IJIRSET
International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

e-ISSN: 2319-8753 | p-ISSN: 2347-6710

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 12, December 2024

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.524

9940 572 462 | 6381 907 438 | ijirset@gmail.com | www.ijirset.com

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.524

|www.ijirset.com |A Monthly, Peer Reviewed & Referred Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 13, Issue 12, December 2024

|DOI: 10.15680/IJIRSET.2024.1312019|

Using Data Visualization and Fingerprinting to Improve Cyber Defense Systems with AI

Dhanush H G, Chethan T Y, Shwetha S

U.G. Student, Department of Computer Engineering, Channabasaveshwara Institute of Technology, Tumkur, Karnataka, India

U.G. Student, Department of Computer Engineering, Channabasaveshwara Institute of Technology, Tumkur, Karnataka, India

Associate Professor, Department of Computer Engineering, Channabasaveshwara Institute of Technology, Tumkur, Karnataka, India

ABSTRACT: With the applications like MalGAN, such cyberattacks enhanced with artificial intelligence (AI) in a broad way across cyber-defense lifecycles successfully take the vulnerabilities of systems at advantage, which are many as these are evading defenses nowadays. Therefore, this methodology proposed a new method which presents the approach of data fingerprinting and visualization for AI-Enhanced Cyber-Defense Systems (AIECDS) for efficiency in detection. AIECDS approach is built combining dynamic reinforcement learning, feature extraction and visualization with Hilbert curves and tornado graphs, real-time data processing. Experimenting using the UNSW-15 dataset shows that even using very small sample sizes it's possible to differentiate malicious sessions from benign ones, meaning a large advancement in AI-driven solutions to cybersecurity, by being more adept at identifying complex threats with simplified machine learning models.

I. INTRODUCTION

Protecting cyber assets from more complex cyberattacks has emerged as an essential concern in the twenty-first century. Cyber risks have surged because of events such as the COVID-19 pandemic and geopolitical crises like the Russian invasion of Ukraine. These attackers exploit weaknesses by evading conventional defenses and using sophisticated tools to attack vital systems. Many cyberattacks that are reinforced with AI and utilize tools like GANs are also extremely problematic. These technologies make it possible for attackers to perform extremely complex attacks, which are hard to trace and to stop.

Because of their dependency on outdated datasets, simulated environments, and limited adaptability to new threats, existing cybersecurity solutions are often inadequate. To be effective, defenses need real-world data and real-time flexibility. One potential cure is visual data representation, which makes complex datasets more efficient and enhances the performance of AI learning.

To overcome these issues, this paper introduces the methodology of AIECDS, which is a data fingerprinting and visualization technique. This technique enhances the efficiency of cybersecurity systems and allows easier detection of harmful patterns by encoding actual network data into visual fingerprints. The UNSW-15 dataset illustrates the potential of this method in enhancing threat identification and mitigating AI-driven attacks.

II. LITERATURE SURVEY

- Sophisticated machine learning (ML) and artificial intelligence (AI)-enhanced solutions have been prompted by the increasing sophistication of cyberattacks. MalGAN and DeepLocker are examples of AI-powered solutions that have exploited the entire lifecycle of cyber-defense, from reconnaissance to malicious execution. These technologies use adversarial tactics and hidden payloads to bypass traditional defenses, thus underlining the need for countermeasures at every stage of the lifecycle, especially during reconnaissance.
- However, there are some significant hurdles that come in the way of today's ML-based cybersecurity solutions. Many use outdated, laboratory-generated datasets that do not represent real-world scenarios. Legacy datasets

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.524

|www.ijirset.com |A Monthly, Peer Reviewed & Referred Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 13, Issue 12, December 2024

|DOI: 10.15680/IJIRSET.2024.1312019|

dominate research, which leads to very low detection rates in real-time scenarios. Moreover, adversarial attacks employ GANs, which exploit weaknesses in anomaly detection systems, or poison the input data to attack machine learning models.

