Modeling the Operational Phases of APT Campaigns

Aimad Berady, Valérie Viet Triem Tong, Gilles Guette, Christophe Bidan, Guillaume Carat

Modeling the Operational Phases of APT Campaigns

Aimad BERADY*, Valérie VIET TRIEM TONG*, Gilles GUETTE*, Christophe BIDAN*, Guillaume CARAT
*CIDRE Team
CentraleSupélec, Univ Rennes, Inria, CNRS, IRISA
Rennes, France
firstname.lastname@irisa.fr

***Abstract*—In the context of Advanced Persistent Threat (APT) attacks, this paper introduces a model, called Nuke, which tries to provide a more operational reading of the attackers' lifecycle in a compromised network. It allows to consider the notions of regression; and repetitiveness of final objectives achievement. By confronting this model with examples of recent attacks (Equifax data breach and TV5Monde sabotage), we emphasize the importance of the attack chronology in the Cyber Threat Intelligence (CTI) reports, as well as the Tactics, Techniques and Procedures (TTP) used by the attacker during his progression.**

***Index Terms*—advanced persistent threat, cyber kill chain, tactics techniques and procedures, cyberspace operations, cyber threat intelligence**

I. Introduction

In the field of computer security, we have been witnessing for years the awareness of the existence of a so-called Advanced Persistent Threat (APT). These attacks, regularly targeting or involving nation-states and large companies, were defined by NIST [1] in 2011. The Advanced Persistent Threat: (i) pursues its objectives repeatedly over an extended period of time; (ii) adapts to defenders' efforts to resist it; and (iii) is determined to maintain the level of interaction needed to execute its objectives.

Since then, in order to model the progression of these attacks, academics and professionals have helped to define the lifecycle of an attacker during his operations and used it to apply mitigations. Different visions oppose to each other or complement one another. They are called kill chains. Unfortunately, these models do not allow us to represent the notion of repetitiveness, which is, however, explicit in the NIST definition. They also do not consider that the attacker may encounter disappointments or be surprised in victim's information system and that he must regularly re-examine his interest in pursuing the attack and assume the costs.

In this paper, after having established the state of the art (section II), we propose a model (section III) that represents the lifecycle of an attacker in a compromised network and that considers a possible regression and introduces the concept of a waiting state, which is essential for long-term actions. Then we propose a confrontation (section IV) between this model and two recent examples of attacks whose progression has been publicly described: the Equifax breach (2017) and the TV5Monde sabotage (2015).

II. State of the art

A. Linear and circular models

With the aim of helping defenders better analyze and respond to the cyber attacks they face, Lockheed Martin described for the first time in 2011 [2] the now famous Cyber Kill Chain® (Figure 1) and its different phases. This concept, borrowed from the military doctrine, lays the foundations for a tactical phasing of cyber offensive operations and proposes countermeasures for each of them in order to stop the attack. However, this model, because of its linearity and its focus on the initial foothold, is not so representative regarding the progression of an advanced attacker inside a network towards his objective, especially the operational states in which he evolves and that are the reflection of his intention.

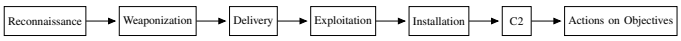

Fig. 1. Phases of Lockheed Martin's Cyber Kill Chain®

It is interesting to note that the same year, in a contribution proposing a global definition of APT attacks, Command Five Pty introduced [3] into their model (Figure 2) the notion of *maintenance*, which translates a need for the attacker to persist in the long term while being disturbed by unforeseen events. They also suggest a possible return to an earlier step *in an attempt to regain access* to his objective.

