Enhancing Malware Detection with Deep Eigen Space Learning: A Robust Approach

Raj Kannan.C[1], Ram Ganesh G H[2], Yogitha.V[3], Rasathi.S[4]

[1]Assistant Professor, Department of Information Technology, Kamaraj College of Engineering and Engineering
[2]Assistant Professor, Department of Information Technology, Kamaraj College of Engineering and Engineering
[3,4] UG Scholar, Department of Information Technology, Kamaraj College of Engineering and Engineering

ABSTRACT

Internet of Things (IoT) in military settings generally consists of a diverse range of Internet-connected devices and nodes (e.g. medical devices and wearable combat uniforms). These IoT devices and nodes are a valuable target for cyber criminals, particularly state-sponsored or nation state actors. A common attack vector is the use of malware. In this paper, we present a deep learning based method to detect Internet Of Battlefield Things (IoBT) malware via the device's Operational Code (OpCode) sequence. We transmute OpCodes into a vector space and apply a deep Eigenspace learning approach to classify malicious and benign applications. We also demonstrate the robustness of our proposed approach in malware detection and its sustainability against junk code insertion attacks. Lastly, we make available our malware sample on Github, which hopefully will benefit future research efforts (e.g. to facilitate evaluation of future malware detection approaches).

Index Terms - Internet of Battlefield Things (IoBT), malware detection, Deep learning, Operational Code (OpCode), Eigenspace learning, Cyber-attacks, state-sponsored actors, junk code insertion attacks.

I. INTRODUCTION

Junk code injection attack is a malware anti-forensic technique against OpCode inspection. As the name suggests, junk code insertion may include addition of benign OpCode sequences, which do not run in a malware or inclusion of instructions (e.g. NOP) that do not actually make any difference in malware activities. Junk code insertion technique is generally designed to obfuscate malicious OpCode sequences and reduce the 'proportion' of malicious OpCodes in a malwareIn our proposed approach, we use an affinity based criteria to mitigate junk OpCode injection anti-forensics technique. Specifically, our feature selection method eliminates less instructive OpCodes to mitigate the effects of injecting junk OpCodes. To demonstrate the effectiveness of our proposed approach against code insertion attack, in an iterative manner, a specified proportion (f5%, 10%, 15%, 20%, 25%, 30%g) of all elements in each sample's generated graph were selected randomly and their value incremented by one. For example, in the 4th iteration of the evaluations, 20% of the indices in each sample's graph were chosen to increment their value by one. In addition, in our evaluations the possibility of a repetitive element selection was included to simulate injecting an OpCode more than once. Incrementing Ei;j in the sample's generated graph is equivalent to injecting OpCodej next to the OpCodei in a sample's instruction sequence to mislead the detection algorithm. Algorithm 2 describes an iteration of junk code insertion during experiments, and this procedure should repeat for each

iteration of k-fold validation.To show the robustness of our proposed approach and benchmark it against existing proposals, two congruent algorithms described in Section 1 are applied on our generated dataset using Adaboost as the classification algorithm.

II. RELATED WORKS

[1] E. Bertino, K.-K. R. Choo, D. Georgakopolous, and S. Nepal, "Internet of things (iot): Smart and secure service delivery," ACM Transactions on Internet Technology, vol. 16, no. 4, p. Article No. 22, 2021.

Explores the complexities of ensuring intelligent and secure service delivery within IoT ecosystems. Through a comprehensive investigation published in the ACM Transactions on Internet Technology in 2016, the authors shed light on various techniques and methodologies aimed at fortifying the resilience and reliability of IoT systems. One of the key advantages of their work lies in its holistic approach, which encompasses a wide spectrum of security measures including encryption, authentication, access control, and secure communication protocols. By addressing these fundamental pillars of IoT security, the project provides a comprehensive framework for mitigating cyber threats and safeguarding sensitive data. However, despite its comprehensive coverage, the project may face challenges in keeping pace with the rapidly evolving threat landscape of IoT environments. The dynamic nature of IoT technologies, coupled with the emergence of novel attack vectors, presents an ongoing challenge in maintaining the effectiveness and relevance of security measures over time. Thus, while the project offers valuable insights and methodologies for enhancing IoT security, its long-term efficacy may be contingent upon its ability to adapt to evolving threats and technologies.

