Agile Approach for IT Forensics Management

Matthias Schopp
Research Institute CODE
Bundeswehr University Munich

Peter Hillmann
Research Institute CODE
Bundeswehr University Munich

Abstract

The forensic investigation of cyber attacks and IT incidents is becoming increasingly difficult due to increasing complexity and intensify networking. Especially with Advanced Attacks (AT) like the increasing Advanced Persistent Threats an agile approach is indispensable. Several systems are involved in an attack (multi-host attacks). Current forensic models and procedures show considerable deficits in the process of analyzing such attacks. For this purpose, this paper presents the novel flower model, which uses agile methods and forms a new forensic management approach. In this way, the growing challenges of ATs are met. In the forensic investigation of such attacks, big data problems have to be solved due to the amount of data that needs to be analyzed. The proposed model meets this requirement by precisely defining the questions that need to be answered in an early state and collecting only the evidence usable in court proceedings that is needed to answer these questions. Additionally, the novel flower model for AT is presented that meets the different phases of an investigation process.

1. Introduction

In the last year's industry, governments and institutions had to face more and more complex and enduring cyber-attacks, such as ThyssenKrupp [1] or RUAG [2]. Mandiant reported attacks on a wide spectrum of industries in 2015 [3]. The median number of days a threat group was active in a network before detection has been 205 days in year 2014, where the longest presence was 2982 days. Such advanced attacks (AT) are long-running, complex and often have the goal to steal or destroy data. The know-how of the attackers is high and therefore the attacks are highly sophisticated and heterogeneous. Furthermore, the attack tool support for the attackers are getting more and more professionalized, whereas the development of forensic software is second-order driven.

2. Problem Description

Digital forensic is defined as analytical and investigative techniques used for the preservation, identification, extraction, documentation, analysis, and interpretation of computer media (e.g. data on SSD/HDD/DVDs, log data on host systems or log data for network activities). When it comes to ATs the amount of computer media that has to be analyzed in a forensic process is increasing. This is partly due to its long duration and persistence. ATs have a higher sophistication level as basic or moderate attacks and are more difficult to clarify. In particular through the cautious and covert approach, whereby slow progress from the attack side is consciously accepted in order not to be discovered.

Forensic investigations of ATs therefore need to adapt to these characteristics. Classical forensic approaches that focus on individual systems (classical post-mortem forensic) and that collect all data in a first step and then try to extract hypotheses from this data is no longer expedient (shotgun forensic vs. sniper forensic [4]). This is because on the one hand not only individual systems are affected but also the communication between them (network forensic). On the other hand the amount of data becomes bigger and therefore big data analysis approaches are required that define questions asked in a first manner and then select the data that can answer them.

Therefore a forensic approach for ATs is needed that is agile and takes network forensic as well as big data into account. Due to the complexity of ATs the new model also aims to fit the different steps a high sophisticated intruder makes during the attack.

3. Related Work

For a digital forensic model, it is important to guide the investigation process in a continuously transparent way. When it comes to ATs the model also needs to be flexible because of the heterogeneity of ATs. In the past, several forensic models have been presented.

Du et al. [5] describe three phases in the current evolution of digital forensic process models. From 2000 to 2010 general models have been defined, such as McKemmish's SAP model (secure, analyze, and present) in [6] or more detailed ones such as the extended model of cyber-crime investigation [7] that enhances the model of Casey [8] that provides a scientific basis for the digital forensic process. These early frameworks have a linear procedure in common and they do not take ATs and big data problems into account.

In the second phase, these models have been refined and the newer approaches focus on particular steps in the investigation process. In [9] Agarwal et al. present a model that focuses on computer fraud and cybercrimes. None of them do treat ATs or focus on big data issues.

The third mentioned phase in [5] describes the recent research in this field. New technologies are taken into account such as cloud computing and internet of things (IoT). Some research has been done when it comes to big data and forensic. In [10] a framework for data reduction and data mining is presented that is not suggested to replace full analysis but serves to provide a rapid triage. They conclude that this approach is time consuming but has its advantages in the field of topic browsing.

All of the presented work are collecting data first and then try to reduce the amount of data or extract the relevant evidence instead of defining the questions asked in a first step. None of the existing models meets the requirements for an agile sniper forensic model. Although these aforesaid approaches can be used as guide in our new model.

4. Research Questions

The forensic process faces many challenges when it comes to ATs. The characteristics of these kind of attacks create a big data scenario that has to be addressed in research. The heterogeneity of ATs also has to be taken into account for a forensic model. Therefore the following research questions are expected to be answered.

1) What does the procedure look like when it comes to the forensic investigation of advanced attacks having regard to the heterogeneity of the attacks?

