ROBUST MALWARE DETECTION FOR INTERNET OF (BATTLEFIELD) THINGS DEVICES USING DEEP EIGENSPACE LEARNING

Bugga Veena Goud[1], Malyala Shivaraju[2],Golla Praveen[3], Mr. Muddam Kotesh[4],Dr.M.L.M. Prasad[5]

[123]UG Student, Department of Computer Science and Engineering - AI&ML, Joginpally B R Engineering College

[4]Assistant Professor, M.Tech (CSE), Department of Computer Science and Engineering - AI&ML, Joginpally B R Engineering College

Email: muddamkotesh@gmail.com

[5]Associate Professor,PhD,Department of Computer Science and Engineering - AI&ML, Joginpally B R Engineering College

Email: mlm.prasad@yahoo.com

Abstract :

With the growing interconnection of technologies in warfare scenarios, devices in the IoBT have become integral to modern-day military operations. And yet, this rising interconnectivity exposes such systems to greater cyber threats, particularly malware sophisticated enough to disrupt missions or steal sensitive information. A robust malware detection solution for IoBT devices is proposed in this paper, which leans on deep Eigenspace learning to detect malicious behavior at the very core. The system models software execution from discrimination-point code execution sequences-the operational code or OpCode-as program behavior into a highly feature-rich vector space. The deep learning model proposed detects very small discrepancies from benign software behavior to detect malicious intent and accurately classify any given software as either malware or benign. On top of this, the model has been put through its paces against common evasion techniques such as junk code insertion, endowing the model with a high degree of immunity while maintaining performance. This will therefore act as a scalable and smart way to shore up the cybersecurity of battlefield-connected systems.

Keywords : Internet of Battlefield Things (IoBT) ,Malware Detection, Deep Learning , OpCode Analysis, Eigenspace Learning,Cybersecurity, Junk Code Insertion ,IoT Security

I . INTRODUCTION

The gradual transformation in modern warfare has helped popularize smart interconnected systems popularly known as the Internet of Battlefield Things (IoBT). The systems involve a diverse array of battlefield-deployed devices-amongst wearable sensors, communication nodes, autonomous vehicles, and medical devices-that monitor, share, and process data in real-time. IoBT, at the same time, increases awareness, operational efficiency, and decision-making in war operations. Since these devices are enhanced through increased interconnectivity and they require less dependency on which cybersecurity issues arise, it gives important places as targets for opposite agendas in a cyber war. Malware is one of the most common threats in the IoBT landscape. Once infected, IoBT nodes can start leaking sensitive information, halt operations, and even hijack mission-critical assets. Because of the strategic value of such systems, malware targeting IoBT is oftentimes produced by nation-state actors with advanced evasion techniques. In such conditions, the conventional method of signature-based detection will almost be unusable, barring the inability to detect a novel or an obfuscated malware variant in action. This puts in emphasis the urgent need for intelligent, dynamic, resilient malware detection solutions aimed at resource-constrained and mission-critical situations posed by the IoBT. Our research presents the first malware detection framework, employing deep Eigenspace learning for recognizing malicious behavior on IoBT devices. The principle behind this is to analyze the sequences of Operational Code (OpCode) which correspond to actual instructions a program executes. Once executed, the program's OpCodes are extracted from the software binaries and mapped into a numerical feature space indicative of the application's behavioral patterns. With deep learning under an Eigenspace framework, the method separates malicious code from benign code even if the intruder employs heavy obfuscation

techniques involving junk code insertion and control flow manipulation. The proposed methodology is called lightweight, scalable, and well-suited to the constraints of an IoBT device. It also has the capability to detect known threats with very high accuracy and can generalize well for unknown malware families. We validate our approach via an experiment that features a curated dataset of real-world samples of malware and benign software, where, for a variety of threat models, our method scores very well. In the end, this research advances the newly formed field of IoBT security by contributing a data-driven, practical malware detection solution. The release of the dataset to the public continues to sustain research and development for the protection of next-generation military networks against emerging cyber threats.

