

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Bodström, Tero; Hämäläinen, Timo

Title: A Novel Method for Detecting APT Attacks by Using OODA Loop and Black Swan Theory

Year: 2018

Version: Accepted version (Final draft)

Copyright: © Springer Nature Switzerland AG 2018.

Rights: In Copyright

Rights url: <http://rightsstatements.org/page/InC/1.0/?language=en>

Please cite the original version:

Bodström, T., & Hämäläinen, T. (2018). A Novel Method for Detecting APT Attacks by Using OODA Loop and Black Swan Theory. In X. Chen, A. Sen, W. W. Li, & M. T. Thai (Eds.), Computational Data and Social Networks : 7th International Conference, CSoNet 2018, December 18-20, 2018, Shanghai, China, Proceedings (pp. 498-509). Springer. Lecture Notes in Computer Science, 11280. https://doi.org/10.1007/978-3-030-04648-4_42

A Novel Method for Detecting APT Attacks by Using OODA Loop and Black Swan Theory

Tero Bodström and Timo Hämäläinen

Faculty of Information Technology, University of Jyväskylä,
P.O. Box 35, (Agora), 40014 Jyväskylä, Finland
{tetabods@student.jyu.fi,timo.hamalainen@jyu.fi}

Abstract. Advanced Persistent Threat(APT) attacks are a major concern for the modern societal digital infrastructures due to their highly sophisticated nature. The purpose of these attacks varies from long period espionage in high level environment to causing maximal destruction for targeted cyber environment. Attackers are skilful and well funded by governments in many cases. Due to sophisticated methods it is highly important to study proper countermeasures to detect these attacks as early as possible. Current detection methods under-performs causing situations where an attack can continue months or even years in a targeted environment. We propose a novel method for analysing APT attacks through OODA loop and Black Swan theory by defining them as a multi-vector multi-stage attacks with continuous strategical ongoing campaign. Additionally it is important to notice that for developing better performing detection methods, we have to find the most common factor within these attacks. We can state that the most common factor of APT attacks is communication, thus environment has to be developed in a way that we are able to capture complete network flow and analyse it.

Keywords: Advanced Persistent Thread(APT), OODA loop, Black Swan theory, Network anomaly detection

1 Introduction

In this paper a novel approach for analysing APT attack kill chain and its lifecycle is proposed to gain deeper insights about the attacks and help the development of detection techniques. The number of these sophisticated attacks is increasing thus leaving societal infrastructures at more risk which can, in worst case, cause lost of human lives. The current attacks are already capable of hiding in cyber environments, bypassing firewalls as well as intrusion detection systems and other countermeasures effectively. The attackers are currently adopting Machine and Deep Learning based solutions to attack campaigns, especially in malware that modify attack vectors and behavioural patterns based on what the malware learns from the cyber environment, making the attacks more dangerous and unpredictable. To effectively protect systems against these

above-mentioned campaigns, one must use similar techniques when developing countermeasures.

Our proposed approach studies APT attack's kill chain and lifecycle through OODA loop and Black Swan theory for describing how the attack behaves. Finally, we propose a method to detect APT attacks in an early phase of the campaign that can be adapted to Deep Learning.

The paper contains following structure: section II presents definitions for APT kill chain and lifecycle from information security organizations and research community and section III introduces the Black Swan theory and related research. In section IV, the OODA loop is described and section V explains thoroughly how the subjects from earlier sections are related to each other. Finally, section VI concludes and describes possible future works.

2 APT attack, kill chain and lifecycle

This section introduces APT attack and its kill chain utilising widely used definitions from research results and information security organizations.

2.1 APT attack

APT attack is a sophisticated network attack which attempts to breach target networks and systems undetected in order to espionage or gain access to privileged information as long as possible. APT attacks can be also used to cause maximal destruction to target networks, systems, critical infrastructure or production. A commonly known malicious worm Stuxnet is an example of APT attack that targets production systems. The developers of these sophisticated attacks are skilled and well funded since they are tailor-made to the target networks and systems[1–5] - the cyber security risk at the targeted side is very high.

Ongoing APT attack uses multiple techniques to masquerade its activity in a network from detection. These attacks have the capability to hide in networks by mimicking legitimate traffic and modify itself during campaign by using random execution intervals and multiple legitimate protocols. The attack can also use more than one attack vectors simultaneously.[1–10] Another weakness that help attackers is the fact that 100% secure systems does not exist. Even well designed and protected systems and networks have their weak spots[8]. Moreover, the growing number of insider threats, where an individual with legitimate system access executes an attack, is a great concern[5, 11].

2.2 Kill chain

The term "*kill chain*" originates from military concept and it was adopted to information security as "*Cyber Kill Chain*" by Lockheed Martin. This commonly used kill chain describes seven stages for an attack, which are i) Reconnaissance, ii) Weaponization, iii) Delivery, iv) Exploitation, v) Installation, vi) Command

and Control and vii) Actions on Objective[1, 2, 7]. Few other similar *kill chain* definitions to mention are: i) LogRhythm, ii) Lancaster, iii) SDAPT and iv) BSI-model [7]. However, these models are practically mechanical execution flows and they do not take into account how sophisticated the attack may be or the fact that attack developers do not have any obligations to follow the phases described in models.

There exist research papers where improvements and different approaches in defining *kill chain* are proposed. Messaoud et al. used the term "*life cycle*" to describe the entire kill chain in their paper. The purpose was to improve earlier definitions by focusing on attack targets as well as different tactics, techniques and methods controlling impact of the attack. This way, instead of understanding only the early stages of the attack which is common, it is possible to understand the attack's life cycle as a whole. They stated that earlier defined *kill chain* falsely assumes that APT attacks are using the same seven phases every time in that exact order. To support this argument, for example Stuxnet has a mechanism to autonomously execute several tasks without command and control (phase vi) and thus it is not following the *kill chain* execution order. Therefore, the authors proposed a new model which is based on six phases, where *kill chain* phases are combined and considered from attackers intention[7]. Bhatt et al. proposed framework for detecting APT's and improvements to *kill chain* in their paper. The purpose was to focus on attack vectors' simultaneous dynamic behaviour and detect them before significant impact to the target. The framework consists of three separate methods: i) multi-stage attack model, its core is based on seven-phase *kill chain*, ii) layered security architecture to delay attack success, which increases detection time and iii) event data collection from various sources and information analysis with Big Data technology. They stated that by using layered security, the entire *kill chain* has to be executed at least once in each layer which helps detection. For data collection they proposed separate sensors to each layer, configured to detect the ongoing *kill chain* phase[8].

Kill chain has not been the one and only approach on analysing APT attacks. For example, few recent papers[5, 10, 12] based their analysis on strategical game theory. Xiao et al. chose cumulative prospect theory (CPT) to improve APT attack detection dynamically. In their work attacker and defender are assumed to not know time intervals for the other opponent's actions, for example system scan for malware detection and attack execution intervals, which creates non-linear sequence of events. Both of the mentioned actions are strategic decisions with a different purpose. They stated that in both sides exists irrationality in strategical decision making under uncertain situations based on human subjective point of view, such as risk taking, thus strategic game theory fits well for APT detection and impact investigation[5, 10]. Zhu et al. proposed combination of three strategic game models to detect APT behaviour and recognize correct countermeasures. APT was defined as a multi-stage and -phase attack and it was divided to three stages: i) infection, ii) stealthy infiltration and iii) causing damage, and for each phase they selected different strategy, while between stages internal transition was implemented. Theory computes optimal behaviour

for both sides, attacker and defender simultaneously, and the theory's main purpose was to "*capture the strategic interactions of an attacker with a sequence of agents in the system*". The authors proposed *Gestalt Nash equilibrium* theory as a core of solution, as it provides a holistic risk assessment theory for APT attacks by adaptive learning methods and designing automated and optimal defence for multiple layers.[12] *Gestalt Nash equilibrium* theory was also analysed in[5] and identified as a best response strategy for opponents, optimizing their long-term objectives.