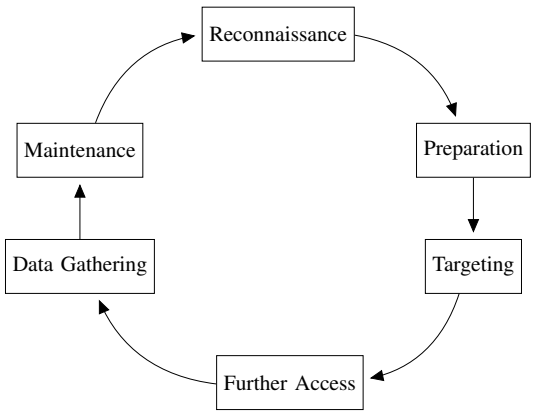

Fig. 2. Phases of Command Five's model

It is the combination of linear and circular shapes for models that has naturally given way to loop ones.

B. Loop models

Since 2013, Mandiant (now FireEye) has used a visualization of the attackers' lifecycles they study through a loop model. It helps to restore a notion that has been left aside in Lockheed Martin's founding publication: repetitiveness. Subsequently, in 2017, Pols (Fox-IT) presents in his PhD thesis [4], a Unified Kill Chain that he built by aggregating the previously proposed models and case studies (Fox-IT's Red Team and APT28). His kill chain consists of 17 optional steps in three loops (Figure 3), which clearly correspond to operational phases.

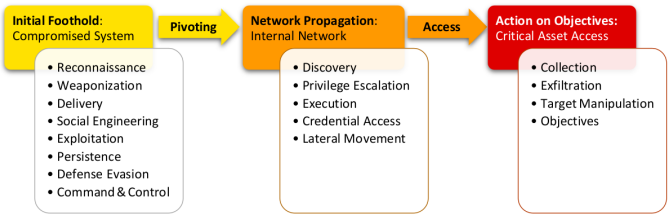

Fig. 3. Pols' Unified Kill Chain

We believe that these loop models are better suited to describe the essential notion of APT campaigns: persistence. But we are convinced that they have two main flaws. First, they do not consider that an attacker can regress during his progression due to actions or countermeasures performed by the defender or by the victim. Secondly, they do not let appear attacker's idle periods yet sometimes described in Cyber Threat Intelligence (CTI) reports, as can be seen in the Figure 4 from the RUAG espionage case [5]. These periods and their context, however, are essential in understanding the attacker's process and capabilities.

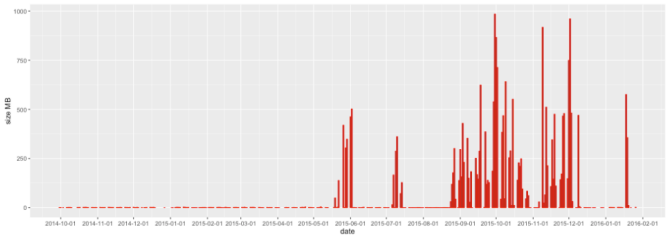

Fig. 4. RUAG espionage case: Data Exfiltration by Day

That is why we modeled a lifecycle in the form of a state machine. We present it in the next section.

III. Proposed Model

This article presents a new modeling of an attacker's lifecycle in a compromised network. Our model is formalized by a state machine with a total of six states. These six states can be divided into three superstates representing three operational phases of an attack. We assume that the attacker has already established his initial foothold, regardless of the involved techniques. Once he has successfully compromised a first asset of the targeted network, the *Exploration phase* begins. During this phase the attacker discovers the network until he succeeds in reaching the asset hosting his objective.

Once the attacker has tamed this asset, the *Exploitation phase* starts. In this phase he regularly checks if he is still able to perform his *ultimate techniques* and provides effects on his objective. What we call here the *ultimate techniques* are those he uses to produce effects on his objective.

These two superstates are tactical phases gathering the states described in the previous models (section II). We add here a last superstate: a *Decision-making phase* during which the attacker evaluates the possibility of aborting the attack.

In our model, the attacker is characterized through his operational capabilities which are techniques similar to those listed in MITRE ATT&CK matrix [6] and represents a part of its Tactics, Techniques and Procedures (TTP) [7]. In the following section, we formalize the fundamental elements involved in our model before presenting the model itself.