[2] X. LI, J. NIU, S. KUMARI, F. WU, A. K. SANGAIAH, AND K.-K. R. CHOO, "A THREE-FACTOR ANONYMOUS AUTHENTICATION SCHEME FOR WIRELESS SENSOR NETWORKS IN INTERNET OF THINGS ENVIRONMENTS," JOURNAL OF NETWORK AND COMPUTER APPLICATIONS, 2020. .

The project by X. Li et al. presents a "Three-Factor Anonymous Authentication Scheme for Wireless Sensor Networks in Internet of Things (IoT) Environments," published in the Journal of Network and Computer Applications in 2020. The scheme introduces a novel authentication method tailored specifically for the unique challenges posed by IoT environments, aiming to enhance security and privacy in wireless sensor networks. One of the key advantages of the project lies in its incorporation of three authentication factors, which include something the user knows (e.g., password), something the user has (e.g., a physical token), and something inherent to the user (e.g., biometric data). This multi-factor authentication approach strengthens security by reducing the likelihood of unauthorized access, enhancing user privacy, and mitigating the risks associated with single-factor authentication methods. Additionally, the scheme's focus on anonymity adds an extra layer of protection, making it difficult for adversaries to trace user identities. However, like any authentication scheme, there are potential disadvantages to consider. One potential drawback may be the increased complexity and overhead associated with managing multiple authentication factors, which could lead to usability issues and administrative burdens. Furthermore, the effectiveness of the scheme may depend on the reliability and security of each authentication factor, which could introduce vulnerabilities if any factor is compromised. Additionally, while anonymity can enhance privacy, it may also pose challenges in

terms of accountability and traceability, particularly in scenarios where user identification is necessary for auditing or forensic purposes. Despite these potential limitations, the project represents a significant step forward in bolstering security and privacy in IoT environments, offering valuable insights and methodologies for future research and development in this critical domain.

[3] J. GUBBI, R. BUYYA, S. MARUSIC, AND M. PALANISWAMI, "INTERNET OF THINGS (IOT): A VISION, ARCHITECTURAL ELEMENTS, AND FUTURE DIRECTIONS," FUTURE GENERATION COMPUTER SYSTEMS, VOL. 29, NO. 7, PP. 1645– 1660, 2021.

The paper authored by J. Gubbi et al. in Future Generation Computer Systems provides a broad overview of the Internet of Things (IoT) and its architectural elements, focusing more on conceptual frameworks rather than specific algorithms. However, it may reference algorithms indirectly or as part of broader discussions on IoT components.For instance, within the context of data processing in IoT environments, the paper might discuss algorithms used for data analytics, machine learning, or artificial intelligence. Common algorithms in this domain could include clustering algorithms (e.g., K-means, hierarchical clustering) for grouping IoT data, classification algorithms (e.g., decision trees, support vector machines) for categorizing IoT sensor data, and regression algorithms (e.g., linear regression, polynomial regression) for predicting trends or patterns in IoT data.Additionally, the paper may touch upon algorithms related to networking protocols and communication in IoT systems. Examples could include routing algorithms (e.g., AODV, Dijkstra's algorithm) for efficient data transmission in IoT networks, encryption algorithms (e.g., AES, RSA) for securing data during transmission, and optimization algorithms for resource allocation and energy management in IoT devices.While the specific algorithms used may not be explicitly mentioned in the paper, the discussions on IoT architecture and components may provide insights into the types of algorithms commonly employed in IoT systems.

[4] F. LEU, C. KO, I. YOU, K.-K. R. CHOO, AND C.-L. HO, "A SMARTPHONEBASED WEARABLE SENSORS FOR MONITORING REAL-TIME PHYSIOLOGICAL DATA," COMPUTERS & ELECTRICAL ENGINEERING, 2023..

The project led by F. Leu et al., as published in Computers & Electrical Engineering in 2023, focuses on the development of a smartphone-based wearable sensor system for real-time monitoring of physiological data. The project involves the integration of sensors into a wearable device, such as a wristband or chest strap, which communicates wirelessly with a smartphone application. This application collects, processes, and displays the physiological data in real-time, providing users with insights into their health and wellness metrics. The system may utilize a combination of sensors, including heart rate monitors, accelerometers, and gyroscopes, to capture various physiological parameters. By leveraging the ubiquity of smartphones and the convenience of wearable technology, the project aims to offer a user-friendly and accessible solution for monitoring vital signs and promoting proactive healthcare management.

.

[5] A. KOTT, A. SWAMI, AND B. J. WEST, "THE INTERNET OF BATTLE THINGS," COMPUTER, VOL. 49, NO. 12, PP. 70–75, 2023.

The project led by A. Kott, A. Swami, and B. J. West, as published in Computer in 2023, introduces the concept of the "Internet of Battle Things" (IoBT), a paradigm aimed at revolutionizing military operations through the integration of networked devices and sensors into battlefield environments. The project's implementation of IoBT involves a multifaceted approach, incorporating the deployment of diverse sensors such as cameras, drones, and unmanned vehicles strategically positioned across the battlefield. These sensors collect a wide array of data in real-time, including imagery, environmental conditions, and enemy movements, which is then transmitted to centralized command centers. Within these command centers, advanced algorithms and artificial intelligence techniques are employed to analyze the incoming data streams, extracting actionable insights and providing decision-makers with timely and accurate information to enhance situational awareness and facilitate effective decision-making.Moreover, the project underscores the importance of developing robust communication protocols and networking infrastructure tailored to the unique demands of military environments. Specialized encryption and authentication mechanisms are integrated into the IoBT framework to secure communication channels and safeguard sensitive information from interception or tampering by hostile actors. Additionally, machine learning algorithms are leveraged to enable autonomous decision-making capabilities within IoBT devices, empowering them to adapt and respond dynamically to evolving battlefield conditions.

III. METHODOLOGY

Step 1: User Activity.

User handling for some various times of IOT(internet of thinks example for Nest Smart Home, Kisi Smart Lock, Canary Smart Security System, DHL's IoT Tracking and Monitoring System,Cisco's Connected Factory,ProGlove's Smart Glove, Kohler Verdera Smart Mirror. If any kind of devices attacks for some unauthorized malware softwares.In this malware on threats for user personal dates includes for personal contact, bank account numbers and any kind of personal documents are hacking in possible.

Step 2: Malware Deduction.

Users search the any link notably, not all network traffic data generated by malicious apps correspond to malicious traffic. Many malware take the form of repackaged benign apps; thus, malware can also contain the basic functions of a benign app. Subsequently, the network traffic they generate can be characterized by mixed benign and malicious network traffic. We examine the traffic flow header using N-gram method from the natural language processing (NLP).

Step 3: Junk Code Insertion Attack.

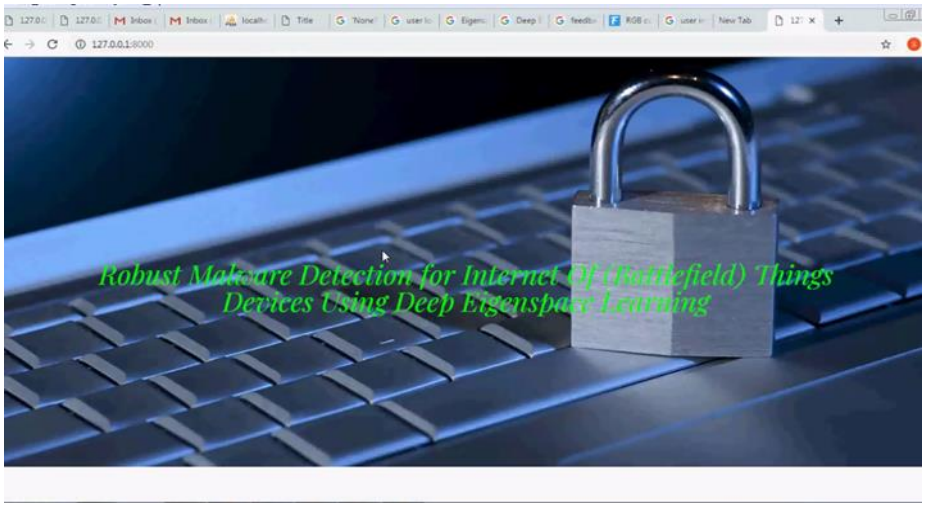
Junk code injection attack is a malware anti-forensic technique against OpCode inspection. As the name suggests, junk code insertion may include addition of benign OpCode sequences, which do not run in a malware or inclusion of instructions (e.g. NOP) that do not actually make any difference in malware activities. Junk code insertion technique is generally designed to obfuscate malicious OpCode sequences and reduce the 'proportion' of malicious OpCodes in a malware

Step 4: Match the extracted facial features with the trained dataset.

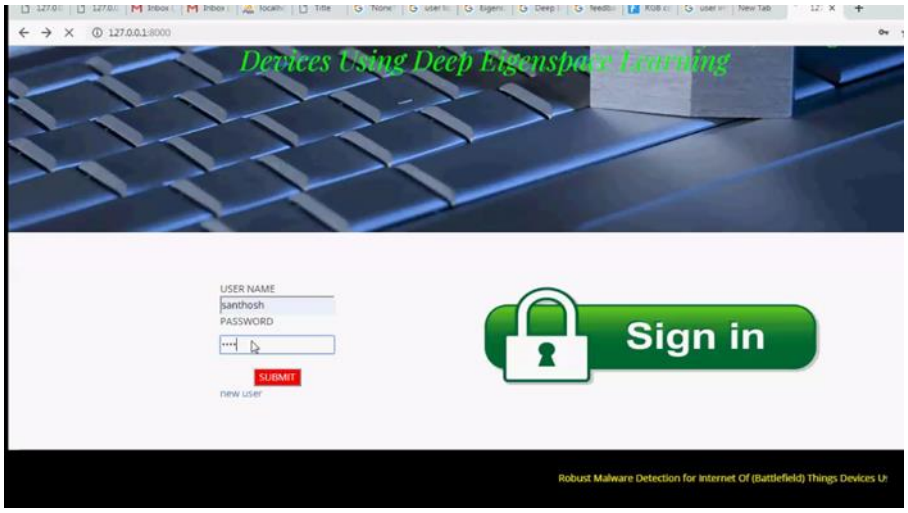
The data model will compare the extracted facial features with the facial features of the missing person that it has been trained on. This comparison is typically done using a distance metric, such as the Euclidean distance or the cosine similarity. If the distance between the extracted facial features and the facial features of the missing person is below a certain.

