Predictive Analysis of Ransomware Attacks using Context-aware AI in IoT Systems

Vytarani Mathane[1], P.V. Lakshmi[2]
Department of Computer Science and Engineering
GITAM University, Vishakhapatnam, India

***Abstract*—Ransomware attacks are emerging as a major source of malware intrusion in recent times. While so far ransomware has affected general-purpose adequately resourceful computing systems, there is a visible shift towards low-cost Internet of Things systems which tend to manage critical endpoints in industrial systems. Many ransomware prediction techniques are proposed but there is a need for more suitable ransomware prediction techniques for constrained heterogeneous IoT systems. Using attack context information profiles reduces the use of resources required by resource-constrained IoT systems. This paper presents a context-aware ransomware prediction technique that uses context ontology for extracting information features (connection requests, software updates, etc.) and Artificial Intelligence, Machine Learning algorithms for predicting ransomware. The proposed techniques focus and rely on early prediction and detection of ransomware penetration attempts to resource-constrained IoT systems. There is an increase of 60 % of reduction in time taken when using context-aware dataset over the non-context aware data.**

Keywords*—Ransomware; IoT; context-aware; machine learning; ontology

I. Introduction

IoT systems are distinct from others in that they are ubiquitous, heterogeneous in capabilities, and usually out in adversarial environments [1]. They are present in Industries, Medical centers, Smart cars, Smart homes, Smart cities, and supply chains [2]. Such IoT systems could be susceptible to multiple categories of attacks like Denial of Service (DoS), botnets, man in the middle, identity and data theft attacks, ransomware attacks given the less than the secured or controlled environment of deployment and quite often limited security capabilities [3]. Among all those ransomware attacks could be more impacting owing to attack methodology where victim systems become unusable until a ransom is paid, typically have attacker-defined timelines to respond, and can cause more monetary loss.

Ransomware attacks, one of the malware attacks affect all types of security issues availability which causes monetary losses, and sensitive information loss [4]. Crypto ransomware, locker ransomware, and hybrid ransomware are common types of ransomware [5][6]. In crypto-ransomware attacks, data files are encrypted and the decryption key is provided only after paying the ransom. In locker ransomware attacks, the resources are blocked and are released only after paying the ransom. In hybrid ransomware attacks, both concepts of crypto-ransomware and locker ransomware are used. BadRabbit, Petya-Esque, Scareware, Screen Lockers, WannaCry are some famous ransomware attacks. By using Botnets, Social engineering, and malvertisement (malicious advertising) ransomware can penetrate IoT devices [6].

Context information of a typical device includes individuality, activity, location, time, and relation [7]. The prediction model has to utilize one or more of these categories to predict a ransomware attack. Location for tracking target and source, time for identifying the time of events occurring on the device, activity to find the set of events that leads to suspicious activity, relation to identifying the dependency between events, and individuality to identify the device through unique characteristics. These features are modeled and used for attack prediction. The context-aware prediction models can use different techniques such as graphs, anomaly detection, classification, clustering, etc.

II. Related Work

AI algorithms are used for cyber defense, malware prevention, and advanced threat detection or prevention [8]. Machine learning is used to learn about the attacks and predict them or machine learning for learning attacks and pattern matching for predicting them [9]. MIT labs developed an AI2 platform to predict cyber-attacks using AI [10]. IoT systems use AI algorithms for attack and anomaly detection [11]. Support Vector Machine (SVM) model as it is good for predicting very specific attacks [9]. According to [12] they use SVM to detect and predict ransomware attacks. SVM is good for detecting zero-day attacks which are unknown [7][13].

Ransomware in IoT can affect the integrity, confidentiality, and availability of the system and can cause monetary losses and loss of sensitive information [14]. In [15] 18 families of ransomware are studied and developed a model for categorizing behavioral characteristics, which can be used to improve the detection of ransomware attacks. [16] uses weighted KNN machine learning technique to detect and predict ransomware attacks on software-defined networking. The author in [17] uses neural networks for detection of ransomware in Industrial IoT where there is a huge risk.

Context-awareness is achieved by [18] using Context ontologies and Ontology description logic to get dynamic context attributes. The author in [19] use known attack context profiles to detect specific attacks that are relevant to a particular context and to avoid false-positive alerts. Known attack context profiles are created using conditional entropy [20][21]. The author in [22] uses sensor ontologies according to the semantic needs of IoT solutions. Ontologies can be

categorized into device ontology, domain ontology, and estimation ontology. Semantic metadata like context, description of the sensor, and its configuration provides contextual information. The proposed paper uses a classification model using contextual features. Section 2 describes the related work on context-aware ransomware attack prediction. Section 3 describes the framework and design of predicting ransomware attacks. Section 4 describes the implementation and Section 5 shows the comparison of solutions using with and without context-aware features.

