed the unified Kill Chain Model, designing a unified kill chain for analyzing comparing and defending against cyber-attacks. In this paper, the author advantages and disadvantages of Laliberte's Kill Chain, Nachreiner's Kill Chain, Bryant's Kill Chain, Malone's Kill Chain, and MITRE ATT&CK Model. Through induction and complementation, the kill chain is divided into 18 steps: Reconnaissance, Weaponization, Defense Evasion, Social Engineering, Delivery, Exploitation, Persistence, Command & Control, Pivoting, Privilege Escalation, Discovery, Lateral Movement, Execution, Credential Access, Action on Objectives, Target Manipulation, Collection, Exfiltration. The unified kill chain model provides attack analysis matrix for Red Team evaluation. It can improve the predictive value of the Red Team assessments.

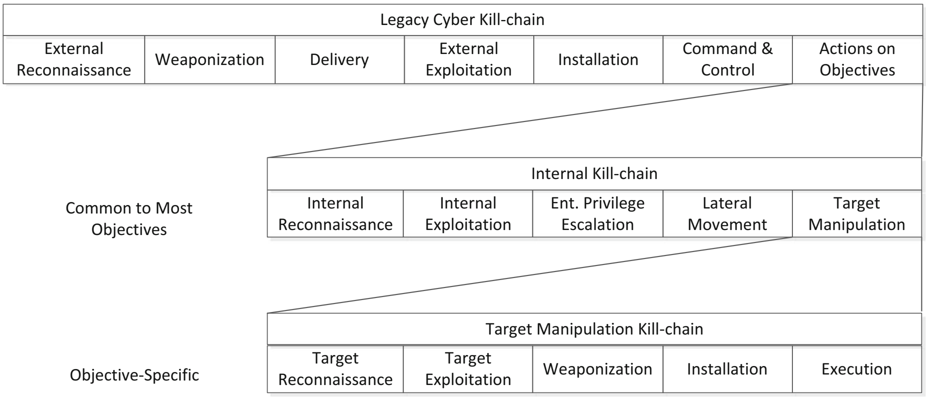

Fig. 6 Malone Kill Chain Model

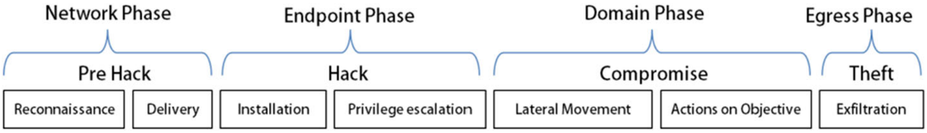

Fig. 7 Bryant Kill Chain Model

2.2.7 Bryant kill chain model

In view of the shortcomings of the above two models (Lockheed Martin Kill Chain and Mandiant Attack Lifecycle), Bryant [6] devised a new Kill Chain model and make it more suitable for attack analysis, Bryant Kill Chain includes 7 phases: reconnaissance, delivery, installation, privilege escalation, lateral movement, actions on objective, exfiltration. And furtherly, these 7 phases can be divided into 4 distinct macro phases: network, endpoint, domain, and egress.

The initial 7 phases of the Bryant Kill Chain were deconstructed into 13 subtasks in order to reflect slight variations between data and attacker behaviors. And each subtask corresponds to different indicators, which give more details about the attack.

The main enhancement made by Bryant Kill Chain is the addition of lateral movement and exfiltration. Bryant Kill Chain makes the attack scenario completer and more meticulous. It aggregates related events into robust alarms and decreases the number of redundant alarms. It uses SQL queries in relational databases to aggregate related data. The specific process of lateral movement as a step in Kill Chain is not appropriate, but it does not reflect the stages of internal exploitation.

2.2.8 Khan kill chain model

Marc Liliberte presents a cognitive and concurrent Cyber Kill Chain model in 2018 [10]. The author adds lateral phase after the command and controls stage, which emphasizes the lateral movement between intermediate nodes within internal network. Furthermore, Khan Kill Chain Model reformed the Cyber Kill Chain model on the basis of Marc Liliberte model, and make it much easier for cognitive analysis.

Firstly, Kill Chain is changed from the original seven stages to the combination of four stages, and the four stages do not differentiate in order. The whole model is regarded as a cycle. Each time point corresponds to an object and a model, and these models are concatenated in series by a timeline.

The difference between Khan model and traditional Kill Chain lies in two aspects. For one thing is that the seven stages of traditional Kill Chain have the same motivation, while in Khan model the first steps are combined into one R step. The second is the sequential relation

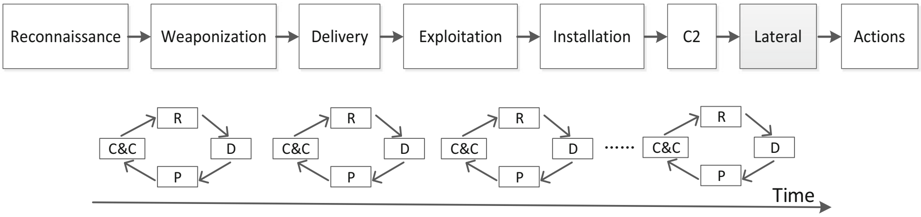

Fig. 8 Marc Liliberte Kill Chain Model

between stages. Khan model is more a sequential analysis model and it is more convenient for metal analysis.

2.2.9 Comparison

As mentioned above, previous research has done good basic work. The Cyber Kill Chain process provides a method to model the sequential steps taken by an adversary to infiltrate target system and either cause permanent damage or infiltrate data from it without leaving a trace. However, the ever-changing threat landscape requires modification in the standard Cyber Kill Chain as well. This is required to accommodate the complex nature of new cyber threats.

Table 1 below compares the features and limitations of the characteristics and threats of the previous models and the proposed modified model (Tables 3, 4 and 5).

In these attacks multiple kill chains can be formed at a time and each one of them can have numerous indicators of attacks. it does not only completely cover all stages of a Cyber Kill Chain but, implements several copies of these chains in parallel and over long-time duration to achieve the desired objective. As for an organized targeted attack, there are multiple possible permutations of the kill chain which can be in effect simultaneously using lateral movement. It is vital for the security researchers to consider the cyber-defense strategies from a holistic point of view rather than the conventional single kill chain model.

However, some researchers [11] believe that methods such as kill chains and the diamond model are not designed to represent the fundamental cyber-attack constructs. Methods such as kill chains and the diamond model are popular amongst the business community but not necessarily amongst the academic community.

Table 3 Comparison of Cyber Kill Chain models

Model	Year	Features	Limitations
Lockheed Martin Kill Chain Model	2011	APT attacks are defined as seven steps for the first time	only for verification and interpretation, lack lateral movement process
Mandiant Attack Lifecycle Model	2013	procedural description of internal network activities	lack of analytical methods
Diamond Model	2013	a formalized way to characterize network intrusions	lack of vulnerability assessment
MITRE ATT&CK Model	2013	more detailed, 11 tactic categories	complex, and the details of attack process expressed by it are too cumbersome,
Malone Kill Chain Model	2016	expand Cyber Kill Chain model into internal kill chain and targeted manipulation kill chain	the common characteristics of internal and external kill chains are not emphasized.
The Unified Kill Chain Model	2017	improve the predictive value of the Red Team assessments	fixed attack description
Bryant Kill Chain Model	2017	decreases the number of redundant alarms, uses SQL query to aggregate related data.	lateral movement does not reflect the stages of internal exploitation
Khan Kill Chain Model	2018	emphasizes the lateral movement between intermediate nodes, attack events are concatenated in series by a timeline	the relationship between sub kill chains is not reflected

Table 4 Fields in *event*

Field	description
e_id	event id
timestamp	event timestamp
pid	process id
pname	Process name
ppid	Parent Process ID
ppname	Parent Process name
sip	source IP address
sport	source port number
dip	destination IP address
dport	destination port number
e_type	type of event
desc	description of event type
source	The source of event

3 Modified cyber kill chain analysis model—MCKC

As discussed in section 2, the traditional Cyber Kill Chain model is more an abstract model. Although there exists several improved models (as shown in section 2), there are still many shortcomings to be noted. Specifically, (1) the first two steps of traditional CKC (information gathering and weaponization) are difficult to identify, (2) the scope of lateral movement is too broad, it cannot reflect detailed activities during movement. (3) they do not distinguish between attack events and alerts at different levels, and (4) they cannot represent the fundamental cyber-attack constructs.

In the practical application of cyber threats investigation, the attack analysis model needs to be combined with specific detection events. A further modification to facilitate APT analysis is required to incorporate practical usage in threat hunting.

To preserve key stages of Cyber Kill Chain and facilitate attack analysis. The author argue that the Cyber Kill Chain should be combined into 5 stages, Reconnaissance (R), Weaponization & Delivery (W&D), Exploitation & Installation (E&I), Command & Control (C2), Action (A).

On the other hand, it can be seen from the existing APTs analysis cases that the fragments of APT attacks are difficult to locate in an attack scenario. We need more information to combine attack events and attack scenarios. For this reason, it is necessary to make further improvement of Cyber Kill Chain models to organize attack fragments.

Table 5 Fields in *alert*

Field	description
a_id	alert id
timestamp	event timestamp
sip	source IP address
sport	source port number
dip	destination IP address
dport	destination port number
a_type	type of alert
src	Source of alert
desc	description of alert type
risk	potential risk level of alert
conf	confidence of alert
phase	the phase of alert

The main innovation of our Modified Cyber Kill Chain model is that the *Action* stage can be the fulfillment of attack goal or the subsequently conducted kill chain. The attack process is generalized as a recursive structure, and the structure is organized as a chain of the sub-chains. The source and destination of the elements in one model point to the same address in the network. For normal business scenarios without network attacks, there is very little evidence of dependence across the kill chain phases. Here we assume that there is only one APT attack against a node in a short period of time. For a victim host of APT attack, there is mutual independence between the ab-normal events, and the occurrence of kill chain events will bring multiple abnor-mal events. The recognition of these abnormal events is the premise of detecting APT attack. The structure of Modified Cyber Kill Chain model is shown in Fig. 9.

Reconnaissance: gather information and plan their method of attack. After identifying potential targets, the first thing to do is gather information about the targets. Commonly used method is gathering information such as TTPs from public sources or other approaches. In addition, attackers also scan for victims, scan for vulnerabilities/weaknesses. Probe internal network and vulnerabilities. Reconnaissance is the preparatory stage of an attack. Its main objective is to identify and locate the target. Perform continuous inspection of network traffic flows to detect and prevent port scans and host sweeps. The difference from the next stage is that there is blindness, usually tentative network behavior. And the new trend is that the attack behavior of attackers adds distributed features to resist detection. Specific attack methods include port scan, network vulnerability scan, web application vulnerability scan, database vulnerability scan.

Weaponization & Delivery: Customize attack tools according to former collected information and deliver them to the target network, gradually approaching the target host. One of the most common ways of doing this is spear phishing attacks with malicious links or attachments. This step is often the most critical part of the successful implementation of an attack. And because of people's participation, it's also the hardest to find out. In the delivery process, it is often the application of social engineering. if possible, even physical delivery method. The difference from the previous stage is that it is the stage of preparation for contact and has a certain understanding of the objectives. The difference from the next stage (Exploitation & Installation) is that destructive events do not really work at this stage, and detection and blocking at this stage can prevent actual losses on the host. Specific attack methods include social engineering、email spam (URL or attachments)、malicious or phishing websites、removable media.

Exploitation & Installation: This stage is the period in which the essence of the attack is carried out. Exploit kit or weaponized document. Usually includes the use of 0-day vulnerabilities, buffer overflow, and install malware in order to conduct further operations. The difference from the first two stages is the need to achieve direct penetration of goals. The difference from the follow-up is that the vulnerability exploitation is still in the initial stage and there is no interaction with the attacker. And the programs run in an automated way, without participation of human beings, Specific attack methods include Privilege escalation, RAT and backdoor software.

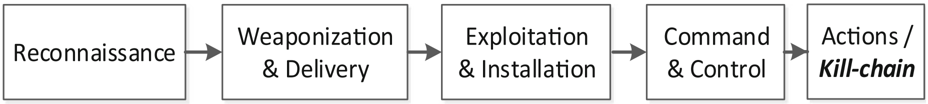

Fig. 9 Modified Cyber Kill Chain model

Command and control (C2): actively control the system, instructing the next stages of attack. The difference from the previous stage is that this stage relies on covert communication. Unlike the next stage, there is no operation on the target host currently. It only transfers command information and internal probe data.