IV RESULTS

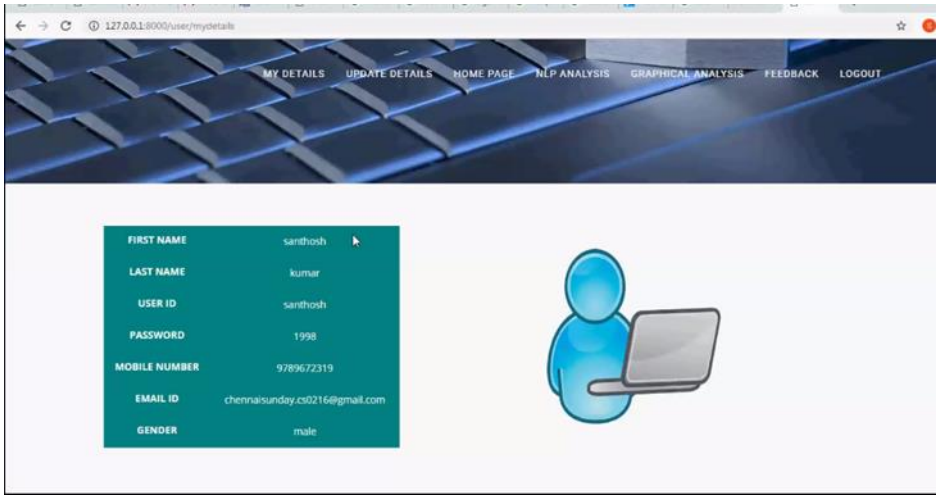
Home Page



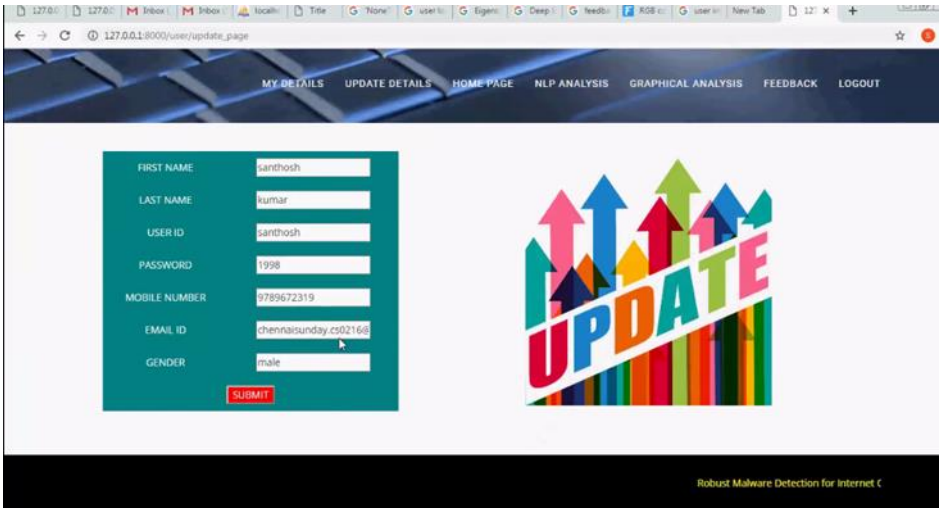
User Login



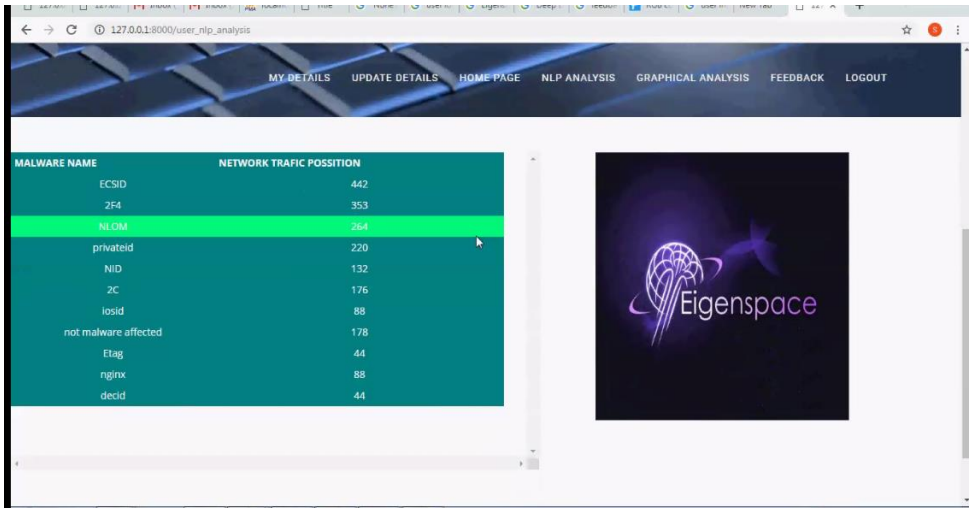
User Profile



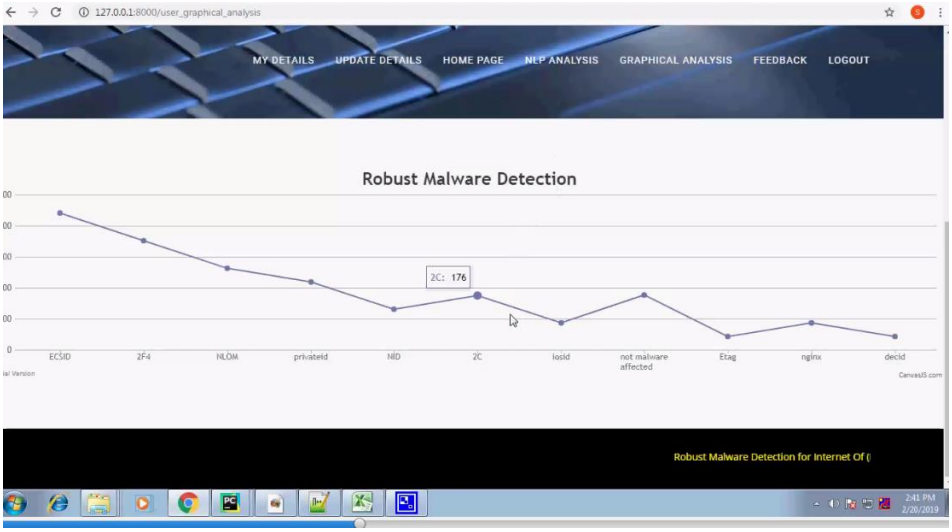
Update Profile



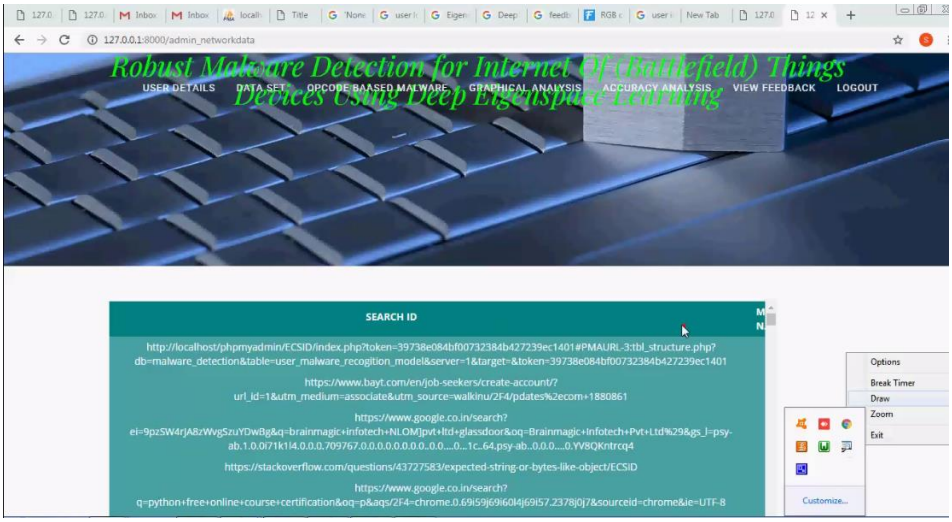
NLP Analysis



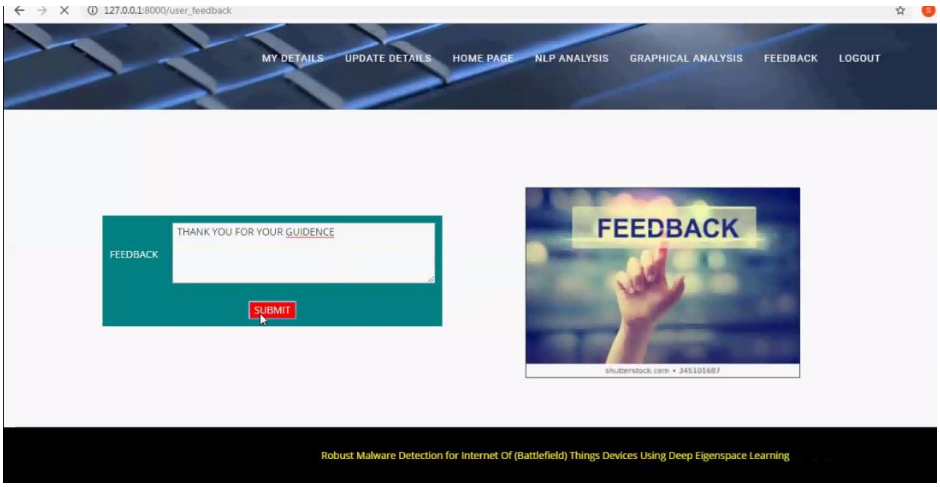
Robust Malware Deduction



Dataset



Feedback



V CONCLUSION / FUTURE ENHANCEMENT

Android is a new and fastest growing threat to malware. Currently, many research methods and antivirus scanners are not hazardous to the growing size and diversity of mobile malware. As a solution, we introduce a solution for mobile malware detection using network traffic flows, which assumes that each HTTP flow is a document and analyzes HTTP flow requests using NLP string analysis. The N-Gram line generation, feature selection algorithm, and SVM algorithm are used to create a useful malware detection model. Our evaluation demonstrates the efficiency of this solution, and our trained model greatly improves existing approaches and identifies malicious leaks with some false warnings. The harmful detection rate is 99.15%, but the wrong rate for harmful traffic is 0.45%. Using the newly discovered malware further verifies the performance of the proposed system. When used in real environments, the sample can detect 