2.3 Lifecycle

The lifecycle of an APT attack begins when an attacker sets the target and intelligence gathering starts to find weak spots from the targeted cyber environment with methods such as open source intelligence, network scanning and social engineering[4, 7]. The lifecycle may end due to various reasons: the attack reaches its purpose, it is detected and interrupted or such countermeasures are implemented that it cannot fulfil its purpose and therefore ceases.

Intelligence gathering activities in a network, local and public, from anomaly detection point of view are difficult to detect, sometimes even impossible[6]. For example network scanning and open source intelligence can be executed completely outside of the target's cyber environment. Further, there can be multiple purposes for a network scanning, thus it is difficult to reason if it was executed by an APT or for some other purpose, even legitimate. Open source intelligence can be gathered from public records, such as DNS, whois records and so forth when targeted cyber environment is connected to public network, thus it does not create any communication to targeted environment. In addition, social engineering based attacks can be executed with methods that does not require access to target's cyber environment. Hence the critical point for network anomaly detection is, when a malware executes[8] first time and makes first communication or attempts it inside the targets cyber environment.

3 Black Swan Theory

According to Taleb's definitions, Black Swan is a surprising highly improbable consequential event, that can change the entire perspective to the subject in question. Black Swans are caused by "*severe limitation to our learning from observations or experience and the fragility of our knowledge*", in other words, one single observation can completely invalidate earlier common beliefs. In addition, there exists positive and negative Black Swans and while effects of positive takes time to appear, effects of negative ones appear fast[13].

To be more precise, Black Swan is an event that has following three attributes, i) it is an outlier that locates outside of the ordinary assumption, or outside of a "*tunnel*", ii) it causes extreme impact and iii) after the event occurs, even being an extreme outlier, there exists a tendency to transfer the event from being unlikely to explainable and predictable one[13].

A rare event is same as uncertainty and to study those events we need to focus on extreme outliers, instead of focusing on normal events. These dynamical sudden events with low predictability and high impact can be seen also as events that should not happen, thus happened exactly because of that. There are multiple reasons for this phenomenon: i) categorizing which reduces true complexity by ruling out sources of uncertainty, thus creating more misunderstanding, ii) by focusing on causes of known Black Swans or a small number of sources of uncertainty, in other words on a "tunnel", iii) huge amount of data can be insignificant occasionally while one piece of data can be very significant, thus invalidate earlier beliefs, iv) reducing dimensions, more random data has higher number of dimensions and it is more difficult to summarize. But as more data is summarized, less random it becomes thus leading to assumption that the environment is less random than it actually is, v) learning is based on historical events at the expense of previously unknown events, leading to ignoring later ones before appearance, vi) future predictions by using tools and methods which exclude rare events and vii) with more detailed knowledge from environment, noise in knowledge increases thus creating false understanding of information. Furthermore, two important attributes which relate to Black Swan events are "duration blindness", that is we cannot predict how long an event will last based on its history, and "the curse of learning", that happens with overlapping information, where less learning happens while having more overlap in information. Above all, any event can have an infinite amount of possible causes[13].

While Black Swans are produced in a messy understanding of an environment, to be more precise, in a gap between what we really know and what we think we know, there are methods to make them predictable to a certain degree and turn unknown unknowns to "Gray Swans". That is, they are somewhat tractable and by being aware of their existence, the element of surprise is lower. More specifically, the event is still rare but also expected. One way to reduce unknown is fractal randomness, it causes some Black Swan consequences to appear but does not give exact answers, in other words one can understand the consequence of an event without knowing the likelihood of an event. It is possible to infer such possible outcomes that are not directly visible in the data. However, instead of ignoring these deductions, they should be taken into account in the set of possible outcomes as well[13]. Zeng et al. proposed a hierarchical Bayesian reliability model framework to reduce unexpected failures in a process, which were presented as Black swans. Their mathematical model took all failures into account, expected and unexpected, and by increasing the knowledge of the environment their test results showed that probability of unexpected failures reduced. In addition, paper showed two important facts: i) "system complexity inherently hides unexpected failures" and ii) as the complexity of environment increases, possibility to unexpected failures increases also, thus "the estimated reliability decreases"[14]. Additionally Arney et al. stated that "Through rare event scenarios that impact the global network, we see how different elements and entities interact with each other to produce even greater impact", that is small change in local condition can cause consequences to entire network[15].