2) How can the relevant data for the investigation process be identified and forensically usable extracted having regard to the relevant questions and considering the big data problems?

3) How do host and network sensors (e.g. intrusion detection, network taps, etc.) have to be placed and forensically secured in the network to capture the needed data for the investigation process?

5. Agile Sniper Forensics for Advanced Attacks Approach

The novel approach focuses on the forensic investigation of advanced attacks and meets the modern challenges of an agile sniper approach. It consists of three parts: The two sub processes *Investigation* and *Documentation* as well as the novel flower model.

5.1 Flower Model

The flower model illustrates the attack procedure presented in Figure 1 taking Mandiant's attack lifecycle [3] and Hutchins et al. intrusion kill chain [11] into account. The model separates the attack in two perspectives, *Attacker* and *Target*. This is because from a forensic process' point of view in the step *Initial Information Gathering* an attacker uses public available information and legitimate tools and therefore this is hard to detect and to prove as part of the attack. Nevertheless, a regularly self-check to delete sensible public data or generatable information higher the attack difficulty for an intruder and should be part of the restoration process after the forensic investigation.

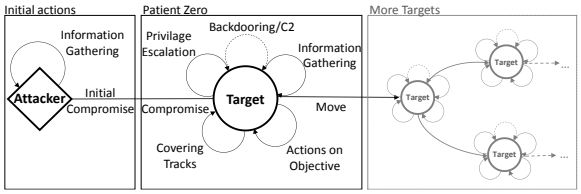

Figure 1. Flower Model for attack step illustration

The initial information gathering results in the initial compromise of the first target system. Each compromised system during the attack is represented as a flower (right hand side of the figure) with different steps as the flowers leaves. These represent the other phases of the attack such as *escalating privileges* after a successful compromise, *maintaining access* and control over the target through *backdoors* and the connection to a *command and control server (C2)*, *information gathering* on the host and/or network, e.g. through several scan techniques, *actions on objective* such as data exfiltration and *covering tracks*. The arrow compound the result either of the attacker's i*nitial compromise* or the action *move* from the perspective of a *Target*. ADs usually include multiple systems that are compromised due to lateral movement steps. These steps are illustrated by the *move* action in the Flower Model where the first compromised system's (patient zero) C*ompromise* action results from the *initial compromise* of the attacker. The model also takes into account that there can be more than one initial compromised system, which leads to multiple *initial compromise* actions, as well as more than one *move* action.

The *Target*'s get connected to an attack graph that represents the involved compromised systems and illustrates the attack chain.

5.2 Invesitgation and documentation method (IDM)

The digital forensic process makes use of the flower model to comprehend the AT. Figure 2 shows this agile sniper forensic approach.

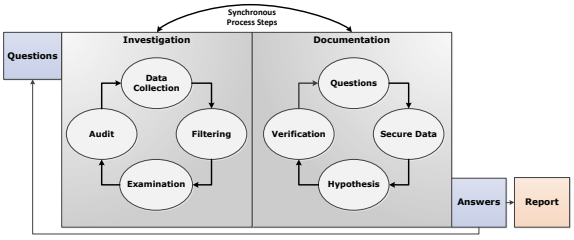

Figure 2. Agile Sniper Forensic Model

It is divided into two sub processes, *Investigation* and *Documentation* where these two processes exist simultaneously. The global management process starts with defining the general *questions* that need to be answered to forensically reproduce the AT. For each found compromised system, the questions related to it are defined and the sub processes are done. According to Greiner [12] examples for such questions are:

- How did the attacker(s) get onto the system?
- What did the attacker(s) do on the system and what did they steal?
- How did the attacker(s) escape / move laterally?
- How did the attacker(s) get the stolen data off the system?

The *Investigation* sub process starts with collecting the data that is needed to answer the defined questions. These data can be rough categorized in host, network and misc. Sources of information are usually log files and temporary files as well as status information like timestamps, user rights, and processes. When it comes to AT the amount of data that needs to be analyzed is big. Therefore in our approach these data sources that are relevant to the defined questions need to be defined and only that data is collected. This subset of data than is filtered to get the specific data for the separate questions. It is examined so hypotheses can be developed that answer the defined questions. The results are then checked with the found data.

In parallel the *Documentation* sub process documents the investigation steps. The questions for extracting the necessary data as well as the steps to collect it are recorded. Therefore an official and common agreed process is mandatory, which is integrity securing. The collected data is stored with checksums and signatures in a matter that its integrity is provable. During the examination in the *Investigation* process the developed hypotheses also need to be documented. The last step is the reporting of the audit actions to verify the hypotheses.