A. Anatomy of a Typical Ransomware Attack on IoT

There has been a significant amount of research on ransomware threats to the IoT segment [14][15][16]17] and it offers very significant insights into ransomware penetration in the area of IoT, attack vectors, methodologies, and few specific implementation details (like Windows APIs) used for attacks. From the analysis of previous ransomware attacks on IoT, the ransomware executes in the following stages:

- Stealth mode where ransomware attacker benefits as long as the attacked system does not detect ransomware.
- Suspicious mode where ransomware starts collecting vital stats required to assess the suitability of specific targets within the system and starts encrypting or locking those.
- Obvious mode where attacker and ransomware display messages to the victim with a chosen mechanism to the victim.

Predominately Windows-based workstations used in IoT grids, Proof of Concept on low-end IoT devices (smart bulbs, smart TV, etc.) are the typical IoT systems that are being attacked.

Crypto, Locker, and Hybrid are different types of ransomware attacks [6]. The attack vectors used are Content distribution, Social engineering, Malvertisement, Downloaders & botnets, Email phishing, and R-a-a-S (Ransomware as a service) [6].

The typical flow of successful ransomware attacks shows certain patterns. Attack made leveraging social engineering goes through a sequence where the victim is made to download ransomware, elevate privileges of ransomware and/or current user, exploit elevated privileges and & locally downloaded ransomware, establish a connection back to command center to make victim submit to demands. Another case of a ransomware attack on network interfaces goes through cyber scanning, enumeration, intrusion attempt, the elevation of privilege, perform malicious tasks, deploy malware/backdoor, delete forensic evidence, and exit.

Other patterns emerging out of existing data are also pointing to increased integrated fingerprinting as a part of mounting ransomware attacks. Such fingerprinting is used to vital data to decide on the usefulness of the content in extortion schemes, Usefulness of content is seen to be analyzed based on a multitude of factors like date & times of content creation, usage of content, location of content in the system, geolocation data, file extensions, file names & entropy of the content.

Since a significant number of attacks were targeted towards Windows OS-based IoT endpoints & IoT servers, there are few studies that leveraged analysis of the Windows APIs used and traversed during attacks to build prediction capabilities. So far most promising models are unfortunately based on very high-level sequences & context, e.g., specific sequences followed by ransomware attacks on network stack can hold the potential key to discovering IoT attacks in real-time.

B. Observations and Deductions from Past Studies

One can make three observations based on the analysis and outcomes of previous studies focused on ransomware attacks on IoT as below [14][15]:

- This anatomy of a typical ransomware attack on an IoT system as described in the prior section allows us to make a safe conclusion that it applies to a very specific section of IoT devices using Windows OS and hence use moderately powered CPUs and other resources. Ransomware attack prediction models built using such data are also applicable largely to such Windows-powered IoT systems.
- Content & hence content analysis plays an important role in the current attack landscape to detect victim system's suitability for exploitation followed by most suitable contents (files, directories, etc.) to execute one of ransomware attack technique (encryption, locking, hybrid).
- Third-social engineering plays a very significant role as an enabler to fetch ransomware into the victim system. The use of social engineering needs to factor in a user being present on the system to intentionally or unintentionally allows download and installation of such malware content. Without such a user being present, the ease of ransomware finding its way to the victim system reduces greatly.

Contrasting these observations of a type of IoT systems attacked, capabilities such systems possess, and attack vectors used with a low-end microcontroller and microprocessor-based IoT provides us a path forward. Such lower-end IoT systems could be a very interesting target because of several reasons:

- These systems could provide a much greater period of stealth and suspicious modes as those typically are unsupervised or do not have a human operator.
- These systems do control vital and critical nodes, operations within a grid and hard to pinpoint for fault analysis given the nature of deployment.
- These systems have tremendous heterogeneity lacking standard & widely used OS capabilities, underlying hardware, need to fine-tune ransomware for each such target system effectively making large scale deployment hard prospect.
- These systems typically do not store data but rather used as control endpoint or sensor endpoints, so the crypto category of ransomware attack does not have meaningful gains.

- These systems also provide smaller attack surface because of limited or none endpoint level user involvement, social engineering vulnerabilities.

Nonetheless these studies, patterns observed, APIs leveraged by ransomware can still be used to make some progress towards ransomware prediction capabilities for lower-end resource constrained IoT devices. Such devices are expected to be spread in a grid, typically control key elements in a grid and if attacked can also bring down large industrial critical infrastructures. It is only to be expected that attackers would want to leverage ransomware to cripple such ground-level IoT devices to maximize damage inflicted to scare victims into paying a ransom.

Segmentation is OSes used in a variety of low-end IoT systems; varying hardware also does not help much ability to build a predictive model as it leads to segmentation of data observed from such systems. One way is to up-level predictive models from specific APIs and capabilities to allow such heterogeneity of implementations.

III. Design Methodology

Building on the previous section, we aim to provide a solution for predicting ransomware attacks in lower-end IoT systems (which has been largely neglected so far) using context-aware AI algorithms methodologies.

A. Building Context Parameters

Context is further defined considering the following factors: target IoT systems, deployment vectors, and attack vectors. Common and specific use cases for target IoT systems are sensor nodes, controllers to a specific function in the power grid, valve controllers on the dam, etc. All such use cases imply that to target specific capability, ransomware needs to collect information about ports, memory addresses, etc.

More prevalent methods of ransomware deployment like social engineering, malvertisement, email phishing does not make any sense to IoT, whereas the following methods can be leveraged to deploy ransomware to lower-end IoT systems: content distribution, downloaders & botnets, and Ransomware as a service (R-a-a-S). Attack Vectors consists of various events, activities, or APIs associated with the above deployment vectors including software, firmware update capabilities (system-specific APIs downloading new firmware packages and overwrite existing memory contents), connection requests /traffic in and out of the system on available transmission protocols (Wi-Fi, BT, etc.), port scanning (scanning for memory input/output port addresses) and use of cryptographic APIs & underlying accelerators or software libraries (typically OpenSSL AES encryption APIs).

B. Scope and Design of the Solution

Further narrowing on the scope of this proposed solution, attacker's entry attempts into an IoT system via TCP/IP or BT-like protocols is focused. This is in line with the strategy that prevention is better than cure and in as early stage as possible. Port scanning, scanning for cryptographic APIs indicate the attacker is already in the system and for the current discussion it is beyond the scope.

Context-awareness is achieved using context ontology and developing attack profiles. The data is provided to the AI models and the attack is predicted. The design of the solution includes the following main components: data collection, Context ontology (for feature extraction), attack context filters, Classification algorithm (for pre-diction), Result alert. Fig. 1 shows the design components of the proposed solution.

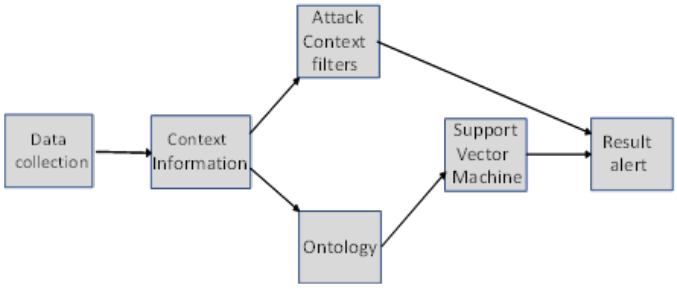

Fig. 1. Design Components of the Solution.

1) Data collection: IoT devices communicate via network stack, packets are collected using a network tool. Thus, collected data is used as training data which comprises benign and ransomware attack traffic. The data set collected from our testbed is represented in a JSON file format.

2) Context ontology: The collected data is represented using context ontology which follows logic and structure and automates the information retrieval. The context ontology uses the data collected from the data collection unit.

As mentioned above only a subset of known ransomware attack vectors (content distribution, downloaders & botnets, and Ransomware as a service) are likely to be used in attacking a typical industrial IoT. In this study, a context ontology for specific attack vectors of downloading ransomware or ransomware infected software images to IoT devices within the network is designed. The activity context information with features such as attacker, target IoT systems, and network events or activity towards downloading ransomware to target IoT device is built. Similarly, one could develop a context for using a content distribution like device configuration or parameters and Ransomware-as-a-service but those are out of scope for the present study.

3) Attack filters: The activity context information is used to create attack context profiles for classifier algorithms. The attack profiles are a set of features that are important to detect the attack. The feature selection is based on a set of rules followed to ensure detecting the attack. The feature vector can be represented using the equation (1):

$$F = \{f1, f2, \ldots, fn\} \quad (1)$$

Following Fig. 2 illustrates how to build attack filters for typical ransomware penetration of an IoT device within a specific IoT network using a download attack vector. Such a scenario comprises an attacker node attempting to impersonate authorized software or content distribution entity which would further attempt to detect possible target IoT devices by doing ipsweep and port scan for devices listening for software or content updates. Once such IoT systems are found, an attacker

node would attempt spoofing as a legitimate content provider and subsequently compromise IoT devices or devices with ransomware infected software or content. As these authors mentioned earlier in this text, detecting and avoiding imminent ransomware attacks is still the best defense against a ransomware attack and this methodology achieves the said purpose.

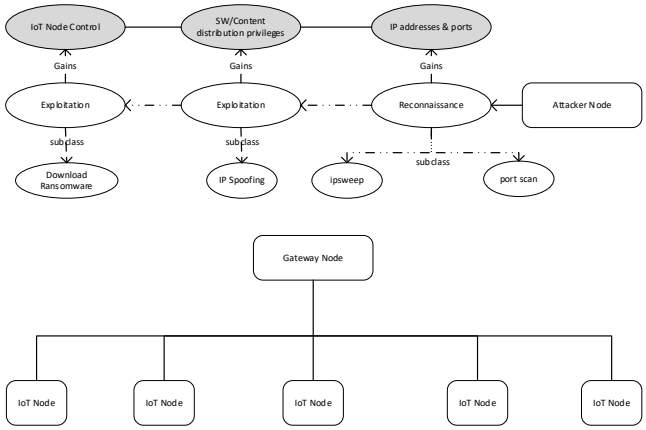

Fig. 2. Building Attack Filters of Ransomware Penetration within IoT Network.

4) Classification algorithm: The feature vectors are fed to classifier algorithms such as SVM to find the attack and give an alert as output. SVM modeling algorithm has to find the optimal hyperplane to classify the data. Optimal hyperplane maximizes the margin of training data. The training data set is a set of n elements (xi, yi) where xi is a p dimensional vector, the definition (2) can be given as smallest ||w|| will be giving biggest margin,

Minimize in (w, b) (2)

||w|| subject to yi (w.xi +b) >= 1

(For any i= 1, …, n)

IV. Results and Discussions

Fig. 3 shows a simplified but typical topological view of a typical industrial IoT network with attack paths/vectors and subsequent events. It involves mater node and several endpoint nodes associated with various data acquisition units. Such an entire deployment is usually managed by a dedicated server. As mentioned in the previous section, deployment paths to build predictive models for determining the probability of ransomware attacks have been focused. Subsequent events like port scanning, encryptions, lockouts of data acquisition units are out of scope for our discussion. Hence our testbed is one such network where the master node is leveraged to deploy attacks on endpoints. The master node would typically connect on TCP/IP or BT interface with its endpoints.

Dataset is collected on the testbed in JSON format. An ontology tool is used to get context-aware data to develop attack filters. We used classifier model SVM and training data with context-aware data and without context-aware data is fed into the model. It is tested with the test dataset. In this method, we overcome the heterogeneity of IoT devices. Context-aware dataset saves time to the tune of 60% compared to the non-aware dataset. Table I presents the time taken by the original dataset and context-aware dataset.

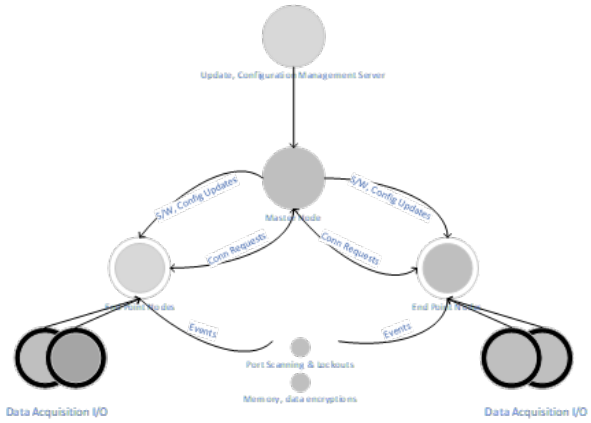

Fig. 3. Topological view of Typical Industrial IoT Network with Attack Paths/Vectors and Subsequent Events.

TABLE I. Time Taken by Original Dataset and Context-aware Data Set

Experiment	Time taken by original dataset and context-aware data set.	
	Original dataset	***Context-aware dataset***
1	3 ms	0.9 ms
2	3 ms	0.8 ms
3	2.8 ms	0.8 ms
4	3.1 ms	1 ms

V. Conclusion

A methodology to build prediction models for ransomware attacks on industrial IoTs is developed by focusing on their specific behavior common to most of such devices to overcome challenges posed by their inherent heterogeneity. In this paper, context awareness is used for identifying the most relevant attack paths, vectors, and resultant events to build more effective prediction capabilities.

Acknowledgment

We would like to thank the GITAM University for proving us with the necessary infrastructure for doing this research.

References

[1] A. Giusto, et al., “The Internet of Things,” Springer, ISBN: 978-1-4419-1673-0, 2010.

[2] L. Atzori, et al., “The Internet of Things: A Survey,” Computer Networks, Vol. 54, Issue 15, 2787-2805, 2010.

[3] T. Aliya and L. Wadha, “Security Framework for IoT Devices against Cyber-Attacks,” Computer Science & Information Technology (CS & IT), 249-266, 2019.

[4] Y. Ibrar, et al., “The rise of ransomware and emerging security challenges in the Internet of Things”, Computer Networks. Volume 129, Part 2, 444-458, 2017.

[5] M. U. Kiru and A. B. Jantan, “The Age of Ransomware: Uniderstading Ransomware and its countermeasures,” Artificial Intelligence and

Security Challenges in emerging networks, R. Abassi, Ed. Pennsylvania: IGI Global, pp. 1–37, 2019.

[6] A. Wani and S. Revathi, "Ransomware protection in loT using software defined networking," International Journal of Electrical & Computer Engineering. Vol. 10 Issue 3, 3166-3175, 2020.

[7] A. Aleroud and K. George, "Contextual information fusion for intrusion detection: a survey and taxonomy," Knowledge and Information Systems, Vol. 52, 563–619, 2017.

[8] Davidson, et al., "Security Gets Smart with AI," SANS Institute, 2019.

[9] H. Martin, et al., "Survey of Attack Projection, Prediction and Forecasting in Cyber Security," IEEE Communications Surveys & Tutorials, 640 – 660, 2018.

[10] V. Petri and L. Martti, "Artificial intelligence in the cyber security environment," The 14th International Conference on Cyber Warfare and Security, Stellenbosch, South Africa, 2019.

[11] H. Mahmudul, et al., "Attack and anomaly detection in IoT sensors in IoT sites using machine learning approaches," Internet of Things, Vol. 7, 2019.

[12] T. Yuki and F. Satoshi, "Detecting Ransomware using Support Vector Machines," Proceedings of the 47th International Conference on Parallel Processing Companion, Article No.1, pp 1–6, Eugene OR USA, 2018.

[13] J. Song, et al., "A generalized feature extraction scheme to detect 0-Day attacks via IDS alerts," Proceedings of the 2008 international symposium on applications and the internet, pp 55–61, Turku, Finland, 2008.

[14] R. Syed, et al., "Ransomware and Internet of Things: A New Security Nightmare", 9th International Conference on Cloud Computing, Data Science & Engineering, Noida, India, 2019.

[15] H. Gavin, et al., "Ransomware deployment methods and analysis: views from a predictive model and human responses", Crime Science, Vol. 8, 2019.

[16] C. Hong-Yi, et al., "Implementation of ransomware prediction system based on weighted-KNN and real-time isolation architecture on SDN Networks," IEEE International Conference on Consumer Electronics, Taiwan, 2019.

[17] A. Muna and S. Elena, "Industrial Internet of Things Based Ransomware Detection using Stacked Variational Neural Network," Proceedings of the 2019 conference on big data and Internet of Things, Melbourn VIC Australia, 2019.

[18] S. Alireza, et al., "A Context-Aware Malware Detection Based on Low-Level Hardware Indicators as a Last Line of Defense," SECURWARE 2017: The Eleventh International Conference on Emerging Security Information, Systems and Technologies, pp 10-19, Rome, Italy, 2017.

[19] A. Ahmed and K. George, "A Contextual Anomaly Detection Approach to Discover Zero-Day Attacks," International Conference on Cyber Security, Washington, DC, USA, 2012.

[20] C.E. Shannon, The Mathematical Theory of Communication. Univ. Illinois Press, 1971.

[21] A. Ahmed and K. George, "A System for Cyber Attack Detection Using Contextual Semantics," 7th International Conference on Knowledge Management in Organizations: Service and Cloud Computing, Vol. 172, pp 431-442, Baltimore, MD 21250, USA, 2013.

[22] M. Gergely, and J. Abonyi, "A Review of Semantic Sensor Technologies in Internet of Things Architectures," Hindawi Complexity, Vol.2019,pp1-21,2019.