II. LITERATURE SURVEY

In the recent past, many enhancement drives have taken place in investigating cyberattacks and malware for IoT environments through advances in DL and ML. Some approaches have been tried to secure IoT systems, mostly restricted by various resources and broadly scattered. Taşcı (2024) has introduced another architecture based on DL for detection of IoT attacks and malware. It is said that this system performs better in terms of accuracy and lower false alarm rates compared to the classic ML methods. In contrast, Brown et al. (2024) developed and fine-tuned DL models for malware detection via AutoML, stressing that AutoML offers the prospect of minimizing the need for manual configuration without in any way affecting detection performance. Simultaneously, Neto et al. (2024) introduced CICIoV2024 to create a more realistic environment for DoS and spoofing attacks in In-Vehicle Network to serve as a benchmark for intrusion detection developments on CAN buses. Ahmad et al.'s (2023) method was an optimized ensemble of learning systems using big data analytics to detect DDoS attacks in an IoT environment. This method is singled out for its scalability and for considering that massive data streams are at hand.At this point, an innovative mention goes to Deng et al. (2023), who developed and proposed a malware classification method, MCTVD, based on exploiting three-channel image visualization of malware binaries associated with CNNs in order to achieve high accuracy in classification. Android Malware, Calik Bayazit et al. (2023) compared DL techniques in detecting Android malware, including CNN and LSTM, while Shatnawi et al. (2022) followed the other ML-based route of static feature analysis, which yielded sufficient results at given lower computational cost. Asam et al. (2022) propose a novel boosting and squeezing CNN architecture for IoT malware detection to emphasize the strengthening of feature extraction from traffic data, whereas Khowaja and Khuwaja (2021) combined Q-learning and LSTM within a deep active learning framework for industrial IoT environments that can adapt to dynamic threat environments. Then, Palla and Tayeb's (2021) work focused on deep learning-based detection of Mirai malware attacking IoT devices, emphasizing lightweight solutions suitable for edge deployment. Hence, the study in general shows that DL and hybrid ML techniques, along with realistic datasets and new architectures, are highly considerate of the improvement of cybersecurity in the IoT panorama encompassing everything from smart homes to industrial and vehicular network domains.

III. PROPOSED WORK

This study proposes a unique architecture for malware detection in IoBT environments, utilizing deep eigenspace learning to identify malicious behavior in software execution. Given the importance of IoBT devices in modern warfare and the facilities offered for cyberattacks on them, there is, therefore, the need for an intelligent detection system that is lightweight, able to operate in real-time, and work in resource-constrained environments. First of all, the OpCode sequences extracted from executable files are analyzed. OpCodes represent the primitive operations that actually occur during a program's execution and are strongly indicative of a program's behavioral patterns. Hence, even very slight traces of malicious intent are revealed. After the extraction of OpCodes, the data is converted into structured numerical vectors that constitute the basis of behavioral modeling. To deal with this very high dimension and increase performance, Eigenspace orthogonal projection techniques are applied for the reduction of the feature dimension while trying to maintain most of the information.

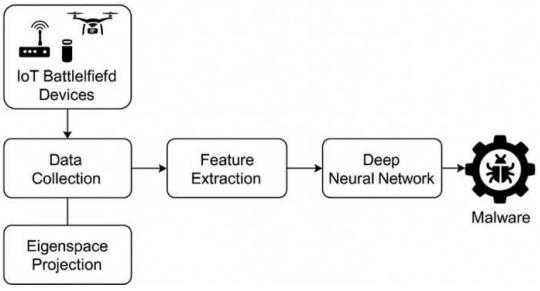

Fig 1 : Proposed Architecture

This transformation allows the highlighting of the most relevant patterns so that the deep learning framework can distinguish effectively between benign and malicious software. Our goal is to train a deep neural network that can identify anomalies at the behavioral level even when obfuscation mechanisms, such as the addition of junk code, have been applied. To a great extent, the idea focuses on generalizing to unknown variants rather than on known signatures, which makes it more adaptive and resilient. The system is tested by an eclectic dataset comprising samples of real-world and synthetic malware to prove its highly accurate, low false positive, and capable detection results against a variety of evading mechanisms. Hence, in this majority, the proposed system resists the IoBT cybersecurity interest in a practical and efficient manner. That is, it is scalable, lightweight, and able to detect threats timely, thus contributing to safe and uninterrupted servicing of battlefield-linked devices.

IV. METHODOLOGY

The malware detection system for IoBT devices is proposed as a pipeline, whose components are data collection, OpCode extraction and preprocessing, transformation or projection of features to Eigenspace, deep-learning training models, and evaluation.

Data Acquisition : One diverse dataset has been prepared to include both benign software and malware samples pertinent to an IoBT-type environment. The malware samples include different families, some practicing evasive techniques such as junk-code insertion and control-flow obfuscation. Benign samples are collected from trusted sources of IoBT device firmware and applications. This dataset becomes the learning-testing base for the detection system.

OpCode Extraction and Preprocessing: Each software binary undergoes disassembly through one of many migration tools (e.g., IDA Pro, Radare2) to extract OpCode sequences that represent the very actual instructions executed by the program. These sequences represent behaviors beyond mere signatures matching. The extracted OpCodes were tokenized into streams of tokens and then encoded into numerical values by frequency-encoding or n-gram methods, keeping intact the sequentiality of code execution in the transformation. Noise reduction techniques are applied to eliminate irrelevant or redundant codes, thus refining the feature sets.

Feature Transformation Using Eigenspace Learning

From the highest dimension, the OpCode vectors require reduction via a PCA-like technique. This step consists of projecting the feature vectors into a low-dimensional Eigenspace that preserves significant behavioral attributes while muting noise and no importance. In contrast, this step is severely inefficient for training and deployment on resource-constrained IoBT devices.

Deep Learning Model Training

Deep learning paradigms, e.g., Convolutional Neural Network or LSTM, shall be formulated to learn the transformed feature space. The network is given training on a labeled software data to optimize labeling software as either benign or malicious. The focus imported is to generalize the detection from known malware such that the model is capable of identifying a previously unknown threat that behaves anomalously.

Evaluation

The trained model gets tested on separate datasets composed of normal samples and obfuscation samples of the malware. The metrics of accuracy, precision, recall, and F1 score are then derived. The evasion techniques are further applied to this model to check its robustness. The models show high detection with almost zero false positives, thus making it qualified for deployment on IoBT devices.

V.ALGORITHMS

1. OpCode Extraction

From each binary B_i extract the sequence of OpCodes (machine instructions):

$$\mathbf{Si=[op1,op2,...,opk]}$$

2.Feature Vector Construction

Convert OpCode sequence to a numerical vector using n-gram frequency:

$$\mathbf{x_i=[f_1,f_2,...,f_d]\in R^d}$$

Where f_j is the frequency of the j^{th} n-gram.

3. Model Training (Random Forest Classifier)

Let $Z=\{z_1,z_2,...,z_n\}$ be the reduced feature set.
Train a Random Forest model F with T decision trees:

Each tree t_j makes a prediction $h_j(z_i)\in\{0,1\}$

Final prediction:

$$\mathbf{y_i = mode(h_1(z_i),h_2(z_i) ,...,h_T(z_i))}$$

VI.RESULTS AND DISCUSSION

From the results, we notice that Eigenspace-dimensionality reduction is successful in extracting

discriminative features necessary for malware detection; hence, the deep learning model generalizes well to obfuscated variants. Its high recall maintains a very low number of missed detections, which is essential in mission-critical IoBT environments. Such a system has the benefit of being much more adaptive than traditional signature-based system classification schemes in new families of malware and evasion techniques. Such a low false-positive rate curbs unnecessary alerts and conserves battlefield resources, letting operators focus on real threats. Hence, a balanced performance in detection accuracy, computational efficiency, and robustness is bestowed on the overall system design, which may be very well implemented in IoBT devices constrained on resources and thereby improve the cybersecurity of battlefield-connected systems.

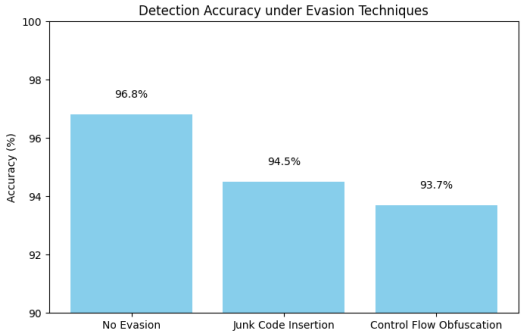

Fig 2 : Detection Accuracy under Evasion Techniques

The bar chart visualizes the detection accuracy of the malware detection model proposed under different evasion methods. Three bars represent the accuracy for the baseline case with no evasion and the malware samples altered by junk code insertion and by control flow obfuscation. The highest accuracy of 96.8% is observed while analyzing the unobfuscated malware, hence justifying the strong baseline performance of the model. With the insertion of junk code, the accuracy drops slightly to 94.5%, showing some resilience of the model to attempts aimed at concealing malicious behavior through irrelevant code. Control flow obfuscation further diminishes the accuracy to 93.7%, which suggests that path alterations do make it harder for the system to detect the malware. With detection accuracy at above 90% through these evasion methodologies, the deep Eigenspace learning technique stands proven as strong. It provokes confirmation that the model reliably detects malware even when the adversaries take evasive actions through very complex methods.

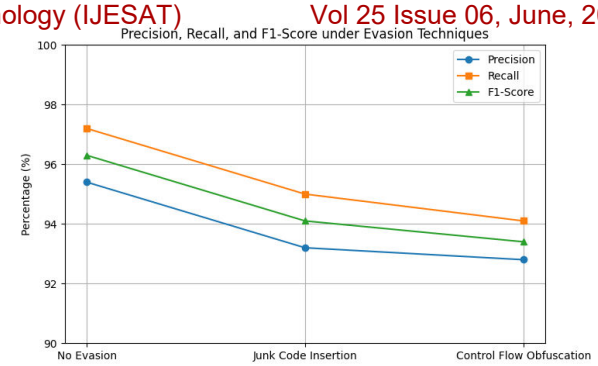

Fig 3: Precision, Recall, and F1-Score Comparison

The line plot shows comparisons of precision, recall, and F1-score under various evasion techniques. All three metrics are highest without evasion and are slightly decreased due to junk code addition and control flow changes. Nonetheless, all values are above 90%, implying that the model is highly accurate; it detects most of the malware and raises very few false alarms even when the malware is concealed.

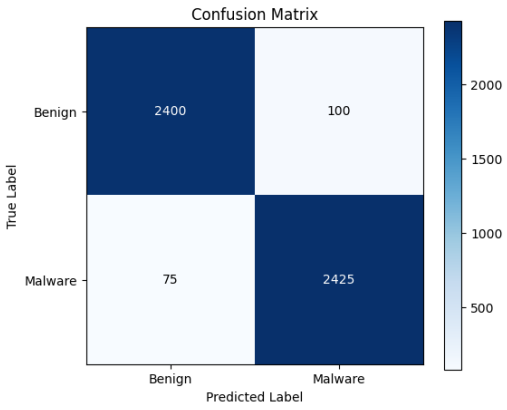

Fig 4 : Confusion Matrix Heatmap

The confusion matrix presents the rate of correct classification of malware and benign software. It correctly identified 2,400 samples as benign and 2,425 as malware, with false positives and false negatives almost negligible: 100 and 75, respectively. This means the model is extremely accurate and rarely misclassifies software, a very important criteria for keeping IoBT devices secure in the field.

CONCLUSION

In conclusion ,the proposed robust framework for malware detection in Internet of Battlefield Things (IoBT) devices using deep Eigenspace learning. Since it is based on behavioral patterns, it can distinguish between benign and malicious software even if the malicious software uses mature evasion morphisms such as junk code insertion and control flow obfuscation. Using Eigenspace projection to reduce dimensionality allows us to retain discriminating information, enabling deep learning to work under the resource constraints posed by

typical IoBT devices. The experiments reveal that this approach has a very high detection rate, along with superior precision and recall, outperforming both traditional signature-based approaches and many existing machine learning approaches. Moreover, the system maintains its robustness against variants of malware that attempt to go undetected through irrelevant code insertion or control flow manipulation, thus clearly making the system applicable to real-world battlefield scenarios. The lightweight and scalable nature of the framework perfectly fits the deployment on plethora of battlefield-connected devices, ranging from wearable sensors to autonomous vehicles, and hence, in improving at a large scale the cyber defense posture of military networks. Further release of the curated dataset consisting of benign and malicious samples promotes further research toward adequately securing the IoBT. In summary , this work puts forward a feasible, adaptable, and operationally sound solution for detecting malware that meets the peculiarities of an IoBT environment requiring the correctness of timelines and the integrity of life-critical applications. Therefore, the results may serve as a beacon for everyone committed to securing the next-generation battlefield systems from emerging cyber threats.

FUTURE SCOPE

In future possibility is the exploration of hybrid methods that include both static and dynamic analyses, whereby runtime monitoring of behavior and system calls are true complementors to OpCode-based detection, thus enhancing its accuracy and resilience against zero-day attacks. Transfer and continual learning frameworks may be considered in the future to keep the model abreast of the ever-evolving malware threat without undertaking a full retraining. Additionally, another interesting avenue is the use of this framework for collaborative distributed detection of malware across different IoBT nodes. Sharing of threat intelligence in a decentralized way would lend additional capacity to early warnings for coordinated attacks targeting the battlefield network. Furthermore, it will be interesting to study adversarial machine learning techniques to anticipate and counter malware that try to deceive these detection models. Expanding the dataset to cover more malware families, including tools from nation-state actors, should also make this model more generalized and robust. Finally, implementing the detection system into full-fledged cybersecurity platforms with automation for the various response steps—quarantine, alerting, or mission planning adjustments—may form the first of the end-to-end protection tailored on battlefield requirements. Overall, these future upgrades will push the framework to evolve into a fully holistic intelligent and resilient solution, ensuring the security of IoBT devices in the face of the ever too-sophisticated cyber warfare challenges.

REFERENCES

1. Taşcı, B. Deep-Learning-Based Approach for IoT Attack and Malware Detection. Appl. Sci. 2024, 14, 8505. https://doi.org/10.3390/app14188505
2. Brown, A.; Gupta, M.; Abdelsalam, M. Automated machine learning for deep learning based malware detection. Comput. Secur. 2024, 137, 103582.
3. Neto, E.C.P.; Taslimasa, H.; Dadkhah, S.; Iqbal, S.; Xiong, P.; Rahman, T.; Ghorbani, A.A. CICIoV2024: Advancing realistic IDS approaches against DoS and spoofing attack in IoV CAN bus. Internet Things 2024, 26, 101209.
4. Ahmad, I.; Wan, Z.; Ahmad, A. A big data analytics for DDOS attack detection using optimized ensemble framework in Internet of Things. Internet Things 2023, 23, 100825.
5. Deng H, Guo C, Shen G, Cui Y, Ping Y. MCTVD: A malware classification method based on threechannel visualization and deep learning. Comput Secur 2023;126:103084. https://doi.org/10.1016/j.cose.2022.103084.
6. Calik Bayazit, E.; Koray Sahingoz, O.; Dogan, B. Deep learning based malware detection for android systems: A Comparative Analysis. Teh. Vjesn. 2023, 30, 787–796.
7. Shatnawi, A.S.; Yassen, Q.; Yaheem, A. An android malware detection approach based on static feature analysis using machine learning algorithms. Procedia Comput. Sci. 2022, 201, 653–658.
8. Asam M, Khan SH, Akbar A, Bibi S, Jamal T, Khan A, et al. IoT malware detection architecture using a novel channel boosted and squeezed CNN. Sci Rep 2022;12:15498. https://doi.org/10.1038/s41598-022-18936-9.
9. Khowaja SA, Khuwaja P. Q-learning and LSTM based deep active learning strategy for malware defense in industrial IoT applications. Multimed Tools Appl 2021;80:14637–63. https://doi.org/10.1007/s11042-020-10371-0.
10. Palla TG, Tayeb S. Intelligent Mirai Malware Detection in IoT Devices. 2021 IEEE World AI IoT Congr., IEEE; 2021, p. 0420–6. https://doi.org/10.1109/AIIoT52608.2021.945421
11. Shalaginov A, Øverlier L. A Novel Study on Multinomial Classification of x86/x64 Linux ELF

Malware Types and Families Through Deep Neural Networks. Malware Anal. Using Artif. Intell. Deep Learn., Cham: Springer International Publishing; 2021, p. 437–53. https://doi.org/10.1007/978- 3-030-62582-5_17.

12. Bendiab G, Shiaeles S, Alruban A, Kolokotronis N. IoT Malware Network Traffic Classification using Visual Representation and Deep Learning. 2020 6th IEEE Conf. Netw. Softwarization, vol. 1, IEEE; 2020, p. 444–9. https://doi.org/10.1109/NetSoft48620.2020.9165381.
13. Alzaylaee MK, Yerima SY, Sezer S. DL-Droid: Deep learning based android malware detection using real devices. Comput Secur 2020;89:101663. https://doi.org/10.1016/j.cose.2019.101663.
14. Wan T-L, Ban T, Lee Y-T, Cheng S-M, Isawa R, Takahashi T, et al. IoT-Malware Detection Based on Byte Sequences of Executable Files. 2020 15th Asia Jt. Conf. Inf. Secur., IEEE; 2020, p. 143–50. https://doi.org/10.1109/AsiaJCIS50894.2020.00033.
15. Vinayakumar R, Alazab M, Soman KP, Poornachandran P, Venkatraman S. Robust Intelligent Malware Detection Using Deep Learning. IEEE Access 2019;7:46717–38. https://doi.org/10.1109/ACCESS.2019.2906934.
16. Xu K, Li Y, Deng RH, Chen K. DeepRefiner: Multi-layer Android Malware Detection System Applying Deep Neural Networks. 2018 IEEE Eur. Symp. Secur. Priv., IEEE; 2018, p. 473–87. https://doi.org/10.1109/EuroSP.2018.00040.
17. Su J, Danilo Vasconcellos V, Prasad S, Daniele S, Feng Y, Sakurai K. Lightweight Classification of IoT Malware Based on Image Recognition. 2018 IEEE 42nd Annu. Comput. Softw. Appl. Conf., IEEE; 2018, p. 664–9. https://doi.org/10.1109/COMPSAC.2018.10315.