- New advancements in the fields of reinforcement learning and deep learning promise hope for conquering these bottlenecks. For example, the idea of using deep reinforcement learning (DRL) together with distributed models has defeated these threats in the face of DDoS attacks. Moreover, by transiting huge volumes of multimodal data into easily understandable patterns, the complex datasets get streamlined through the visualization process in using tornado graphs and Hilbert curves. In this way, efficiency in the detection is greatly enhanced for threats associated with sparse sample sizes.
- Although much progress has been achieved, modern malware detection systems (MDS) and intrusion detection systems (IDS) still cannot successfully work in dynamic situations or identify unknown threats. The revolutionary technique that can be used to fill the gap is said to be fingerprinting data and visualization, which helps reduce model complexity by encoding real-world patterns into visual representations, thereby improving threat classification.

III. METHODOLOGY

- The suggested methodology of AIECDS, which combines data visualization with fingerprinting, to fight AI-driven cyber threats includes five main stages:
- Real-Time Data Capture: Network packets are gathered using Packet Capture (PCAP) technology to ensure relevance in the real world.
- IP addresses, ports, protocols, among other essential features of interest, are retrieved and temporarily retained for processing. Such a process is termed feature extraction and buffering.
- Data preparation involves tagging and enhancing the extracted features with information to create an organized dataset in fingerprinting.
- Data Fingerprinting and Visualization: Fingerprinting simplifies multimodal danger patterns by encoding real-world data into visual representations using Hilbert curves and tornado graphs.
- Threat Detection: To classify threats, maximize detection, and adapt to hostile attacks, a reinforcement learning model analyzes fingerprints.

IV. EXPERIMENTAL RESULTS

The UNSW-15 dataset, which was comprised of 10,240 network sessions with both benign and malicious activity, was verified using the suggested AIECDS approach. A methodology was followed to produce visual fingerprints that could distinguish benign from malicious occurrences by tracking patterns in protocol conversation and sent data.

1. Clustering malicious fingerprints

Malicious fingerprints were grouped using k-means clustering according to similarities within threat categories. The method identified eight out of nine dangerous threat types in the dataset and exhibited distinct visual patterns that distinguish malicious from benign network sessions.

2. Analysis of Transmitted Data

Fingerprint grids were encoded using a 128x128 Hilbert curve that showed particular patterns for every session. Benign and malicious fingerprints were compared with the help of Frobenius distance. It revealed significant variance in the pattern of data being transmitted. However, fingerprints that are associated with unknown threats, even reconnaissance malware and shellcode, have discernible patterns.

3. Protocol Discourse Analysis

The packet exchanges and communication sequences were recorded in the protocol discourse section. Patterns found throughout the setup, transfer, and teardown stages highlighted the distinctions between benign and malicious fingerprints. Even reconnaissance malware had unique packet interaction sequences, albeit with low detection percentages using conventional techniques.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