4 OODA loop

OODA loop, the decision cycle was presented by Colonel John Boyd and it was originally developed for observing and examining fighter pilots in aerial combat. Later on it has been adopted widely to business, law enforcement, military as well as to information security, as a decision making strategy[16–19].

Boyd stated that the people's ambiguity and the environments randomness creates a lot of uncertainty among us. However, the bigger problem is our inability to properly understand changing reality, that is, we find it difficult to change our perspectives according to prevailing conditions thus we tend to keep continuing with the existing mental concept. He also pointed out "*that trying to understand a randomly changing space with pre-existing mental concepts can only lead to confusion, ambiguity, and more uncertainty*". Previous idea is based on three principles: i) Gödel's Incompleteness theorem ii) Heisenberg's Uncertainty Principle and iii) 2nd Law of Thermodynamics. First, Gödel's Incompleteness theorem, states that every logical model of reality is incomplete, inconsistent and must adapt constantly to new observations. Second, Heisenberg's Uncertainty Principle, states that it is impossible to define the velocity and position of a particle at the same time. By applying this to environment, we will obtain more precise observations to exact domain but on the other hand we obtain more uncertainty to the other one. Third, 2nd Law of Thermodynamics, states that "*isolated system*" will have increasing entropy. By transferring this to organization, Boyd assumed that if individuals or organizations does not communicate with outsiders for getting new information, they will create mentally closed environment which leads to distorted perception of surroundings[16]. Therefore, it is important to frequently do situational assessments to gain awareness about the things that are going on in the environment[19].

The decision cycle has four phases, Observe - Orient - Decide - Act. Each phase in an OODA loop is a representation of a process, which is interacting with its environment. Its purpose is to resolve the earlier mentioned randomly changing space problem and increase awareness about the surrounding environment. Observe is a process for acquiring information about the environment through observing and interacting with it. It is guided and controlled by Orient while receiving feedback from Decide and Act. Orient is the process of filtering the information gathered in observe phase, taking into account the possible Orient-phases from previous loops. The filtering may include finding such correlations and dependencies that can be used in further decision making, therefore irrelevant information can be rejected. "*It shapes the way... we observe, the way we decide, the way we act*". The process that selects which hypotheses will be executed based on environment's situation, is Decide. It is guided by input from Orient and delivers feedback to Observe. In Act-process, the selected hypotheses are tested by interacting with the environment. It is guided and controlled by Orient, it receives feedback from Decide and it provides feedback to Observe[16].

It is important to remember that both sides, attacker and defender in our case, are supposedly using their OODA loops or comparable decision making strategies[5]. Boyd emphasized that by operating in faster decision cycle tempo

than an opponent, there the probability of winning is higher, or even better there exist a chance to get inside in an opponent's OODA loop. In other words, defender's cycle should perform with faster tempo than attackers[11, 16, 19]. Ma et al. pointed out that OODA loop is a robust encounter strategy for randomly changing environments and they also stated that effective OODA cycles generate encounter effectiveness[18]. Dapeng et al. executed OODA loop effectiveness simulation in their research. They stated that in theory, *"more timely and complete the information is more accurate and rapid the estimation, decision-making and action will be"*. That is, the opponent who gain more effective information, such that can be exploited on further attack phases, in certain time period, has more effective OODA cycle. Function table revealed that while gaining increasing amount of effective information, the accuracy of estimation increased as well. Their simulation showed that when opponents have the same tactical and technical performance, an opponent with more effective, or faster, OODA loop will win. An essential conclusion was, that it is more important to prevent the opponent gaining access to effective information than to improve tactical and technical performance[17]. However, Révay et al. stated that instead of faster OODA cycle tempo, Boyd actually meant rapid random changes in an OODA cycle tempo, which will create surprising and ambiguous behaviour thus confusing the opponent[16].