If more data is needed to answer the questions, other filters need to be applied, hypotheses are refuted, or new hypotheses are developed that need to be proven, these sub processes will be repeated until the answers to the questions are proven with the data.

The whole process is also repeated until the initial attack vector for the compromised system is found and its origin (system or attacker) is found. The model is designed to adapt the questions depending on the found evidence on a system. This is to be agile on the steps taken by the attacker. Each compromised system is treated this way during the digital forensic investigation. The documentation builds the separate flowers from the flower model with its leaves and connects them so the complete attack is forensically reproducible in the report.

6. Evaluation

We analysed 21 technical reports on ADs ([2], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32]). These reports analysed one or more attacks that used similar types of malwares and tools per report. To show that an agile approach is needed we examined the amount of targets involved in such an attack and the steps taken by the attackers.

The reports show that in all 21 cases multiple targets were part of the attacks. The six types of actions on a target – the leaves of the flower model – were also present on the analysed reports. The actions and tools used by the attackers were heterogeneous and targeted on their objectives. Therefore the forensic analysis process needs to be agile to the attack. The presented approach will be evaluated in more detail in a practical manner in future work.

7. Conclusion and Next Steps

In this proposal, the need for a digital forensic management approach for advanced attacks was pointed out. With the agile sniper forensic model processes were presented as a first step towards such a management approach. For the ease of reproducibility also the flower model was introduced that can be used for documentation purposes.

In the next steps, this forensic approach has to be refined and tested in a practical manner. Therefore a specific tool support will be developed. It also has to be analyzed what techniques can be used to identify and extract relevant data and also where to place sensors to capture the relevant data.

8. References

[1] L. Mathews, 2016, ThyssenKrupp, "Attackers Stole Trade Secrets In Massive Hack", https://www.forbes.com/sites/leemathews/2016/12/08/thyssenkrupp-attackers-stole-trade-secrets-in-massive-hack/,
Access Date: 05 February 2019.

[2] GovCERT.ch, 2016, "APT Case RUAG", https://www.melaniadmin.ch/dam/melani/en/dokumente/2016/technical%20report%20ruag.pdf.download.pdf/Report_Ruag-Espionage-Case.pdf,
Access Date: 05 February 2019.

[3] Mandiant, 2016, "M-Trends 2015: A view from the front lines", https://www2.fireeye.com/rs/fireye/images/rpt-m-trends-2015.pdf,
Access Date: 05 February 2019.

[4] C. E. Pogue, 2009, "Sniper forensics – one shot, one kill", https://digital-forensics.sans.org/summitarchives/2010/2-newell-spiderlabs-sniper-forensics.pdf,
Access Date: 05 February 2019.

[5] X. Du, N. Le-Khac, and M. Scanlon, "Evaluation of Digital Forensic Process Models with Respect to Digital Forensics as a Service," CoRR, vol. abs/1708.01730, 2017.

[6] R. McKemmish, "What is forensic computing?" Trends and Issues in Crime and Criminal Justice, vol. 118, 2002, https://aic.gov.au/publications/tandi/tandi118,
Access Date: 05 February 2019.

[7] S. O. Ciardhuain, "An extended model of cybercrime investigations," vol. 3, no. 1. International Journal of Digital Evidence, 2004.

[8] E. Casey, "Digital Evidence and Computer Crime", San Diego: Academic Press, 2000.

[9] A. Agarwal, M. Megha Gupta, M. Saurabh Gupta, Y. GUPTA, and C. Gupta, "Systematic digital forensic investigation model", pp. 2011–118, 2011.

[10] D. Quick and K.-K. R. Choo, "Data Reduction and Data Mining Framework for Digital Forensic Evidence: Storage, Intelligence, Review and Archive," vol. 480, pp. 1–11, 2014.

[11] E. M Hutchins and M. J Cloppert and R. M Amin, "Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains", Leading Issues in Information Warfare & Security Research, 2011.

[12] L. Greiner, "Sniper forensics", netWorker 13, 2009.

[13] GROUP-IB and FOX-IT, "Anunak: APT against financial institutions", 2014, https://www.group-ib.com/resources/threat-research/Anunak_APT_against_financial_institutions.pdf,
Access Date: 12 February 2019.

[14] Kaspersky Labs - Global Research & Analysis Team, "Carbanak APT -The Great Bank Robbery", 2015, https://media.kasperskycontenthub.com/wp-content/uploads/sites/43/2018/03/08064518/Carbanak_APT_eng.pdf,
Access Date: 12 February 2019.

[15] F-Secure Labs, "Cozyduke", 2012, https://www.f-secure.com/documents/996508/1030745/CozyDuke,
Access Date: 12 February 2019.

[16] Cylance, "Operation Cleaver", 2014, https://www.cylance.com/content/dam/cylance/pdfs/reports/Cylance_Operation_Cleaver_Report.pdf,
Access Date: 12 February 2019.

[17] Kaspersky Labs - Global Research & Analysis Team, "The Darkhotel APT A Story of unusual hospitality", 2014, https://media.kasperskycontenthub.com/wp-content/uploads/sites/43/2018/03/08070903/darkhotel_kl_07.11.pdf,
Access Date: 12 February 2019.

[18] Kaspersky Labs - Global Research & Analysis Team, "Equation group: Questions and answers", 2015, https://wikileaks.org/ciav7p1/cms/files/Equation_group_questions_and_answers.pdf,
Access Date: 12 February 2019.

[19] RSA, "RSA Incident Response: Emerging Threat Profile Shell_Crew", 2014, https://www.emc.com/collateral/white-papers/h12756-wp-shell-crew.pdf,
Access Date: 12 February 2019.

[20] C. Raiu and M. Golovkin, "The Chronicles of the Hellsing APT: the Empire Strikes Back", 2015 https://securelist.com/analysis/publications/69567/the-chronicles-of-the-hellsing-apt-the-empire-strikes-back/,
Access Date: 12 February 2019.

[21] Kaspersky Labs - Global Research & Analysis Team, "The 'Icefog' APT: A Tale of Cloak and Three Daggers", 2013, https://media.kaspersky.com/en/icefog-apt-threat.pdf,
Access Date: 12 February 2019.

[22] Kaspersky Labs - Global Research & Analysis Team, "The Regin Platform - Nation-State Ownage of GSM Networks", 2014, https://media.kasperskycontenthub.com/wp-content/uploads/sites/43/2018/03/08070305/Kaspersky_Lab_whitepaper_Regin_platform_eng.pdf,
Access Date: 12 February 2019.

[23] K. Baumgartner and M. Golovkin, "The MsnMM Campaigns – The Earliest Naikon APT Campaigns", Kaspersky Lab, 2015, https://media.kasperskycontenthub.com/wp-content/uploads/sites/43/2018/03/07205555/TheNaikonAPT-MsnMM1.pdf,
Access Date: 12 February 2019.

[24] GReAT, "OlympicDestroyer is here to trick the industry", Kaspersky Lab, 2018, https://securelist.com/olympicdestroyer-is-here-to-trick-the-industry/84295/,
Access Date: 12 February 2019.

[25] McAfee Labs and McAfee Foundstone Professional Services, "Protecting Your Critical Assets – Lessons Learned from 'Operation Aurora'", 2010, https://www.wired.com/images_blogs/threatlevel/2010/03/operationaurora_wp_0310_fnl.pdf,
Access Date: 12 February 2019.

[26] FireEye Labs, "APT28: A Window Into Russia's Cyber Espionage Operations?", 2014, https://www.fireeye.com/content/dam/fireeye-www/global/en/current-threats/pdfs/rpt-apt28.pdf,
Access Date: 12 February 2019.

[27] FireEye Labs, "APT30 and the Mechanics of a Long-Running Cyber Espionage Operation", 2015, https://media.kasperskycontenthub.com/wp-content/uploads/sites/43/2015/05/20081935/rpt-apt30.pdf,
Access Date: 12 February 2019.

[28] Clearsky - Cyber Security, "Thamar Reservoir - An Iranian cyber-attack camapaign against targets in the Middel East", 2015, https://www.clearskysec.com/wp-content/uploads/2015/06/Thamar-Reservoir-public1.pdf,
Access Date: 12 February 2019.

[29] R. Dela Paz, "The HeartBeat APT Campaign", Trend Micro, 2012, https://docplayer.net/17881843-The-heartbeat-apt-campaign.html,
Access Date: 12 February 2019.

[30] Kaspersky Labs - Global Research & Analysis Team, "The Desert Falcons Targeted Attacks", 2015, https://media.kasperskycontenthub.com/wp-content/uploads/sites/43/2018/03/08064309/The-Desert-Falcons-targeted-attacks.pdf,
Access Date: 12 February 2019.

[31] Symantec, "The Waterbug attack group", 2016, https://www.symantec.com/content/en/us/enterprise/media/security_response/whitepapers/waterbug-attack-group.pdf,
Access Date: 12 February 2019.

[32] N. Villeneuve and J. T. Bennett and N. Moran and T. Haq and M. Scott and Kenneth Geers, "Operation 'Ke3chang': Targeted Attacks Against Ministries of Foreign Affairs", FireEye Labs, 2015, https://www.fireeye.com/content/dam/fireeye-www/global/en/current-threats/pdfs/wp-operation-ke3chang.pdf,
Access Date: 12 February 2019.

