Adversarial Attacks and Defenses Using Machine Learning for Cybersecurity in Corporates

1st R. Gopinath
Department of Computer Science and Engineering
K S.Rangasamy College of Technology
Tiruchgode, Tamil Nadu, India
gopinath@ksrct.ac.in

2nd C. Sathiyamoorthy
Department of Commerce
Saveetha College of Liberal Arts and Sciences
Chennai, Tamil Nadu, India
rcsathiya81@gmail.com

3rd D. Sugumaran
Department of Information Technology
Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology
Chennai, Tamil Nadu, India
Sugumaran_dhanda@rediffmail.com

4th K.S. Giriprasath
Department of Computer Science and Engineering
PSG Institute of Technology and Applied Research
Coimbatore, Tamil Nadu, India
ksgiriprasath@gmail.com

5th Neha Tripathi
Graphic Era Deemed to be University
Dehradun, Uttarakhand, India
nehagarg.february@gmail.com

6th Mohammad Arif
School of Computer Science & Engineering
Vellore Institute of Technology
Vellore, Tamil Nadu , India
arif_mohd2k@yahoo.com

2024 International Conference on Distributed Systems, Computer Networks and Cybersecurity (ICDSCNC) | 979-8-3503-7544-2/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICDSCNC62492.2024.10939963

***Abstract*—This article suggests a novel method for protecting corporate cybersecurity systems from malevolent attacks, based on Capsule Networks (CapsNets). The enhancement of hierarchical feature learning by Capital Networks is a critical component of its capacity to differentiate between authentic and fraudulent data. Robust optimization techniques and adversarial training are implemented to develop a model. The training seeks to be more resilient and beneficial in a larger environment by introducing perturbations one capsule at a time. CapsNets executed an effective operation, achieving 95% accuracy and 97% precision. In terms of managing adversarial assaults, CapsNets outperform baseline models greatly. The proposed approach exhibits potential as an improved cybersecurity defense method, as a result of its exceptional resilience and precision. This study demonstrates the efficacy of CapsNets in improving cybersecurity and also offers a glimpse into the adversarial defenses used in enterprise machine learning applications.**

Keywords—capsule networks, adversarial attacks, cybersecurity, machine learning, resilient defense.

I. Introduction

Machine learning plays a crucial role in enhancing the cybersecurity measures of contemporary organizations [1]. As a result of the constant evolution of cyber threats, there is a significant demand for innovative solutions [2]. This work presents an innovative defense paradigm that strengthens cybersecurity measures against malicious intrusions [3]. This is accomplished through the implementation of a novel deep learning framework known as Capsule Networks (CapsNets) [4]. The increasing complexity of adversarial attacks requires a transition to more robust models, despite the evident utility of machine learning in the field of cybersecurity [5]. CapsNets provide an innovative approach to feature learning, vis-à-vis more traditional architectures such as convolutional neural networks (CNNs) [4]. Enhancing discriminatory capability and comprehending hierarchical connections among characteristics render it an attractive contender for tackling evolving cybersecurity challenges [6]. Consciously manipulating data to deceive machine learning algorithms constitutes an adversarial attack, which has the potential to cause severe damage to cybersecurity systems. CapsNets' comprehensive hierarchical feature learning capabilities offer a potentially effective strategy for mitigating the impacts of such attacks. By capitalizing on the interrelationships among features and the understanding of spatial hierarchies, CapsNets enhance the durability and interpretability of models. To augment generalization and reduce overfitting, robust optimization techniques are employed to fine-tune the model. Using a variety of techniques, the CapsNet-based model can be trained to perform exceptionally well on pure data and to withstand detrimental perturbations. The flexibility of the proposed model is of utmost importance when one takes into account the extensive range of cybersecurity challenges that may arise in corporate settings [7]. It is of the utmost importance that the datasets utilized in our research accurately represent actual cybersecurity issues [8]. In order to enhance comprehension of the model's performance across diverse scenarios, it is advisable to incorporate both benign and malevolent data. The utilization of evaluation parameters such as recall, accuracy, precision, and F1-score facilitates a thorough evaluation of the model's efficacy in both benign and malicious environments [9]. Additional validation for the distinctive capabilities of CapsNets in the realm of cybersecurity is achieved through comparisons with foundational models. CapsNets boast remarkable capabilities in hierarchical feature learning, which establish them as superior alternatives to more traditional architectures [10]. In accordance with the findings of this study, capsule networks may significantly improve enterprise cybersecurity. The objective of this research is to make a scholarly contribution to the ongoing discourse surrounding novel countermeasures against adversarial attacks through the implementation of more resilient and interpretable machine learning models in cybersecurity. This is accomplished by capitalizing on the intrinsic benefits of CapsNets.