Fusano et al. tested OODA loop robustness with game theory by developing multi-agent combat simulation based on framework derived from game theory model and their multiple simulations conclusion supports Révay et al. lastly mentioned statement, opponent cannot win only by improving the performance based on observations of the other. To win, performance also requires modifications to the rules of orientation[20]. Another game theory research which supports same statement was simulated by Bilar et al., they developed a defence framework in order to identify malicious activities in a network. Their framework was based on fake targets: while the purpose was to detect suspicious behaviour, their intention was also to undermine opponents decision structure[21]. Thus OODA loop can be unified with a game theory in order to optimize own performance and create false understanding of the environment, which causes confusion to the opponent.

5 Deep analysis of APT attacks

In this paper, we propose a novel approach for evaluating the dynamics of APT attacks. Due to the complexity of attacks, our method tries to capture the multi-dimensionality of these attacks that multiple ongoing attack vectors, possibly in different stages, present. Instead of using strict flow controls or pure strategical theories we combine APT attack with OODA loop and Black Swan -theory in order to find the most common factor within these attacks.

5.1 APT attack and Black Swan Theory

APT attack can be considered as a rare event, thus it can be considered as a Black Swan. Although APT attack numbers are growing, they continue to be rare and for this there are two possible reasons, i) ongoing APT attacks are not detected or ii) APT attack detections are not reported publicly. In this section we focus on the first reason, where ongoing Black Swans are not detected. By comparing Black Swan three attributes and APT attacks, we can state i) APT attack locates outside of ordinary assumption, ii) APT attack causes high impact to target on purpose and iii) after discovery there exist a tendency to prove that APT attack is explainable and was predictable.

To detect these sophisticated attacks, we can consider Black Swan phenomenon reasons i-vii listed earlier in section III, "*duration blindness*" and "*the curse of learning*". When considering a detection method, we must recognize earlier mentioned problems and develop a solution from a new perspective, that is, i) keep true complexity of an environment, ii) expand focus also to outside of a "tunnel", iii) focus on all data, not just clusters, iv) not reduce data dimensions in order to reduce computational complexity, v) expand focus from historical events also to unknown events, vi) develop new tools which does not exclude rare events and vii) focus on data in original non-processed format. Even though there are known APT attacks that last from a month up to four years, it is not possible to predict with certainty that how long an attack continues. That is, the duration of one APT attack can be less than a day while another one may try to continue campaign as long as possible. Therefore, there does not exist a fixed duration for an attack campaign. To avoid "*the curse of learning*" a mechanism must be considered that learns from shorter time periods thus decreases overlap in information. Another important issue is that one has to know the environment correctly instead of making best guesses and creating countermeasures based on those.

5.2 APT attack and OODA loop

APT attack uses multiple simultaneous attack vectors which can be in different phases and different vectors may depend on the output of another vector. Therefore, we should think APT as multiple possibly dependent simultaneous OODA loops inside multiple attack vectors. This scheme is visualized in figure 1.

From the figure we can observe how these vectors 1-n have their dependent OODA loops. Timely execution of a vector's loop can depend on the results of earlier loop, the random execution interval of the vector and those loops from other attack vectors that send input data and execution commands. Before continuing to next stage, a vector can be in a halt until input from another vector arrives.

With this presented method, there is no necessity for knowing "*kill chain*" phases, thus we can observe that APT is a complex dynamical problem and we can state that APT a multi-vector multi-stage attack with continuous strategical ongoing campaign.