A. Model's elements

a) Attacker's final objective and targeted network assets: The attacker's ultimate objective may be to exfiltrate data or modify behavior on the target's critical assets. The assets can be computers (workstations or servers), software, data, users, business processes, etc. Thereafter, we will simply note $\mathbf{e}$ an asset and $\mathbf{E}$ all the different assets and we will note by $\mathbf{e_{final}} \in \mathbf{E}$ the asset hosting the attacker's final objective. $\mathbf{e_{init}} \in \mathbf{E}$ will note the intruder's initial implantation point inside the compromised network. It is not excluded that $\mathbf{e_{init}} = \mathbf{e_{final}}$.

b) Attacker's knowledge on the targeted network: The knowledge that the attacker has about his victim's network influences his behavior during the attack. We will use $\mathbf{K}_{Atk_n}$ to note the attacker's knowledge about the victim's environment at time n. This knowledge consists of a subset of the target network assets. We assume here that the attacker's initial knowledge is restricted to the initial implantation point. We also assume that his knowledge can only increase. In other words, it means that $\mathbf{K}_{Atk_0} = \{\mathbf{e_{init}}\} \neq \emptyset$ and $\mathbf{K}_{Atk_n} \subseteq \mathbf{K}_{Atk_{n+1}} \subseteq \mathbf{E}$.

c) Attacker's technical capabilities and their effects on the target: The attacker's technical capabilities are all the techniques an attacker masters and can use against a system. A lot of these techniques have a reference number proposed by MITRE. For example, $(\mathbf{t}_{1075})$ is *Pass the Hash (PtH)*, which is *a method of authenticating as a user without having access to the user's cleartext password*. From a more formal point of view, we note here $\mathbf{T}_{Atk}$ the set of all the techniques $\mathbf{t}_j$ mastered by the attacker and thus describing its operational capabilities. The application of a $\mathbf{t}$ technique by an attacker with the knowledge $\mathbf{K}_{Atk_n}$ on a system itself represented by its assets $\mathbf{E}$ is called an *action* and it induces:

- a set of artifacts $\mathbf{X}$ that are unwittingly disseminated by the attacker in the network and inside its assets;
- a set of newly discovered assets $\mathbf{e_i} \in \mathbf{E}$ enriching the attacker's knowledge of the surrounding environment.

These sets may be empty if the applied technique fails or if the attacker covers his tracks (*e.g.* by cleaning the logs). We will note that empty logs may be a characteristic artifact of an attempt to evade defense. The artifacts are the only elements that are observable by the defender. The comprehension of the attack is an extrapolation of these.

B. Nuke

With Nuke we intend to model the attacker's lifecycle in a compromised network as a state machine where each state is defined by the assets of the targeted network $\mathbf{E}$, the implantation point $\mathbf{e_{init}}$, the final objective $\mathbf{e_{final}}$, the set of techniques $\mathbf{T}_{Atk}$ mastered by the attacker, the current knowledge of the attacker $\mathbf{K}_{Atk_n}$, and the set of the artifacts $\mathbf{X}$ unwittingly disseminated by the attacker and the application of a technique induces a transition between two states.

The state machine is composed of six states. Nuke fixes the sequence of these as detailed in this section and as represented on Figure 5. These steps are part of three phases:

- *Exploration phase*,
- *Exploitation phase*,
- *Decision-making phase*.

In the following we describe each state and the events inducing a transition towards another state. The transition table is given on Figure 6.