Action on objective/ Kill Chain: data exfiltration, destruction of critical infrastructure, to deface web property, or to create fear or the means for extortion. In the Lockheed Installation phase, it is simply the installation of a backdoor to the compromised system. In our phase it can be as simple as assessing the information on the compromised system to beginning to reach out across the network to determine other systems to exploit or gathering information such as available services. The attacker must reach out and attempt to gather information about the network. The actual attack events are performed in this stage. Another case is that the attack is not completed and there is no actual destruction or theft, but the next CKC is the target of the front attack events.

As shown in Fig. 10, node A conduct kill chain attack after scanning the network, and establish network connection to node B. The kill chain 1 is then connected by the movement of B to C. Here we use Attack Scenario Graph (ASG) to represent the APT attack scenario. The relevant definitions in ASG are as follows.

Node Elements in node set N include PC, server, router, smartphone, PDA and other computing equipment with functions of program execution and network communication.

Edge Edge is a pair of nodes in a graph,$e = (n_s, n_t, KC(n_s, n_t)) \in E$, where n_s, $n_t \in N$. An edge represents 5 stages of a complete Cyber Kill Chain summarized above. Edges in ASG represent the intrusion relationship between nodes of the network.

Attack scenario graph Attack scenario graph G consists of nodes and edges. Attack graph and fault tree are two commonly used structures for representing attack scenarios. In this paper, we use graph structure composed of kill chains for attack scenarios representation.

This is an integrated, end-to-end process which is described as a "chain" because any one deficiency will interrupt the entire process.

Attack alerts are generated by anomaly detection algorithms and other security devices. Once an alert occurs between nodes, analysis engine will map it to the stage of kill chain and an edge will be activated. Continuously added into the network topology. The edge between

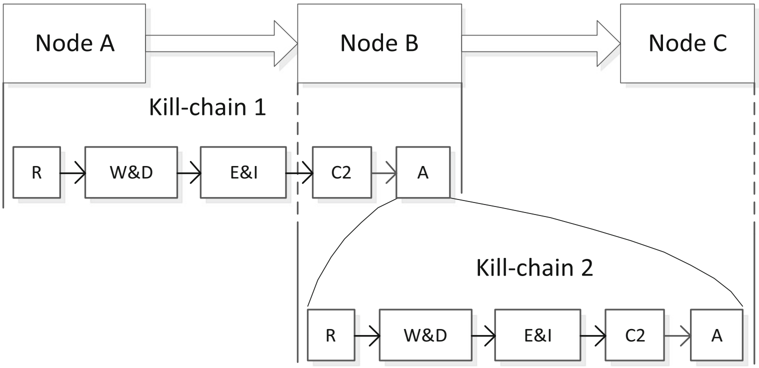

Fig. 10 Expanded version of Modified Cyber Kill Chain model—MCKC

nodes is represented by Kill Chain. The definitions of event, alert, phase, kill chain and attack scenario are as follows.

Event APT detection begins with events come from raw audit logs and security alerts. Security events are the traces of the APT attack. And with the rapid growth of data volume, security log processing has become a big data problem.

Alert Events correspond to the raw audit logs and alerts contain more security semantics. Security alerts mainly come from two sources, security alarms generated by security devices and anomalies generated from audit logs. As we know, these alerts may not be real attack and false alarms are the main problem that bothers us. The accuracy rate of alert generation will influence follow-up analysis. We need to discover and update feature bases. Generate alerts/alarms, and matching feature, AI technologies will simplify this process. Generated alerts are higher level security events compared with row events.

Phase Attack phase is the tag of alerts. This tagging process is implemented based on expert knowledge. The mapping function is mainly realized by CAPEC [5] The Common Attack Pattern Enumeration and Classification (CAPEC) provides a publicly available catalog of common attack patterns that helps users understand how adversaries exploit weaknesses. Different types of events found can be mapped to the corresponding phases in the kill chain. The attack phases express the further recognition of the attack. The main challenge is the design of map function and this is the main direction of future research.

Kill chain The Kill Chain expresses the local integrity of the attack process. It links the attack phases occurring on the same hosts together. The rules for linking the hosts are based on the same host or similar features. And then, we can conduct correlation analysis approach after the basic kill chain is formulated,

$$kill\text{-}chain(KC) = \{R,\ \ W\&D,\ \ E\&I,\ \ C2,\ \ A\}$$

Scenario An attack scenario is a more comprehensive description of the attack connected by KCs (kill chains) according to the relationship of attack events. In which $KC(\mathrm{m})$. $A = KC(\mathrm{m} + 1)$ where$1 \leq \mathrm{m} \leq \mathrm{n} - 1$, or there is parallel relationship between KCs. Based on the analysis of existing APT attack cases, we assume that n is no more than 3.

$$attack\text{-}scenario = \{KC(1), KC(2), \cdots, KC(\mathrm{n})\}$$

4 Cognitive analysis approach with MCKC

Cognitive analysis with MCKC is the key to correlate multi-source security data into attack scenarios. It mainly relies on the combination of forward analysis and backward analysis. At first, forward analysis partial kill chain through matching from raw log data and establish connections between sub-chains. Backward anal-ysis is a further supplement to forward analysis. It populates the undetected and unacknowledged events to improve the confidence of detection. Flowchart of cognitive analysis with MCKC model is as shown in Fig. 11.

4.1 Forward analysis

The purpose of forward analysis is to assess the impact of APTs, by starting from an entry point and discovering all the possible effects dependent on the entry point. Kill Chain represents the penetration process between nodes in the network. Some steps may be missing in detection. It is possible that an attack that does not occur in itself or that part of the attack has occurred but has not been detected. Attack scenario cascade Kill Chain between nodes into a new chain. The action of the intermediate node is actually another Cyber Kill Chain. And the process of forward analysis relies on classify algorithm.

Firstly, raw data is collected and gathered from network, system and other data sources. After preprocessing, raw data is represented as formatted events and stored on the platform. Different data sources correspond to different stages of an attack. It can be found from the previous research [2] that the volume of events is too large and the event relationships contained in the original log are not clear. Therefore, it is necessary to analyze the multi-source security events to produce anomaly alerts with security semantics. However, only low-level security information is included in anomaly alerts and there are false alarms that will mislead attack analysis. Other semantic security data generated by security devices such as IDS, antivirus software and firewall also provide information about attacks. Correlation analysis will connect the dots into a larger campaign. Guided by expert knowledge, security alerts can be mapped into attack phases. The design of mapping function is the most important part in security systems. On the basis of the previous work, attack scenarios can be reconstructed from kill chains (Fig. 12).