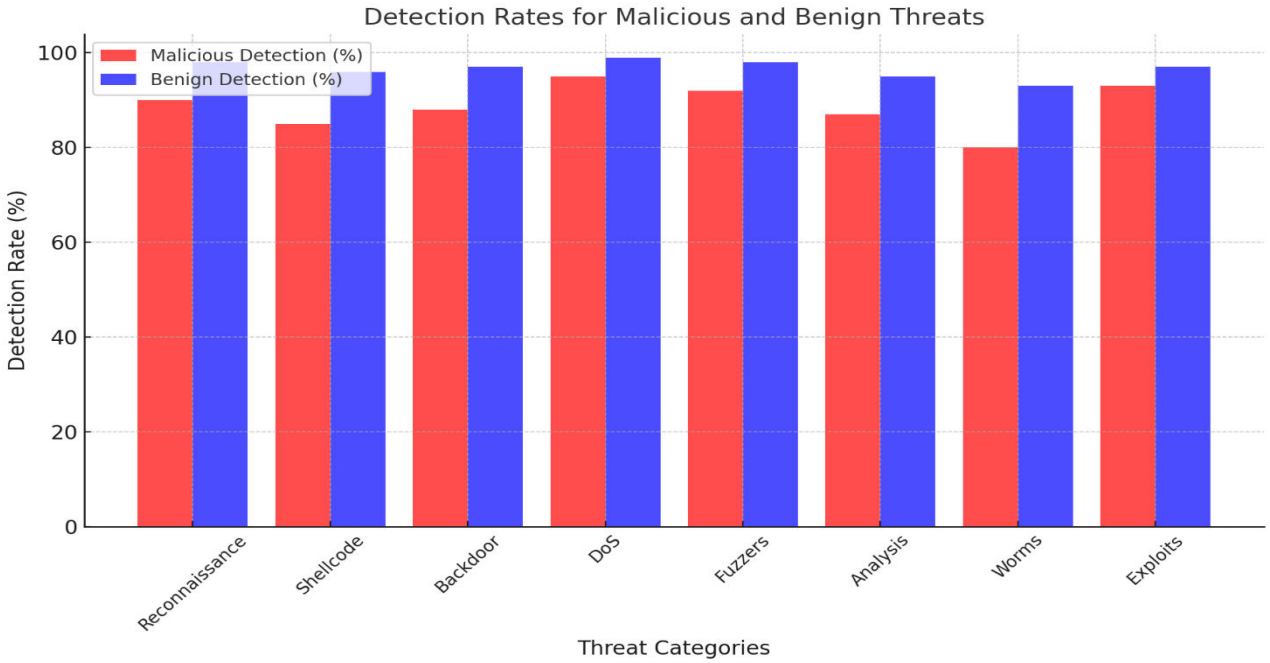
Impact Factor 8.524

|www.ijirset.com |A Monthly, Peer Reviewed & Referred Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 13, Issue 12, December 2024

|DOI: 10.15680/IJIRSET.2024.1312019|

Fig. 1 Model Performance



V. CONCLUSION

The proposed paper introduces the AIECDS technique that relies on data fingerprinting and visualization for improved cybersecurity against AI-driven threats. The concept simplifies complicated data patterns and improves the detection of threats through methods of encoding real-world network data as visual fingerprints, for instance, tornado graphs and Hilbert curves. Even for threats with scant sample data, AIECDS proved very effective in distinguishing between malicious and benign network sessions when crossvalidated with the UNSW-15 dataset. The strategy addressed important issues such as the need for real-time adaptability and the limitations of traditional databases. Moreover, visual fingerprints enhanced detection accuracy, reduced model complexity, and built robust decision limits. This approach offers a good foundation for developing AI-powered cybersecurity solutions, especially in terms of countering new and hostile cyberthreats.

REFERENCES

[1]. Aslan, Ö.A. and Samet, R., 2020. A comprehensive review on malware detection approaches. IEEE access, 8, pp.6249-6271.
[2]. Malialis, K., 2014. Distributed reinforcement learning for network intrusion response (Doctoral dissertation, University of York).
[3]. Duddu, V., 2018. A survey of adversarial machine learning in cyber warfare. Defense Science Journal, 68(4), p.356.
[4]. Klopper, C. and Eloff, J., 2023, February. Fingerprinting Network Sessions for the Discovery of Cyber Threats. In International Conference on Cyber Warfare and Security (18(1), pp. 171-180).
[5]. Heaney, C., Li, Y., Matar, O., Pain, 2024. Applying Convolutional Neural Networks to data on unstructured meshes with space-filling curves. Neural Networks, 175: 106198 (2024)
[6]. Keim, D.A., 1996. Pixel-oriented database visualizations. ACM Sigmod Record, 25(4), pp.35-39.
[7]. Yi , L., Xiaochong, T., Dali, W., Chunping, Q., He, L. and Youwei , Z., 2023. W-Hilbert: A W-shaped Hilbert curve and coding method for multiscale geospatial data index. International Journal of Applied Earth Observation and Geoinformation, 118, p.103298
[8]. Eschenbach, T.G., 2006. Constructing tornado diagrams with spreadsheets. The Engineering Economist, 51(2), pp.195-204.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi cross ref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com

www.ijirset.com

Scan to save the contact details