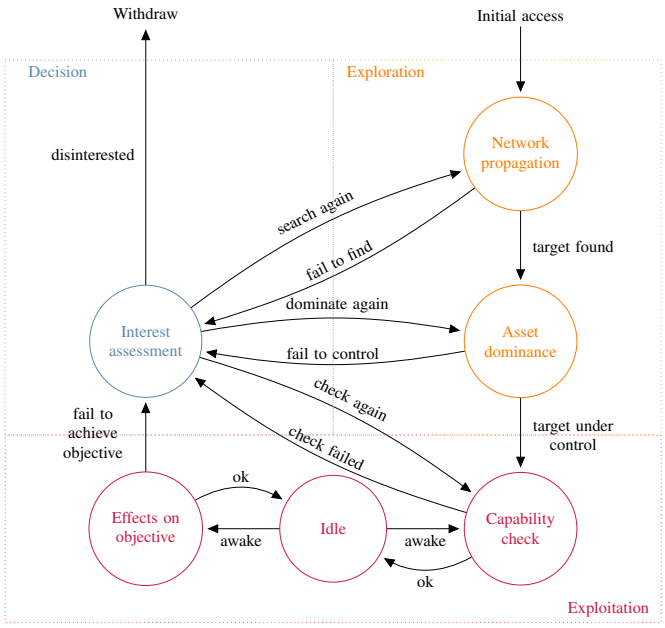

Fig. 5. Representation of the proposed model: Nuke

1) Exploration phase: This phase has two states namely *Network Propagation* and *Asset Dominance* where the attacker performs techniques on the compromised network until he dominates the asset hosting his final objective.

- **Network Propagation**: From his initial implantation point $\mathbf{e_{init}}$, the attacker starts to discover what is around him. But also, thanks to lateral movements, he becomes aware of the technical and social assets that make up the victim's network. This process from cognitive sciences is called *Situational Assessment*. It leads him to a level of knowledge called *Situational Awareness* [8]. In this state, the attacker is very noisy. His dissonant and hazardous technical actions denote from legit events. He disseminates artifacts all over the network. While attacker's final objective has not been discovered: $\mathbf{e_{final}} \notin \mathbf{K}_{Atk_n}$, he conducts *actions* from all techniques $\mathbf{t}_j$ from $\mathbf{T}_{Atk}$ on all the network's assets $\mathbf{e}$ of $\mathbf{K}_{Atk_n}$. During this step, the attacker has no particular focus. He looks for the asset hosting his objective in the entire network.
- **Asset Dominance**: when the attacker's final objective has been discovered: $\mathbf{e_{final}} \in \mathbf{K}_{Atk_n}$. He tries to tame the asset hosting his final objective to be able to perform his *ultimate techniques* efficiently. To do this, he must choose among his capabilities the most appropriate techniques or learn new ones. While the attacker's *ultimate techniques* are not a part of his capabilities, he conducts *actions* by applying techniques he masters $\mathbf{t}_j$ from $\mathbf{T}_{Atk}$, focusing on his objective $\mathbf{e_{final}}$. Before leaving this state, the attacker defines the best route from his initial implantation point $\mathbf{e_{init}}$ to the asset hosting his final objective $\mathbf{e_{final}}$.

During this phase, the attacker's knowledge ($\mathbf{K}_{Atk_n}$) about the targeted network increases.

2) Exploitation phase: Once the attacker has succeeded in dominating the targeted asset, he quits the *Exploration phase* to enter *Exploitation phase* during which he regularly checks if he is still able to perform effects on his final objective. We mainly distinguish the state where the attacker checks his capabilities on the network (*Capability check*), the state where the attacker stays dormant (*Idle*) and the state where the attacker produces effects on his final objective. Whatever the *Exploitation state* in which they are performed, technical actions are optimal. It means that they are homogenous, seem legit and the footprint in the victim's network is low.

- **Capability check**: when the asset hosting his final objective is under control of the attacker, he will check if he is able to perform his *ultimate techniques* on the asset hosting his objective. If one of these techniques fails, he moves to the *Interest assessment state* of the *Decision-making phase* described below. If these techniques succeeds, he moves to an *Idle state*.
- **Idle**: in this state, the attacker is just waiting. Regularly, he awakes and moves back to the *Capability check* or he moves towards the state *Effects on objective*.
- **Effects on objective**: using the defined route from his initial implantation point $\mathbf{e_{init}}$ to the asset hosting his final objective $\mathbf{e_{final}}$, the attacker performs primitive actions with the most appropriate techniques in order to achieve his objective. As for the *Capability check state*, if this technique fails, he moves to the *Interest assessment state* of the decision-making phase. If this technique succeeds, he moves back to the *Idle state*.

3) Decision-making phase: The decision-making phase has only one state: *Interest assessment*. In this phase, the attacker does not produce technical actions on the network. However, he will mainly confront two parameters: his level of remaining resources available (i.e. time, money, people, tooling), which represents the actual cost of his operation; and his interest in continuing the attack (motivation).

- **Interest assessment**: following an unexpected event encountered by the attacker while being in any state, he could end up in a non-technical state where he will put in perspective the estimated gain compared to the real cost of the attack. This event could be the result of a countermeasure implemented by the defender or because of an unavoidable technology for which the cost of developing skills is too high. If he considers that it is still interesting to reconquer lost territories, he moves back to one state of the *Exploration phase*; otherwise he withdraws. In this model, we assume that as long as the attacker is able to produce effects on the goal, he maintain presence in the compromised network.

Event / State	e_{final} found	e_{final} owned	Failure	Awakening	Positive Review
Network Propagation	Asset dominance		Interest assessment		
Asset dominance		Capability check	Interest assessment		
Capability check			Interest assessment		Idle
Idle				Capability check Effect on objective	
Effect on objective			Interest assessment		Idle
Interest assessment					Network Propagation Asset dominance Capability check

Fig. 6. Nuke transition table

IV. Comparing the model with two concrete APT

In this section, we confront our model to two APT campaigns that were reported publicly. We choose to work on these attacks because their public reports capture the timeline of the attack unlike the vast majority of reports that focus on the technical means used by the attackers.

These two chosen attacks are the data breach suffered by Equifax (a US credit bureau company) and the sabotage inflicted on TV5Monde (a French public television).

Despite these approaches describing a large number of the technical events of the attack, the reports are incomplete because they do not consider the issues that the attackers may encounter and evoke only the malicious events suffered by the victim.

Through these examples, we find that most of the information needed to complete the exploration and exploitation phases is present. But we also presume the presence of intermediate steps bringing the attacker back into a *Decision-making phase* and these can occur during any phase of the attack.

A. Equifax data breach (2017)

1) Context and initial foothold: Equifax is a US credit bureau company, stock traded, employing 11K people worldwide. It was founded in 1899 and had 3.1 billion USD in annual revenue in 2017. That same year, the IT security of the company was compromised [9]. A lot of the personal data of its customers was stolen during this attack.

Figure 7 details the representation of this breach according to our model. In this figure, the transitions between states that are explicitly described in the public report are noted in solid lines, on the contrary the transition that we can only presume are noted by dashed lines.

On May 13, 2017, the attacker exploited a known vulnerability on Apache Struts, which were installed on a public-facing application 1 ($\mathbf{t}_{1190}$) of Equifax's system called ACIS (*Automated Consumer Interview System*). It allowed customers to change incorrect information on their records.

The attacker established its first foothold in the ACIS environment. He installed a webshell ($\mathbf{t}_{1100}$) on this compromised web server. The only asset he knew inside the network is this server. $\mathbf{K}_{Atk_0} = \{\mathbf{e_{init}}\} = ACIS\_WS1$. Its final objective was to collect personally identifiable information ($\mathbf{e_{final}} = PII\_data$). At that moment, the attacker did not know the nature of this objective in the victim's network (files, databases, network flows...).

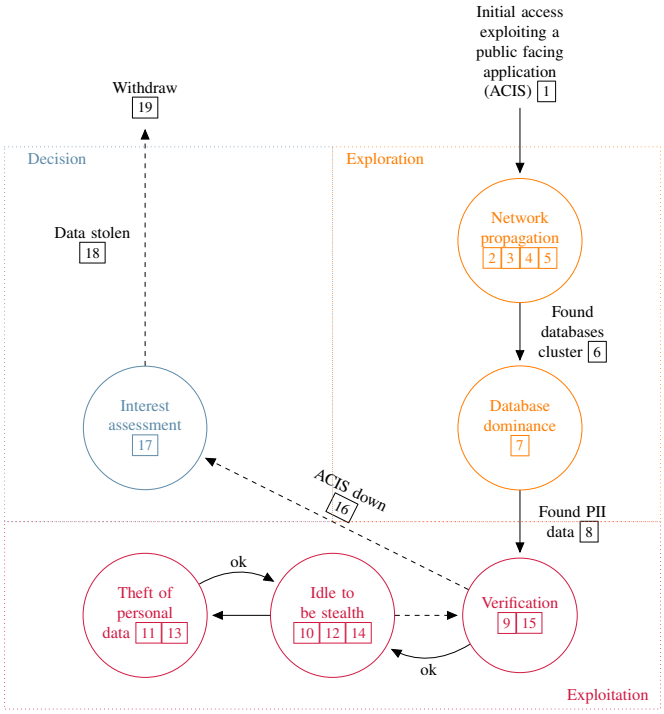

Fig. 7. Nuke model instantiation for the Equifax data breach

2) Exploration phase:

- **Network propagation**: The attacker began to discover the neighborhood of this compromised web server $\mathbf{e_{init}}$. He discovered that the ACIS environment was actually

two webservers and two application servers [2]. He installed webshells [3] ($\mathbf{t}_{1100}$) on all these newly discovered servers. Then he accessed a mounted file share ($\mathbf{t}_{1039}$) and discovered unencrypted credentials stored in a configuration file database ($\mathbf{t}_{1081}$). He continued his situational assessment and discovered ($\mathbf{t}_{1018}$) ($\mathbf{t}_{1046}$), outside the ACIS environment (but inside the network), 48 unrelated databases [4]: $\mathbf{K}_{Atk_2} = \mathbf{K}_{Atk_1} \cup \{DB1, DB2, ..., DB48\}$. The attacker reused application credentials to gain access to these databases and confirmed his capability to dig in it [5]. He knew that his final objective was actually the cluster of all these 48 databases but at this time, he didn't know how to reach it [6]. In order to own his objective, the attacker ran approximately 9000 queries to identify the type of database [7], by getting the metadata of specific tables.

- **Asset dominance** (database): Finally, the attacker found a table with Personally Identifiable Information (PII) [8]. He determined that the appropriate set of techniques to perform his objective were to run queries to retrieve the data from the database and to store it in files ($\mathbf{t}_{1005}$), file compression ($\mathbf{t}_{1002}$), put them on an accessible web server and download it out of the network ($\mathbf{t}_{1048}$).

3) Exploitation phase:

- **Capability check** (Verification): He checked his process and used a webshell to exfiltrate some of the data [9].
- **Idle** (to be stealth): We know that the attacker remained in the network for 76 days without being detected [10].
- **Effects on objective** (Theft of personal data): He has extracted and exfiltrated [11] slowly [12] [13] [14] personal information of 145.5 million US, UK and Canadian customers. For that he made 275 queries on the table which was his final objective. He did it on a long period in order to avoid volumetric-based detection.

4) Decision-making phase:

- **Interest assessment**: A few days after Equifax discovered the attack, the defenders decided, on July 30, 2017, to shut down the ACIS web portal [16]. Because it was the attacker's only entry point [15] and because he had already collected what he wanted [17], it ended the cyberattack. He withdrew [19]... with the stolen data [18].

B. TV5Monde sabotage (2015)

1) Context and initial foothold: TV5Monde is a French public television company founded in 1984 and broadcasting a dozen channels in nearly 200 countries with more than 350 million potential viewers. In 2015, the company was hit by an attack [10] (transcribed by Suiche [11]). This has been materialized in the real world with black screens for several hours for all but one of its channels. Figure 8 is this attack representation according to our model. As in the case of Equifax, in this model instantiation, the transitions between states that are explicitly described in the public report are noted in solid lines and the transition that we can only presume are noted by dashed lines.

After a first attempt to compromise the network via an isolated server in February 2015, he penetrated for the first time the TV5Monde network using the legitimate VPN ($\mathbf{t}_{1133}$) with a service provider account [1] ($\mathbf{t}_{1078}$). It provided him a foothold in the zone that allowed him to stay below the radar and be stealth. $\mathbf{K}_{Atk_0} = \{\mathbf{e_{init}}\}$ with $\mathbf{e_{init}} = VPN$. His main objective was to sabotage the broadcasting of television channels. But at that time, he still didn't know which asset was best suited for this action. However he did know that $\mathbf{e_{final}} \notin \mathbf{K}_{Atk_0}$.

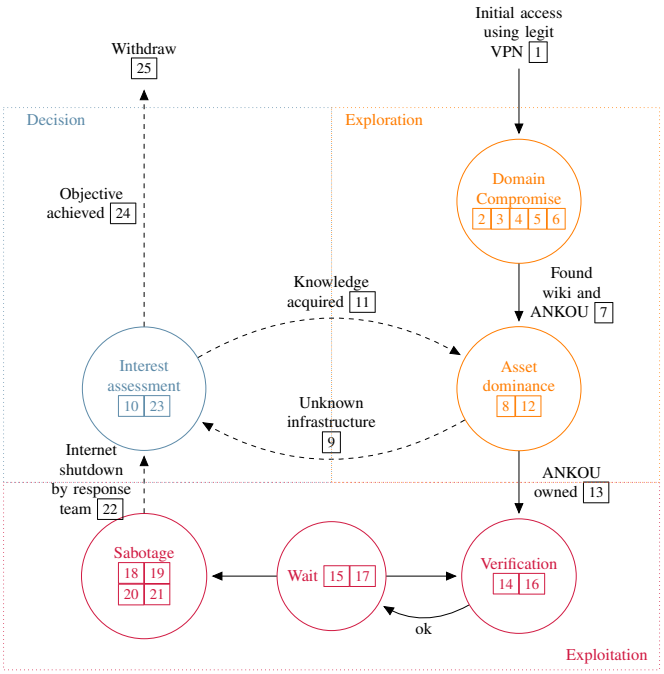

Fig. 8. Nuke model instantiation for the TV5Monde sabotage

2) Exploration:

- **Network propagation** (domain compromise): From his implantation point, he started to scan the VPN clients' LAN and the other TV5Monde internal networks [2] ($\mathbf{t}_{1046}$). On February 6th, he discovered two Windows computers called ROB1 and ROB2. He installed his RAT on these two new compromised assets [3]. Then he used another service provider account which was domain administrator to compromise the domain controller on February 11th [4], next he created a new domain administrator [5] ($\mathbf{t}_{1136}$) in order to preserve his privileges and ensure his progression. His knowledge of the network grew, he had discovered and compromised four computers. Then he targeted the TV5monde wiki [6] in order to collect information ($\mathbf{t}_{1083}$) and credentials ($\mathbf{t}_{1081}$) about TV5Monde's business. In these documents, he discovered an asset, called ANKOU [7], that he considered like his technical target because of its privileges and its location.

- **Asset dominance**: After the ANKOU asset discovery, he had all the technical and technological information that enabled him to learn how to achieve his objective [8]. Until mid-March, he performed searches outside the target network, probably because he didn't know this kind of infrastructure and he needed to build an attack scenario [9], the cost of these studies is not expensive for him [10], he came back in the compromised network a few days later [11]. He implanted ANKOU with his RAT [12]. At that moment, he fully owned the asset that would allow him to achieve his objective [13]

3) Exploitation:

- **Capability check** (Verification): A few days after his research, he came back to test that the accounts and the technical information he had were correct and allowed him to act on and from the final asset [14].
- **Idle** (Wait): Then there was no activity for a few weeks [15]. April 8, 2015, at 3:40p.m., he performed a last check to connect to multiplexers, encoders, switches and routers to be sure that he was able to destroy the network [16]. Then he waited for a few hours [17].
- **Effects on objective** (Sabotage): At 7:57p.m. he destroyed the IP configuration ($\mathbf{t}_{1498}$) of encoders and multiplexers [18]. At 8:58p.m. he defaced ($\mathbf{t}_{1491}$) website [19]. At 9:48p.m. he erased firmwares ($\mathbf{t}_{1495}$) of switches and routers [20] to finish his sabotage. At 10:40p.m. he deleted a few VM on ESX ($\mathbf{t}_{1485}$) including the messaging server [21]. At 11:50p.m. response team decided to shut down Internet connection. It made network access impossible for the attacker [22].

4) Decision-making phase:

- **Interest assessment**: He knew defenders were hunting him, the cost-interest ratio was too high [23], because he achieved his objective [24], he withdrew [25].

These confrontations allowed us to confirm that even with the few elements available in public reports, it was possible to represent lifecycles through the different phases of our model. We also find that all transitions to the *Interest assessment state* can be made from any other state, depending on the event.

V. Conclusion

Our model highlights that APT campaigns are divided into several phases, some using technical means and others involving third-party factors that make it possible to compare the cost of the attack with its actual interest. Moreover, it confirms the cyclical aspect and thus better perceive the so-called long-term concept of APT attacks. Unfortunately, the few publicly available reports with attacks' chronologies only allow us to suppose that these transitions to these new states we have introduced exist. In addition, these reports are biased since they present a threat eradication approach.In an active cyber defense perspective, our vision allows a defender to consider forcing the attacker to return to a previous state and to incite him to reveal his operational capabilities or simply to deter him from continuing the attack.

References

[1] R. S. Ross, "Managing Information Security Risk: Organization, Mission, and Information System View | NIST", Special Publication (NIST SP) - 800-39, 2011.

[2] E. M. Hutchins, M. J. Cloppert, and R. M. Amin, "Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains", 2011.

[3] Command Five Pty Ltd, "Advanced Persistent Threats: A Decade in Review", 2011.

[4] P. Pols, "The Unified Kill Chain: Designing a Unified Kill Chain for Analyzing, Comparing and Defending against Cyber Attacks", PhD thesis, Cyber Security Academy (CSA), 2017.

[5] MELANI:GovCERT.ch, "APT Case RUAG: Technical Report" [online]. Available: https://www.melani.admin.ch/dam/melani/en/dokumente/2016/technical%20report%20ruag.pdf [Accessed: 24-May-2019].

[6] B. E. Strom, J. A. Battaglia, M. S. Kemmerer, W. Kupersanin, D. P. Miller, C. Wampler, S. M. Whitley, R. D. Wolf, "Finding cyber threats with ATT&CK-based analytics", MITRE technical report, 2017.

[7] F. Maymí, R. Bixler, R. Jones, and S. Lathrop, "Towards a definition of cyberspace tactics, techniques and procedures", in 2017 IEEE International Conference on Big Data (Big Data), 2017, pp. 4674-4679.

[8] M. R. Endsley, "Toward a Theory of Situation Awareness in Dynamic Systems", Hum Factors, vol. 37, no. 1, pp. 32-64, 1995.

[9] U.S. House of Representatives, "The Equifax Data Breach", pp.31-39,54, 2018.

[10] ANSSI, "Retour technique de l'incident de TV5Monde" [video], 2017, Available at: https://static.sstic.org/videos2017/SSTIC_2017-06-09_P09.mp4 [Accessed 14 Sep. 2019].

[11] M. Suiche, "Lessons from TV5Monde 2015 Hack", Medium, 15-Jun-2017. [Online]. Available: https : / / blog . comae . io / lessons-from-tv5monde-2015-hack-c4d62f07849d. [Accessed: 14-Sep-2019].