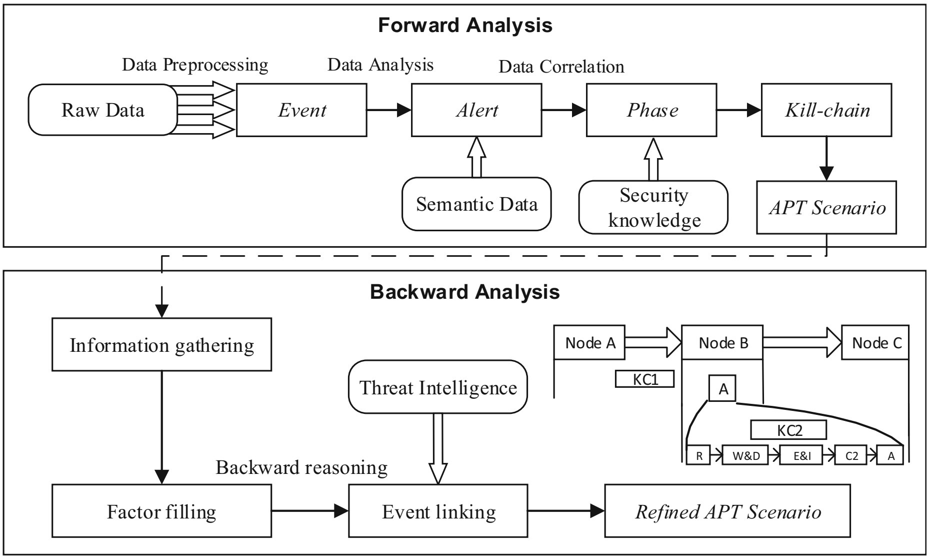

Fig. 11 Flowchart of cognitive analysis

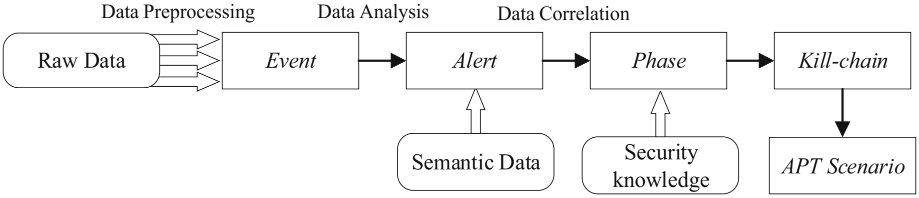

Fig. 12 The process of forward analysis

Forward analysis approach relies on a layered framework for analysis. The layered frame-work of forward analysis is shown in Fig. 13. The structure of layered frame-work corresponds to the process of forward analysis. From bottom to top are raw data layer, event layer, security alert layer, attack phase layer, and kill chain layer.

In the layered framework of forward analysis, the raw data layer provides in-itial data source for the whole system, and it is the basis of analysis and detection at upper layers. The original log data can be converted into formatted events through log parsing, and these formatted events are the input of anomaly detection algorithm. The output of the anomaly detection algorithm, together with the alerts generated by existing security devices, are the source of subsequent attack analysis. In the attack phase layer, different types of security alerts can be mapped to different stages in the kill chain. And the alerts are organized into kill chains through alert correlation.

Concretely, the process of forward analysis is described in Algorithm 1. Three kinds of data are the input of Algorithm 1: audit logs L, security alerts W and security knowledge K. Audit logs are expressed as security events through log parsing. Anomaly score of security event is computed according to anomaly detection function $\_function(e)$. And if the anomaly score is larger than threshold θ, the security event is generated as an alarm. This aggregation method greatly reduces the number of security incidents and relieves the pressure of follow-up analysis. By annotating event types of security alerts, various security events are mapped to

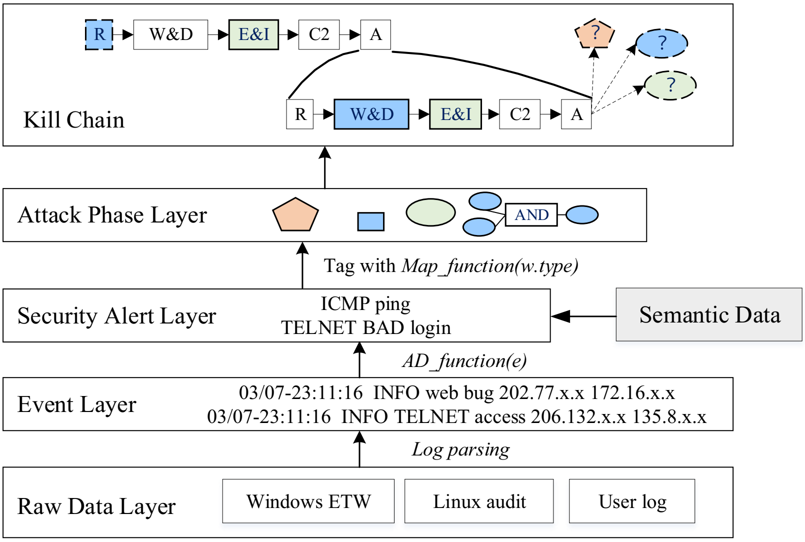

Fig. 13 Layered framework of forward analysis

kill chains. By association of kill chains and adding hypothetical events, kill chains are filled step by step. Finally, attack scenarios are concatenated by kill chains.

Algorithm 1. Forward analysis algorithm

```
Input: audit logs $L$, threshold $\theta$, security alerts $W$, security knowledge $K$
Output: attack scenario $S$
while $L$ is not empty
    for all log entry $l$ in $L$, do
        $e = resolve\_function(l)$
        add $e$ to event set $E$
    end for
    for all event $e$ in $E$, do
        $s(e) = AD\_function(e)$                // compute anomaly score of $e$
        if $s(e) > \theta$                      // $\theta$ is set according to security level
            then add $e$ to the $W$
        if $e.type$ is not in event type set
            then add $e.type$ to event type set    //according to security requirement
    end for
    for all alert $w$ in $W$, do
        $w.node = (e.src, e.dst)$, $w.edge = KC(e.src, e.dst)$
        activate $w.node$ and $w.edge$
        $w.tag = map\_function(w.type)$
    end for
    for $killchain$ in set $KC$
        if $1, e2 \in killchain$ , add $Hypo\_evt\ e'$ to $killchain$
        add $killchain$ to $S$
    end for
    for each $killchain_i, killchain_j$ in set $KC$
        if $killchain_i.A = killchain_j$
        add $S(killchain_i).attend(killchain_j)$
    end for
end while
return $S$
```

4.2 Backward reasoning

Because of the hidden behavior of attackers, it is often difficult to detect certain attacks in time. Generally, the APT detection is a forward analysis process, but in afterwards attack analysis, backward reasoning is more often used. After anomalous events were detected by forward analysis, the complete attack scenario may not be recovered. By combining backward reasoning, missing attack elements can be added and hypothetical knowledge can supplement the cognition of APT attacks.

Here we refer to the method of diamond model about pivoting analysis. The Diamond Model fundamentally supports analytic pivoting and is one of its strongest characteristics.

Pivoting analysis is the analytic technique of extracting a data element and exploiting that element, in conjunction with data sources, to discover other related elements.

The results of forward analysis may be incomplete because of two reasons. Undetected attacks may occur due to insufficient data sources or missing events. And intentional concealment of personal behavior by attackers also leads to the omission of attack detection events. The implementation of a single attack process does not necessarily guarantee the expected attack target. It usually requires a complete APT attack with multiple hosts as foothold. Attack scenarios need to be restored and described from the perspective of attack implementation.

Forward analysis can organize alerts into an kill chain, and forward analysis also makes it easier to connect sporadic attacks. Backward reasoning with MCKC is another view of attack analysis. It can supplement undetected missing attack elements. According to the continuity of APT attack implementation, missing elements can be added to this incomplete kill chain. By organizing attack events into kill chains, it is easier to connect scattered attack events. Backward reasoning can further enrich the kill chain and enhance the detected attack scenario.

Take a typical APT attack scenario as an example, the process of backward reasoning is shown in Fig. 14. As shown in Fig 14, two events in attack scenario were not detected after the forward detection, R in *killchain*$_1$ and C2 in *killchain*$_2$. According to the implementation rules of APT attacks, we add hypothetical events $e_1 = \text{R}$ and $e_2 = \text{C2}$ to the detected kill chains, Further, alerts that may not have been detected need to be traced back from alert layers. Further validation analysis of events corresponding to event level will be performed by analysts.

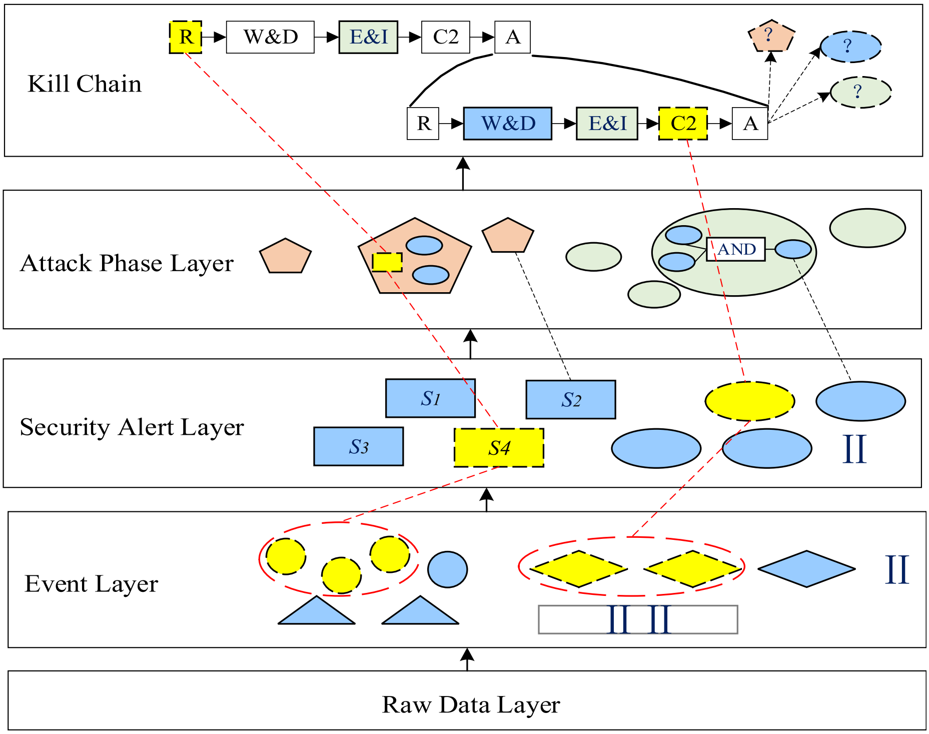

Fig. 14 Layered framework of backward reasoning

5 Case study

In practical usage, forward analysis and backward reasoning were combined to generate APT attack scenario. In this section, we give a case study on typical APT attack scenario reconstruction to illustrate the effectiveness of our approach. Furthermore, an example of WannaCry attack is presented to prove the effectiveness of the MCKC model in an actual APT attack traceback. From the view of both forward analysis and backward reasoning the of our proposed MCKC model can be verified to be effective for APT attack detection and investigation.

5.1 Typical APT attack scenario

The APT attack scenario described in the Figure 15 is as follows.

- **S1**: reconnaissance

An attacker locates the target and sends a phishing email to an internal user. The email contains an attachment with malicious code.

- **S2**: delivery

The victim user download and open the attachment from mail client. Malicious code hidden in the attachment is executed.

- **S3**: exploitation

Backdoor malware installation and execution.

- **S4**: c2

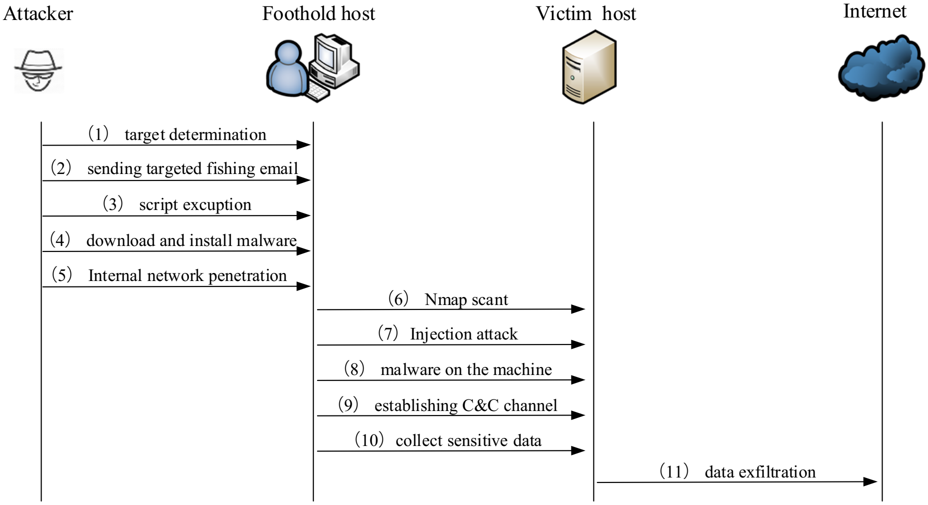
Attacker
Foothold host
Victim host
Internet
(1) target determination
(2) sending targeted fishing email
(3) script excuption
(4) download and install malware
(5) Internal network penetration
(6) Nmap scant
(7) Injection attack
(8) malware on the machine
(9) establishing C&C channel
(10) collect sensitive data
(11) data exfiltration

Fig. 15 A typical APT attack scenario

Obtaining privileges in the target network. The attacker establishes C&C channel with foothold node.

- **S5**: action

Sub Kill Chain conducted by node B (s6-s10).

- **S6**: reconnaissance

The attacker performs a port scan on the foothold host using *Nmap*.

- **S7**: delivery

The attacker uses the attack script to exploit the victim machine (using "*ExtraBacon*").

- **S8**: exploitation

On successful exploitation, the attacker injects malware into the victim machine.

- **S9**: c2

Establishing C&C channel between victim host and external attack host.

- **S10**: action

Data exfiltration. Encryption software and keys get downloaded into the victim machine. The next task from the malware is the detection of sensitive files and their encryption using the downloaded tool.

Forward analysis detection result is shown as in Fig 16. In the forward analysis process, multiple detection events (**S2, S3, S4**) are identified by IDS and other security devices. The Kill Chain 1 implemented by node A to node B is activated. Subsequently, attack events on Node C are detected (**S6-S8, S10**), the Kill Chain 2 implemented by node B to node C is activated. These two kill chain are connected to form a more complete APT attack scenario, as the target host of Kill Chain 1 is the initiator of Kill Chain 2. Then, two attack hypothetical events (**S2, S9**) are added to the attack scenario and the ASG construction process is shown as follows. Based on the scoring results, backward reasoning identifies the set of possible source events for hypothetical events and submits them to the analyst for further verification.

5.2 WannaCry attack

Furthermore, to prove the effectiveness of MCKC model in the actual attack analysis, the trace analysis of WannaCry attack is taken as an example to illustrate the process of kill chain expression and APT detection. The APT attack scenario described in Figure 17 is as follows.

- **S1**: reconnaissance

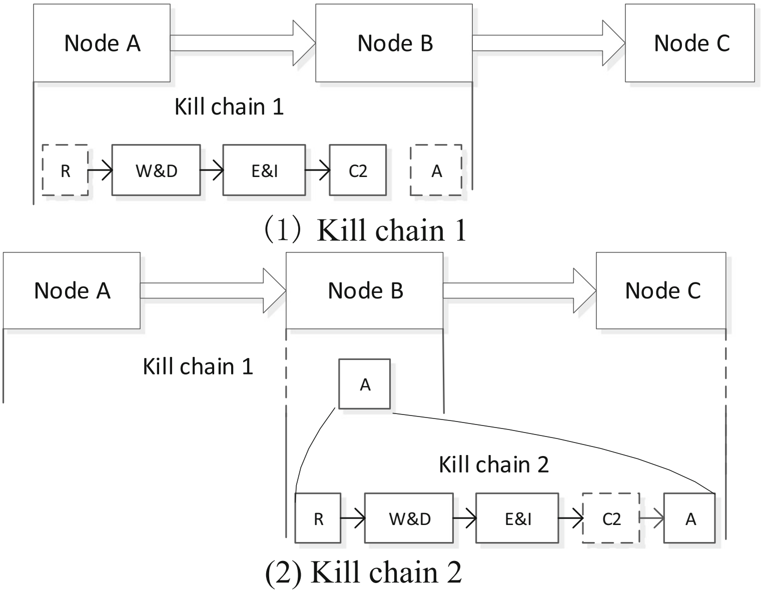

Fig. 16 Forward analysis detection result

An attacker spread phishing email with malware over the Internet.

- **S2**: delivery

Malware mssecsvc.exe infect the host through the vulnerability of eternal blue MS-017.

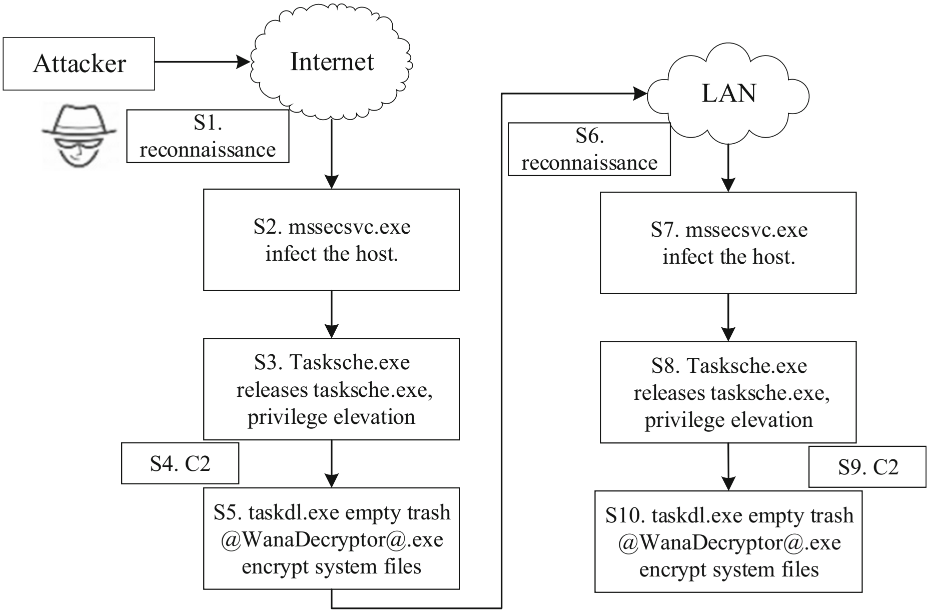
Attacker
Internet
S1. reconnaissance
S2. mssecsvc.exe infect the host.
S3. Tasksche.exe releases tasksche.exe, privilege elevation
S4. C2
S5. taskdl.exe empty trash @WanaDecryptor@.exe encrypt system files
LAN
S6. reconnaissance
S7. mssecsvc.exe infect the host.
S8. Tasksche.exe releases tasksche.exe, privilege elevation
S9. C2
S10. taskdl.exe empty trash @WanaDecryptor@.exe encrypt system files

Fig. 17 WannaCry attack

- **S3**: exploitation

The main program mssecsvc.exe releases tasksche.exe, and tasksche.exe releases Taskse.exe to achieve an elevation of privilege

- **S4**: c2

The main program mssecsvc.exe establishes C&C connection.

- **S5**: action

Ransomware @WanaDecryptor@.exe encrypt system files and begin extortion. At the same time, Taskdl.exe empty trash and further spread malware in internal LAN. This stage also includes the Sub Kill Chain conducted by node B (s6-s10).

- **S6**: reconnaissance

The main program tasksche.exe spread malware in internal LAN

- **S7**: delivery

Malware mssecsvc.exe infects the host through the vulnerability of eternal blue MS-017.

- **S8**: exploitation

The main program mssecsvc.exe releases tasksche.exe, and tasksche.exe releases Taskse.exe to achieve an elevation of privilege

- **S9**: c2

The main program mssecsvc.exe establishes C&C connection.

- **S10**: action

Ransomware @WanaDecryptor@.exe encrypt system files and begin extortion. At the same time, Taskdl.exe empty trash and eliminate attack traces.

In the actual scenario, we often use the method of backward analysis to re-store the attack scene gradually. As for the detection of WannaCry, we start with malicious code analysis. By judging the parameters of the main program and the height of the code in memory, the corresponding process is executed, and its propagation mode is analyzed. After analyzing the kill chain of WannaCry, we can prevent further damage to other hosts in the network by patching the vulnerability, closing the port and create a mutex. This is also the idea of blocking the kill chain in time.

As the case study shows, the correlation of multiple actual events can be realized by adding hypothetical edges between discovered events. In this way, we can improve the detection rate and avoid difficulties due to a lack of detection events.

5.3 Result and analysis

Weakest Link problem will bring obstacles to the attacker. That means, an intruder must generate events across the kill chain to accomplish his mission. However, the independence suggests that corresponding chains of events may be infrequent under non-attack conditions. The correlation of multiple events on the chain can help analysts reconstruct attack scenarios, and correlation analysis based on multi-source events is also the main idea of our paper.

From the above analysis, we can see that the analysis based on kill chain has unique advantages on describing the attack scenario and helping to understand. Moreover, with the description of kill chain, we can see the key steps in the process of attack implementation, and quickly block them in defense.

To represent a dynamic security system with time characteristics, the following parameters are given in the model,

- Penetration time Pt: time from the beginning of the invasion to the success.
- Detection time Dt: time from the beginning of the attack to the detection of it.
- Response time Rt: time from the attack is detected to the system begin to respond.
- Exposure time Et: it is the time when the system is in an unsafe condition.

Due to the characteristics of big data analysis methods, correlation analysis and detection of APTs often lag behind the occurrence of real attacks. From the perspective of dynamic security model, the time of attack discovery may be longer than that of attack success. That is $Et = Dt + Rt - Pt > 0$. Only if $Pt > Dt + Rt$ can the security of system be guaranteed. Through the application of this model, the time of attack discovery Dt can be shortened, so as to ensure the security of the system.

The timeline of attack detection and prevention is shown in Fig 18.

Furthermore, in order to illustrate the effectiveness and innovation of the model proposed in this paper, we make a comparative analysis between the MCKC model and previous studies. Here we will first explain the evaluation indexes (C1-C10):

C1 How many stages the model is divided. The number of model phases.
C2 Whether retrieval analysis for alarm events are supported.
C3 Whether lateral movement of internal network is considered.
C4 Whether quantitative analysis is supported.
C5 Whether it is close to human cognitive intelligence

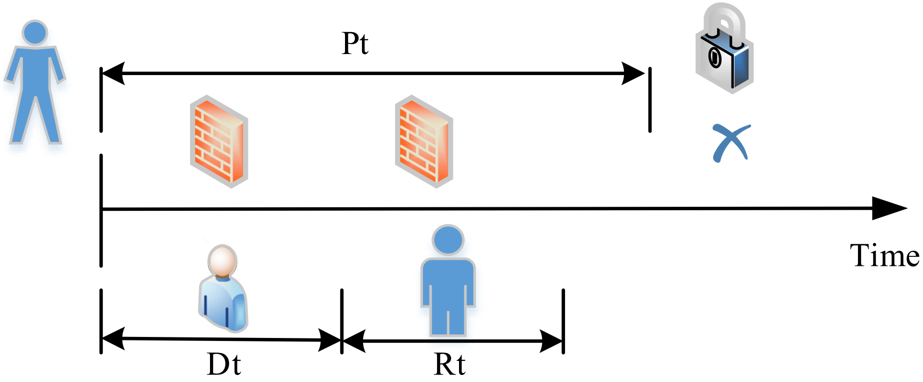

Fig. 18 The timeline of attack detection and prevention

C6 Whether it is lightweight enough for cognitive analysis.
C7 Whether loop iteration analysis is supported
C8 Whether it is convenient for analysists to map data layer by layer
C9 Whether meta-analysis is supported.
C10 Three levels of alert event compression: H(high), M(mid), L(low)

Table 6**.** Performance and efficiency comparison among relevant attack models.

From the above analysis, we can see that MCKC model is superior to the previous methods in most indicators. Our model simplifies the attack process to five stages, which is a lightweight model. MCKC model retains the expression of the lateral movement of internal network in a recursive structure. Our bidirectional analysis method supports meta-analysis and loop iteration analysis, which is closer to human's cognition.

However, MCKC model is insufficient in quantitative analysis, and this is also an important problem in the field of APT attack investigation. Attacks need to be novel enough to avoid detection, and business anomaly and probability anomaly are often inconsistent, which makes it a difficult problem to solve. The measurement of attack and risk assessment with ATT&CK model is one of our future research directions.

The main contributions of this dissertation are concluded as follows:

1. A lightweight kill chain model is proposed. To compress irrelevant stages, we revise the Cyber Kill Chain model and reduce the kill chain into 5 detectable stages. We aggregate the similar stages in attack detection, combing Weaponization and Delivery, Exploitation and Installation into one part. Compared with the existing models, this kind of combination is more suitable for cognitive analysis of APT attack.

2. The attack can be organized as a recursive structure as the Action stage of kill chain model can be a sub-chain of the whole attack scenario. Attack scenario can be represented as a chain of the sub-chains. Compared with existing models, this kind of representation is dynamically adjustable as there are no limits on the number of attack phases.

3. A bidirectional analysis method is proposed, which can be used to analyze APT attacks in both forward and backward directions. The method proposed in this paper explicitly maps attack events to kill chain phases and organizes them into attack scenario. At the same time, the missing attack events can be supplemented by backward reasoning.

Table 6 Performance and efficiency comparison among relevant attack models

Model	Year	Evaluation Criteria									
		C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
Lockheed Martin Kill Chain Model	2011	7	✓	×	×	×	×	×	×	×	L
Mandiant Attack Lifecycle Model	2013	8	✓	✓	×	✓	×	×	×	×	L
Diamond Model	2013	–	✓	–	–	✓	×	✓	×	✓	M
MITRE ATT&CK Model	2013	11	✓	×	×	×	×	×	×	×	M
Malone Kill Chain Model	2016	5/7	✓	✓	×	✓	×	✓	✓	×	M
The Unified Kill Chain Model	2017	18	✓	✓	✓	×	×	×	✓	×	M
Bryant Kill Chain Model	2017	4	✓	×	×	✓	✓	×	×	×	H
Khan Kill Chain Model	2018	8	✓	✓	×	✓	×	✓	×	✓	M
MCKC Model in our paper	2020	5	✓	✓	×	✓	✓	✓	✓	✓	H

4. Close to human cognitive intelligence. The forward analysis method from bottom up and the backward analysis from top down are given. It facilitates logical data correlation in attack detection. It’s a user-friendly threat analysis method.

To summarize, our model is simpler and easier to understand. This Modified Kill Chain Model can be applied in practical analysis, as we mentioned in the supplementary case study. The idea of this paper is different from that of predecessors in that it is to segment attacks and connect fragments into a whole scene. In this way, we can avoid one-sidedness in correlation analysis. The measurement of attack and risk assessment with ATT&CK model is one of our future research directions.

6 Conclusion and future work

Attack modeling and analysis techniques are important for APT attack discovery and provenance tracing. Especially kill chain model for describing APT attack penetration process. Much research has been done to make it better for representing attack activity. Firstly, this paper summarizes the previous research techniques. To overcome the deficiencies of the standard Cyber Kill Chain model, a Modified Cyber Kill Chain model (MCKC) was developed through literature study and case studies. Different types of threats are also categorized with each stage of Cyber Kill Chain. Finally, a case study for bi-directional analysis method based on MCKC model is presented to describe the effectiveness of the proposed MCKC model. The result of case study shows that, this revised kill chain model can express APT attack process flexibly, and it is more convenient to understand reasoning and detection.

Modelling cyber attacks and predicting threat activities is an important issue for securing enterprise multimedia network. Based on previous research of kill chain, the MCKC model presented in this paper extends the process modeling method for network threats. Compared with existing models, the MCKC model represents both internal kill chain and lateral movement. The attack events are concatenated in series by kill chain, and redundant alarms can be decreased through data aggregation. More importantly, the proposed MCKC model is flexible attack expression and it supports bi-directional analysis for both forward analysis and backward reasoning. Forward analysis provides a layered framework for entire-process attack detection and kill chain filling. Backward reasoning further complements and improve the results of forward detection, and it helps security experts to trace back the APT attacks with a cognitive analysis approach.

In future research, the MCKC model could be further refined through adding attack assessment measures. Additional case studies and analytical tactics could be made to optimize the process of APTs provenance tracing. It is also necessary to combine knowledge computing technology to strengthen automated reasoning. More measures should be made to formulate and adjust corresponding defensive strategies against APTs.

References

1. Amirul Aslam Ahmed, Anazida Zainal (2017): Cyber Attack Profiling Towards Critical Inftastructures Using Modified System Fault Risk Framework. UTM Computing Proceedings Innovations in Computing Technology and Applications 2
2. Albeshri, A., Thayananthan, V.J.I.J.O.I.T, Making, D. (2018): Analytical techniques for decision making on information security for big data breaches

3. Bada, M, Sasse, AM, Nurse, JRJ (2019). A.P.A.: Cyber security awareness campaigns: Why do they fail to change behaviour
4. Kiran Bandla, David Westcott, https://github.com/aptnotes/data
5. Bayley, I (2014): Challenges for a formal framework for patterns. https://link.springer.com/chapter/10.1007%2F978-3-319-04447-7_4
6. Bryant B, Saiedian H (2017) A novel kill-chain framework for remote security log analysis with SIEM software. Computers & Security 67:198–210
7. Caltagirone, S, Pendergast, A, Betz, C (2013): The diamond model of intrusion analysis. Center For Cyber Intelligence Analysis and Threat Research
8. FireEye, https://www.fireeye.com/blog/threat-research/2018/04/m-trends-2018.html
9. Hutchins EM, Cloppert MJ, Amin RM (2011) Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains. Leading Issues in Information Warfare & Security Research 1:1–14
10. Khan, MS, Siddiqui, S, Ferens, K (2018): A Cognitive and Concurrent Cyber Kill Chain Model. Computer and Network Security Essentials, pp. 585–602
11. Lallie, H.S., Debattista, K., Bal, J.: An empirical evaluation of the effectiveness of attack graphs and fault trees in cyber-attack perception. IEEE Transactions on Information Forensics and Security \, 1110–1122 (2018)
12. Mandiant APT (2013) 1 Report
13. Siadati, H, Memon, N (2017): Detecting structurally anomalous logins within enterprise networks. the 2017 ACM SIGSAC Conference, pp. 1273–1284
14. Strom, BE, Battaglia, JA, Kemmerer, MS, Kupersanin, W, Miller, DP, Wampler, C, Whitley, SM, Wolf, RD (2017): Finding Cyber Threats with ATT&CK™-Based Analytics. MITRE Technical Report MTR170202. The MITRE Corporation, 2017. URL https …
15. Syed, Z, Padia, A, Finin, T, Mathews, ML, Joshi, A (2017): UCO: a unified Cybersecurity ontology. In: AAAI Workshop: Artificial Intelligence for Cyber Security
16. Wikipedia, https://en.wikipedia.org/wiki/Advanced_persistent_threat