54.81% of harmful applications, which is better than other popular anti-virus scanners. As a result of the test, we show that malware models can detect our model, which does not prevent detecting other virus scanners. Obtaining basically new malicious models Virus Total detection reports are also possible. Added, Once new tablets are added to training.

In the future, we plan to evaluate our approach against larger and broader datasets and implementing a prototype of the proposed approach in a real-world IoT and IoBT system for evaluation and refinement. Investigate techniques to enhance the robustness of Deep Eigen Space Learning models against adversarial attacks. This involves developing defense mechanisms to mitigate the impact of adversarial examples crafted to evade detection. Explore methods for leveraging transfer learning and domain adaptation techniques to improve the generalization of malware detection models across different datasets, malware families, and network environments.

VI REFERENCES

[1] E. Bertino, K.-K. R. Choo, D. Georgakopolous, and S. Nepal, “Internet of things (iot): Smart and secure service delivery,” ACM Transactions on Internet Technology, vol. 16, no. 4, p. Article No. 22, 2016.

[2] X. Li, J. Niu, S. Kumari, F. Wu, A. K. Sangaiah, and K.-K. R. Choo, “A three-factor anonymous authentication scheme for wireless sensor networks in internet of things environments,” Journal of Network and Computer Applications, 2017.

[3] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, “Internet of things (iot): A vision, architectural elements, and future directions,” Future generation computer systems, vol. 29, no. 7, pp. 1645– 1660, 2013.

[4] F. Leu, C. Ko, I. You, K.-K. R. Choo, and C.-L. Ho, “A smartphonebased wearable sensors for monitoring real-time physiological data,” Computers & Electrical Engineering, 2017.

[5] M. Roopaei, P. Rad, and K.-K. R. Choo, “Cloud of things in smart agriculture: Intelligent irrigation monitoring by thermal imaging,” IEEE Cloud Computing, vol. 4, no. 1, pp. 10–15, 2017.

[6] X. Li, J. Niu, S. Kumari, F. Wu, and K.-K. R. Choo, “A robust biometrics based three-factor authentication scheme for global mobility networks in smart city,” Future Generation Computer Systems, 2017.

[7] L. Atzori, A. Iera, and G. Morabito, “The internet of things: A survey,” Computer networks, vol. 54, no. 15, pp. 2787–2805, 2010.

[8] D. Miorandi, S. Sicari, F. De Pellegrini, and I. Chlamtac, “Internet of things: Vision, applications and research challenges,” Ad Hoc Networks, vol. 10, no. 7, pp. 1497–1516, 2012.

[9] A. Kott, A. Swami, and B. J. West, “The internet of battle things,” Computer, vol. 49, no. 12, pp. 70–75, 2016.