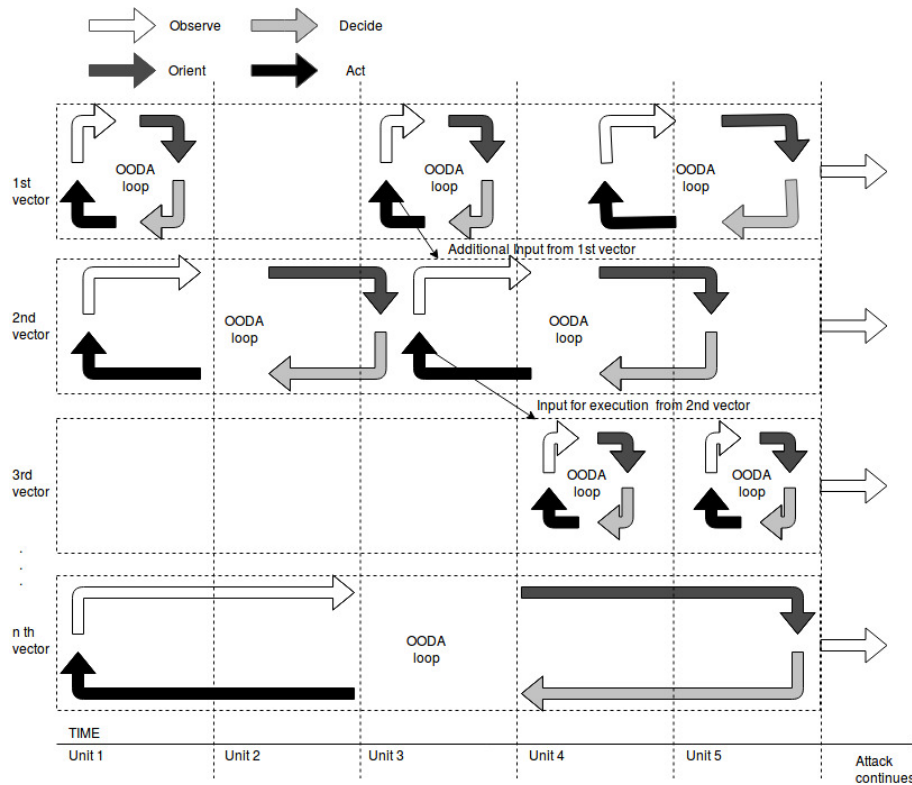


Fig. 1. Multi-vector multi-stage attack

5.3 From Black Swan to strategic decision

APT's were considered as a Black swans earlier, however, we can remove the surprise element, or at least reduce it with hypothesis "we have an ongoing APT attack in a network". By being aware of the possibility of attacks, we transfer those to Grey Swans and thus make them somewhat predictable, that is, attacks continue to be unknown unknowns as there is no actual detection available, but it is possible to set detection methods in place.

Considering APT attack behaviour, it can be divided to two types: i) programmed to fit environment, in other words tailored to known environment and ii) programmed to learn from environment and modify itself during campaign to fit to the environment. Difference is that in the first option, attacker has to know the environment completely during the development of malware, Stuxnet is one example of this type attack. In the second option, with the help of machine or deep learning an attack learns and modifies itself during the campaign. Both of these has one common functionality, after the execution they start to follow campaign strategy, which can cause behavioural patterns that can be detected.

5.4 Performance cost

When considering the proposed method from a technical point of view, we can state that it requires computational power, possibly a cluster of GPU's (Graphical Processing Unit) for running Deep Learning algorithms. Other resources include Random Access Memory (RAM) for buffering incoming network flow and also enough of hard disk space for saving outliers in a database.

As mirroring or replicating entire network flows require more network devices and planning, network capacity becomes also one concern. Furthermore, APT attacks are targeted to high value cyber environments thus it might not be beneficial to implement it to lower value environments until the prices of technical devices are low enough. In other words, cost benefit is an important factor when considering implementing the method.