II. Literature Review

Verma et al. [11] analyzes the advantages and disadvantages of utilizing AML to examine a PE malware categorization system as one of its methods. This research paper evaluates a number of malware mitigation strategies by employing AML intelligence in a Black-Box attack. Moustafa et al. [12] assesses Explainable Artificial Intelligence (XAI) approaches utilized in IoT networks for the purpose of anomaly-based intrusion detection. Extensive methodology is employed. The research centers on techniques for enhancing XAI models, approaches for anomaly detection using historical data, and the interaction among these three domains and the Internet of Things. Furthermore, an examination of the

efficacy in discerning manipulated data from authentic data. The model's resilience is significantly improved through the careful incorporation of adversarial training and robust optimization, which showcases its capability to adjust to ever-changing cyber threats. The evaluation metrics and comparative analysis underscore the model's remarkable precision, recall, accuracy, and F1-score, which substantially surpass those of the baseline models. The statistical significance assessment reinforces the superiority of CapsNets. This research not only makes a valuable contribution to the field of adversarial machine learning in cybersecurity but also emphasizes the effectiveness of CapsNets as a robust solution for protecting corporate environments. Due to their remarkable resilience, CapsNets are regarded as a valuable resource in the continuous effort to develop machine learning-driven cybersecurity measures that are both effective and resilient.

REFERENCES

[1] K. He, D. D. Kim, and M. R. Asghar, "Adversarial Machine Learning for Network Intrusion Detection Systems: A Comprehensive Survey," IEEE Communications Surveys & Tutorials, vol. 25, no. 1, pp. 538-566, Firstquarter 2023.

[2] T. Khodadadi, M. Zamani, S. S. Chaeikar, Y. Javadianasl, M. Talebkhah, and M. Alizadeh, "Exploring the Benefits and Drawbacks of Machine Learning in Cybersecurity to Strengthen Cybersecurity Defences," in 2023 IEEE 30th Annual Software Technology Conference (STC), MD, USA, 2023, pp. 1-1.

[3] J. Singh, M. Wazid, A. K. Das, V. Chamola, and M. Guizani, "Machine learning security attacks and defense approaches for emerging cyber physical applications: A comprehensive survey," Computer Communications, vol. 192, 2022.

[4] G. Almashaqbeh, K. Kelley, A. Bishop, and J. Cappos, "CAPnet: A defense Against Cache Accounting Attacks on Content Distribution Networks," in 2019 IEEE Conference on Communications and Network Security (CNS), Washington, DC, USA, 2019, pp. 250-258.

[5] I. H. Sarker, M. H. Furhad, and R. Nowrozy, "AI-driven cybersecurity: An overview security intelligence modeling and research directions," SN Comput. Sci., vol. 2, no. 3, pp. 1-18, 2021.

[6] G. Apruzzese, M. Colajanni, L. Ferretti, A. Guido, and M. Marchetti, "On the effectiveness of machine and deep learning for cyber security," 2018 10th International Conference on Cyber Conflict (CyCon), Tallinn, Estonia, 2018, pp. 371-390.

[7] R. Colbaugh and K. Glass, "Predictive defense against evolving adversaries," in 2012 IEEE International Conference on Intelligence and Security Informatics, Washington, DC, USA, 2012, pp. 18-23.

[8] M. Girdhar, J. Hong, and J. