Detecting Multi Stage Attacks Using Sequence to Sequence Model

First Author:Mrs. Sk.Karimunni
Second Author:K.Sumanthi
1.Assistant professor,Dept of MCA,
Audisankara college of engineering and technology (AUTONOMOUS), Gudur,AP, India
2.PG Scholor, Dept of MCA, Audisankara college of engineering and technology
(AUTONOMOUS) Gudur ,AP, India

ABSTRACT:

Multi-stage attacks are one of the most critical security threats in the current cyberspace. To accurately identify multi-stage attacks, this paper proposes an anomaly-based multi-stage attack detection method. It constructs a Multi-Stage Profile (MSP) by modeling the stable system's normal state to detect attack behaviors. Initially, the method employs Doc2Vec to vectorize alert messages generated by the intrusion detection systems (IDS), extracting profound inter-message correlations. Subsequently, Hidden Markov Models (HMM) are employed to model the normal system state, constructing an MSP, with relevant HMM parameters dynamically acquired via clustering algorithms. Finally, the detection of attacks is achieved by determining the anomaly threshold through the generation probability (GP). To evaluate the performance of the proposed method, experiments were conducted using three public datasets and compared with three advanced multi-stage attack detection methods.

1. INRODUCTION

Nowadays, the Internet expands rapidly by merging the traditional operation techniques (OT) with the new information techniques (IT) and gradually interconnects nearly all the things of our physical society into the cyberspace, forming a new Internet of things (IoT) paradigm (Carruthers, 2016). With this trend, a large number of industrial control systems (ICS), such as SCADA, PLC etc., previously running on an isolated and local industrial environment are now connecting to the Internet for a more convenient operation and efficient communication (Aijaz and Sooriyabandara, 2018). However, since these ICSs are usually served on the critical infrastructures of our national economy and safety such as the energy supplies, nuclear powers, medical treatments and financial markets, it is necessary to take the network security as a frontier prerequisite when they open to the public Internet (Wolf and Serpanos, 2017). As a result, the network protections (e.g., intrusion detection systems or firewalls) have been widely deployed in the entrance (i.e., edge servers or network gates) from the Internet to these ICSs (Mitchell and Chen, 2014), hence largely limiting the intrusion risks to the underlying infrastructures.

However, even the network infrastructures are well protected by intrusion detection systems (IDS) or firewalls, they are still attracting the hackers' or the hacking organizations' main attention due to their high value in economics and the potentially large impacts on the physical society (McLaughlin et al., 2016). Especially for the case of Internet battle at the national level, the enemy country's ultimate target is usually on the rival's critical infrastructures that are definitely under strict protections (Jones, 2016). To bypass these protections, attackers are evolving to be more intelligent to discover the protection's weakness and design flaws and sometimes exploit social engineering tricks for penetration, which results a single intrusion task with multiple attacking stages, namely multi-stage attacks (Navarro et al., 2018). For example, the Havex attack reported by the ICS-CERT 2014 (NCCIC ICS-CERT, 2014) can be considered as a kind of typical multi-stage attack. In particular, the Havex completes its intrusion with at least four stages: first, it takes reconnaissance to discover the ICS's vendor site that is not well protected and poison a backdoor to the site's softwares. Second, it bypasses the targeted ICS's protection by the ICS downloading the polluted software. Third, it compromises the whole ICS by exploiting a zero-day vulnerability reside in the OLE for Process Control (OPC) component. Fourth, it chooses to hide in the ICS for further control or do the real damage to the underlying physical infrastructures (Rrushi, 2018). The havex has been found effective to avoid the IDSs and firewalls deployed in a wide area of energy, aviation, pharmaceutical and petrochemical industries in the United States and Europe, causing more then millions of dollars economic losses world wide. Moreover, many other well known ICS intrusion events such as the Stuxnet, Flame, Duqu, BlackEnergy etc. Willems (2019) are also organized with multiple stages and usually last for a long duration, resulting in an advanced persistent threat (APT) to their targeted industries and finally making a huge impacts on our human society (Cole, 2012).

Compared with the intrusion detection for a single attacking action, detecting multi-stage attacks is more challenging. On the one hand, by organizing a single intrusion mission into a sequence of attacking actions (i.e., stages), the attackers can intelligently masquerade their true intrusion purpose by interleaving irrelevant packets into different stages or launching different actions in a long period of intervals (Salah et al., 2013). On the other hand, the detection of the intrusion with multiple stages is necessarily stateful. That is, to recognize the relationships and dependencies among different stages, the detectors must remember the past alerts in the observation sequence and progressively investigate all the possible state permutations with the historical sequence growing up. As a result, the smart attackers may purposely enlarge the dispersion between different attacking actions in order to increase the length of stage dependency and make the condition of permutation more complex. State-of-the-art detection solutions to multi-stage attack usually build on a hidden markov model (HMM) (Chadza, Kyriakopoulos, Lambotharan, 2020, Chen, Guan, Huang, Ou, 2012, Holgado, Villagrá, Vazquez, 2017, Kholidy, Erradi, Abdelwahed, Azab, 2014, Shawly, Elghariani, Kobes, Ghafoor, 2019) that considers the observed alerts as a markov chain with hidden states (including but not limited to attacking stages). They discover the most likely stage sequence by iterating the possible hidden state permutations, hence not able (error-prone or complexity unacceptable) to effectively detect the long-term stage dependency.

2. LITERATURE SURVEY

In recent years, various methods have been proposed to address issues related to multi-stage attack detection, with Bayesian models and HMM being the primary approaches for detecting multi-stage attacks. Ren et al. [15] introduced a multi-stage attack detection method based on Bayesian models, which divides the detection process into two stages. Firstly, it employs Bayesian networks to automatically extract correlations and constraints between alerts, testing different features to find the most accurate descriptors for attack stages. Then, based on the selected features, it extracts attack scenarios from the alert stream. Marchetti et al. [16] proposed the use of Bayesian models to calculate alert correlations, identify whether alerts belong to the same attack scenario, and generate an alert correlation graph. However, these methods reconstruct attack scenarios through alert correlations, requiring a significant amount of prior knowledge and increasing the complexity of maintaining a secure system. Moreover, these methods merely replicate attack scenarios and do not detect the stages at which attacks occur.

As early as 2003, HMM was employed to address the issue of MSA detection. HMM is a dual stochastic process [17], and in the field of statistical machine learning, it is considered one of the most suitable techniques for multi-stage attacks detection. The main reason for this is its mathematically tractable form for analyzing input-output relationships and generating transition probability matrices based on training datasets. D. Ourston et al. [18] utilized HMM for detecting multi-stage attacks and compared it with two other classical machine learning algorithms, decision trees, and neural networks. The results showed that HMM outperforms decision trees and significantly surpasses neural networks in multi-stage attacks detection. Chen et al. [19] proposed the introduction of HMM in the cloud for attack sequence detection, defining the detection of multi-stage attacks as a state-based classification model. Holgado et al. [20] provided a more detailed introduction to how HMM can be applied to multi-stage attacks. They defined states based on Common Vulnerabilities and Exposures (CVE) statistics, combining multi-stage attack data with the CVE database. They explained how to construct HMM using supervised and unsupervised learning methods, namely using the Baum-Welch algorithm or statistical frequency methods to train model parameters. The Viterbi algorithm and forward-backward algorithm can be used to determine the most likely attack stages. Suratkar et al. [21] introduced an HMM-based Host Intrusion Detection System (HIDS) model that consists of an anomaly detection module built using Long Short-Term Memory (LSTM) and multiple HMM modules for multi-stage attack detection. Due to the significant impact of HMM parameters on detection performance, Chadza et al. [22] designed an effective detection framework combining transfer learning and HMM. They trained HMM on labeled data and transferred the learned parameters to new tasks. Unlike other discrete modeling techniques, HMM excel in hidden states and transitions, thereby eliminating the need for complete information before attack detection.

Furthermore, With the advancement of big data technology [23], some deep learning methods have been applied to multi-stage attack detection. Deep learning approaches can overcome some limitations of traditional shallow machine learning, capturing deep-seated features within the data [24], and enhancing detection performance [10]. Vinayakumar et al. [25] introduced a deep learning framework for detecting zombie networks, which operates at the application layer of DNS services. This framework works by distinguishing between normal behavior and zombie network behavior. Sudheera et al. [26] conducted research on multi-stage attacks and proposed a distributed multi-stage attack detection method. Their work addressed the spatiotemporal challenges of zombie network attacks. They used alert-level and pattern-level information as features and employed machine learning methods to identify various attack stages within generated alerts. Xu et al. [27] designed an LSTM network based on multiple feature layers. They introduced a stage feature layer to store and compute historical data to identify different stages of multi-stage attacks with varying durations. Then, they used a time series feature layer to link independent attack stages and analyze whether the current data is within a certain attack cycle. However, these methods struggle to detect unknown attack paradigms and have lower detection effectiveness for new attack behaviors.

In the method proposed in this paper, we leverage the relative stability of the system's normal state to construct an MSP of the normal state. This MSP allows us to label alerts that do not conform to the normal state as attacks, thereby achieving multi-stage attack detection. The method introduced in this paper can directly build statistical detection models from the raw data of alerts without the need for additional expert knowledge or specific attributes. Additionally, it can detect the specific stages at which alerts occur, while also addressing the limitation of traditional methods in detecting unknown attack paradigms.

3. PROPOSED WORK

Firstly, in phase 1, First, in Phase 1, we designed an automated data acquisition method to obtain alert information from network traffic. Deployed IDS in the network continuously analyze traffic data captured from the network environment and generate alerts when suspicious packets are detected based on predefined rules. IDS may not detect complete multi-stage attacks, but when attackers attempt to infiltrate through multiple attack stages, IDS may capture individual attack actions and issue corresponding alerts. In Phase 1, the system primarily faced performance pressure stemming from IDS traffic analysis, and therefore, we adopted an offline analysis strategy to avoid impacting the overall system performance.

However, it is worth noting that IDS systems generate a significant number of false positives. These alerts result from the inability of the alert generation rules to distinctly differentiate between normal and malicious activities within the network, and thus, do not represent genuine security threats [28]. Nevertheless, these non-attack stage alerts often carry information about the system's activity patterns, serving as a means to describe the system's normal state. Alerts generated by IDS are stored in the Alerts Database, which encompasses alert data from the system's normal state (referred to as non-attack stage alerts) and alert data from multi-stage attack states (referred to as attack stage alerts in this paper).

In phase 2, we introduce a method for alert preprocessing with the goal of transforming the text-style alert data generated in Phase 1 into data that can be used by machine learning algorithms. Since HMM cannot directly process alert information, we convert the alert data from the Alerts Database into vectors using the Doc2Vec algorithm. This allows us to extract deep-seated information from the alerts and analyze the associations between alert entries for further processing. The performance overhead in Phase 2 primarily stems from the training of the Doc2Vec model. Therefore, similar to Phase 1, we adopt an offline training strategy in Phase 2.

Then, the proposed MSP is constructed in Phase 3. We initially use a clustering approach to automatically obtain the stage division of normal alert vectors, which is then mapped to hidden states in the HMM to complete the construction of MSP based on HMM. Training the clustering model and HMM introduces a significant performance overhead in this phase. Therefore, similar to previous phases, we also employ an offline training approach.

Finally, in Phase 4, we perform online detection of alert data using the constructed MSP. The probability generated by the MSP is used as the basis for determining anomalies. This probability is compared to a predefined threshold to decide whether the sequence is anomalous. If it is, then the alert is marked as an attack stage alert. In contrast to Phases 1, 2, and 3, the detection process in Phase 4 is conducted online, using the MSP model obtained in the offline training of Phase 3. Online detection is nearly real-time, and it imposes relatively low performance overhead.

Alert preprocessing

The semantic description of alerts can be seen as a sequence of statements, and if the context of two alert descriptions is similar, it can be considered that they have similar semantics. In multi-stage attacks, the attacker’s actions are intentional, and the alerts from attack stages also exhibit certain characteristics. Similar attack methods result in similar alert information. Therefore, by learning alert semantic representations from a large number of alert sequences, it is possible to effectively represent alerts.

In order to extract the semantic description of alerts and use it for further computation, we need to represent them in vectorized form. There are various methods for vectorizing semantic descriptions, such as the Bag-of-Words model, One-hot Encoding, and others. However, these methods do not capture the relationships between words in the alert information, and their sparse representation can lead to the curse of dimensionality. Word2Vec addresses the dimensionality issue but loses sequence information by averaging word vectors. When using Word2Vec to compute text similarity, keyword extraction algorithms may not perform accurately. To address these issues, the alert preprocessing model proposed in this paper utilizes the Doc2Vec model to transform alert descriptions into low-dimensional continuous values, mapping semantically similar alert descriptions to nearby positions in the vector space, thus extracting semantic knowledge from the alert descriptions.

4. CONCLUSION

In this paper, we present an anomaly-based multi-stage attack detection method. By modeling the normal state of a stable system and constructing an MSP, our aim is to detect attack behaviors. Our objective is to develop a model capable of detecting unknown pattern attacks, not just common ones like DoS. Our approach starts by vectorizing alert information to better capture the deep-seated information within alerts. Next, we process the vectorized data of non-attack stage alerts, using clustering and HMM to build the MSP. Finally, we perform detection on the alert data collected by IDS, using the alert's fit to the MSP's generated probability as the basis for judgment to determine if the alert belongs to an attack stage.

5. REFERENCES

1. Ingale S, Paraye M, Ambawade D. A survey on methodologies for multi-step attack prediction. In: 2020 Fourth International Conference on Inventive Systems and Control (ICISC). IEEE; 2020. p. 37–45.

2. Navarro J, Deruyver A, Parrend P. A systematic survey on multi-step attack detection. Computers & Security. 2018;76:214–249. doi: 10.1016/j.cose.2018.03.001 [CrossRef] [Google Scholar]

3. Kotenko I, Gaifulina D, Zelichenok I. Systematic literature review of security event correlation methods. IEEE Access. 2022;10:43387–43420. doi: 10.1109/ACCESS.2022.3168976 [CrossRef] [Google Scholar]

4. Wang X, Gong X, Yu L, Liu J. MAAC: Novel alert correlation method to detect multi-step attack. In: 2021 IEEE 20th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom). IEEE; 2021. p. 726–733.

5. Husák M, Komárková J, Bou-Harb E, Čeleda P. Survey of attack projection, prediction, and forecasting in cyber security. IEEE Communications Surveys & Tutorials. 2018;21(1):640–660. [Google Scholar]

6. Mao B, Liu J, Lai Y, Sun M. MIF: A multi-step attack scenario reconstruction and attack chains extraction method based on multi-information fusion. Computer Networks. 2021;198:108340. doi: 10.1016/j.comnet.2021.108340 [CrossRef] [Google Scholar]

7. Cheng Z, Sun D, Wang L, Lv Q, Wang Y. MMSP: A LSTM Based Framework for Multi-Step Attack Prediction in Mixed Scenarios. In: 2022 IEEE Symposium on Computers and Communications (ISCC). IEEE; 2022. p. 1–6.

8. Zhou P, Zhou G, Wu D, Fei M. Detecting multi-stage attacks using sequence-to-sequence model. Computers & Security. 2021;105:102203. doi: 10.1016/j.cose.2021.102203 [CrossRef] [Google Scholar]

9. Shawly T, Elghariani A, Kobes J, Ghafoor A. Architectures for detecting interleaved multi-stage network attacks using hidden Markov models. IEEE Transactions on Dependable and Secure Computing. 2019;18(5):2316–2330. [Google Scholar]

10. Dhasarathan C, Shanmugam M, Kumar M, Tripathi D, Khapre S, Shankar A. A nomadic multi-agent based privacy metrics for e-health care: a deep learning approach. Multimedia Tools and Applications. 2023; p. 1–24. doi: 10.1007/s11042-023-15363-4 [PMC free article] [PubMed] [CrossRef] [Google Scholar]

11. Fraser N, O'Leary J, Cannon V, Plan F. Apt38: Details on new north korean regime-backed threat group. FireEye, October. 2018;3. [Google Scholar]

12. Ye N, et al. A markov chain model of temporal behavior for anomaly detection. In: Proceedings of the 2000 IEEE Systems, Man, and Cybernetics Information Assurance and Security Workshop. vol. 166. Citeseer; 2000. p. 169.

13. Chadza T, Kyriakopoulos KG, Lambotharan S. Analysis of hidden Markov model learning algorithms for the detection and prediction of

multi-stage network attacks. Future generation computer systems. 2020;108:636–649. doi: 10.1016/j.future.2020.03.014 [CrossRef] [Google Scholar]

Author's Profile,:

Mrs. SK.karimunni currently she is working assistant professor in Audisankara college of Engineering and technology Gudur,Tipati(dt).

She is done M.Tech from Quba college of Engineering and technology ,Venkachalam in 2015.

KATURU SUMANTHI is pursuing MCA from Audisankara college of engineering and technology (Autonomous) NH-5 Bypass Road, Gudur, Tirupati (dt) Andhrapradesh, India