6 Conclusion and Future Works

We can state as a fact that an APT attack uses some sort of communication in a network, otherwise malware would be an isolated malware in a hardware without further purpose defined for an APT. The communication can be, but not limited to, command & control to outside environment all the way to the attackers service. Campaign can also act independently inside a network and in this case, the communication happens between devices. However, communication is based on pre-programmed software logic, or logic through machine learning process from environment, and an APT starts to follow strategic instruction how to proceed with a campaign. Although an attack uses multiple simultaneous vectors with different phases, masquerades communication data, changes execution time intervals randomly, uses horizontal and vertical connections, mimics legitimate traffic, it communicates which leaves traces to network flow. Due to earlier mentioned sophisticated stealth techniques, an APT traffic can be statistically close to a normal traffic but it causes anomalies. However, it might be necessary to look deep into the binary level to find such anomalies. When detecting APT attacks, the focus should be in outliers, even in the tiniest ones, since an impact may cause huge damage to an environment.

We propose an approach to detect anomalies commonly present in APT attacks directly from network flow. When considering APT attacks, or more precisely their random execution intervals and long durations, real-time detection might not be possible nor necessary. Instead, the focus is to drop the detection time from years or months to an acceptable one, that is, days. To detect these complex attacks, there are few issues that must be considered: i) dimension reduction and overlapping information may cause outliers to vanish and ii) taking into account earlier detected outliers from historical data. Our earlier research results[22, 23] showed that deep learning methods have a high potential to resolve the considerations i) and ii). One concern is to locate sufficient amount of good quality network data, for executing training and benchmarking tests.

Based on these observations, we can state also that the most common factor of APT attacks is communication, thus environment has to be developed in

a way that we are able to capture complete network flow and analyse it for outliers. Additionally, we can setup decoys, for example honeypots, to create diversion from environment which can cause an attack to expose itself more easily. Furthermore, we have to consider how to detect attacks from legitimate outliers, to avoid false positives and even more serious false negative detections.

As a future works, we continue to study the proposed approach and to work on implementing such APT anomaly detection method that uses the ideas presented in the paper as well as determine which data types support detecting APT attack.

References

1. Brogi, G., Tong, V.V.T.: TerminAPTor: Highlighting Advanced Persistent Threats through Information Flow Tracking. In: 2016 8th IFIP International Conference on New Technologies, Mobility and Security (NTMS) (2016). <https://doi.org/10.1109/NTMS.2016.7792480>
2. Vukalović, J., D. Delija, D.: Advanced Persistent Threats Detection and Defense. In: 2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), pp. 1324-1330 (2015). <https://doi.org/10.1109/MIPRO.2015.7160480>
3. Chandran, P. Hrudya, P., Poornachandran,P.:An Efficient Classification Model for Detecting Advanced Persistent Threat. In: 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI) pp. 2001-2009 (2015). <https://doi.org/10.1109/ICACCI.2015.7275911>
4. Settanni, G., Shovgenya, Y., Skopik, F., Graf, R., Wurzenberger, M., Fiedler,R.: Acquiring Cyber Threat Intelligence through Security Information Correlation. In: 2017 3rd IEEE International Conference on Cybernetics (CYBCONF) (2017). <https://doi.org/10.1109/CYBConf.2017.7985754>
5. Hu, P., Li, H., Fu, H., Cansever, D., Mohapatra, P.: Dynamic Defense Strategy against Advanced Persistent Threat with Insiders. In: 2015 IEEE Conference on Computer Communications (INFOCOM) pp. 747-755 (2015). <https://doi.org/10.1109/INFOCOM.2015.7218444>
6. Ussath, M., Jaeger, D., Cheng, F.: Advanced Persistent Threats: Behind the Scenes. In: 2016 Annual Conference on Information Science and Systems (CISS), (2016). <https://doi.org/10.1109/CISS.2016.7460498>
7. Messaoud, B., Guennoun, K., Wahbi, M., Sadik, M.: Advanced Persistent Threat: new analysis driven by life cycle phases and their challenges. In: 2016 International Conference on Advanced Communication Systems and Information Security (ACOSIS) (2016). <https://doi.org/10.1109/ACOSIS.2016.7843932>
8. Bhatt, P., Yano, E.T., Gustavsson, P.M.: Towards a Framework to Detect Multi-Stage Advanced Persistent Threats Attacks. In: 2014 IEEE 8th International Symposium on Service Oriented System Engineering , pp. 390-395 (2014). <https://doi.org/10.1109/SOSE.2014.53>
9. Vance, A.: Flow based analysis of Advanced Persistent Threats detecting targeted attacks in cloud computing. In: 2014 First International Scientific-Practical Conference Problems of Infocommunications Science and Technology, pp. 173-176 (2014). <https://doi.org/10.1109/INFOCOMMST.2014.6992342>

10. Xiao, L., Xu, D., Mandayam, N.b., Poor, H.V.: Attacker-Centric View of a Detection Game Against Advanced Persistent Threats. In: IEEE Transactions on Mobile Computing (2018). DOI 10.1109/TMC.2018.2814052
11. Eidle, D., Ni, S.Y., DeCusatis, C., Sager, A.: Autonomic Security for Zero Trust Networks. In: 2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON), (2017). <https://doi.org/10.1109/UEMCON.2017.8249053>
12. Zhu, Q., Rass, S.: On Multi-Phase and Multi-Stage Game-Theoretic Modeling of Advanced Persistent Threats. In: IEEE Access (Volume: 6), pp. 13958-13971 (2018). <https://doi.org/10.1109/ACCESS.2018.2814481>
13. Taleb, N.: The Black Swan: The Impact of the Highly Improbable. Random House, New York (2007)
14. Zeng, Z., Zio, E.: Modelling Unexpected Failures with a Hierarchical Bayesian Model. In: 2017 2nd International Conference on System Reliability and Safety (ICSRS), pp. 135-139 (2017). <https://doi.org/10.1109/ICSRS.2017.8272809>
15. Arney, C., Coronges, K., Fletcher, H., Hagen, J., Hutchinson, K., Moss, A., Thomas, C.: Using Rare Event Modeling & Networking to Build Scenarios and Forecast the Future. In: 2013 IEEE 2nd Network Science Workshop (NSW), pp. 29-64 (2013). <https://doi.org/10.1109/NSW.2013.6609191>
16. Révay, M., Líška, M.: OODA Loop in Command & Control Systems. In: 2017 Communication and Information Technologies (KIT), (2017). <https://doi.org/10.23919/KIT.2017.8109463>
17. Dapeng, G., Jianming, H., Yuhu, Guoqian, X., Nainiang, Z.: Research on Combat SD Model based on OODA Loop. In: 2015 2nd International Conference on Information Science and Control Engineering, pp. 884-888 (2015). <https://doi.org/10.1109/ICISCE.2015.201>
18. Ma, L., Zhang, M., Zhou, Z.: The OODA loop robustness evaluation based on OSOS combat network. In: 2014 International Conference on Information and Communications Technologies (ICT 2014), (2014). <https://doi.org/10.1049/cp.2014.0583>
19. Blasch, E.P., Breton, R., Valin, P., Bosse, E.: User Information Fusion Decision Making Analysis with the C-OODA Model. In: 14th International Conference on Information Fusion, (2011).
20. Fusano, A., Sato, H., Namatame, A.: Study of multi-agent based combat simulation for grouped OODA Loop. In: SICE Annual Conference 2011, pp. 131-136, (2011).
21. Bilar, D., Saltaformaggio, B.: Using a Novel Behavioral Stimuli-Response Framework to Defend against Adversarial Cyberspace Participants. In: 2011 3rd International Conference on Cyber Conflict, (2011).
22. Bodström, T., Hämäläinen, T.: State of the art literature review on Network Anomaly Detection. In: to be published in the 18th International Conference on Next Generation Wired/Wireless Advanced Networks and Systems NEW2AN 2018, (2018).
23. Bodström, T., Hämäläinen, T.: State of the art literature review on Network Anomaly Detection with Deep Learning. In: to be published in the 18th International Conference on Next Generation Wired/Wireless Advanced Networks and Systems NEW2AN 2018, (2018).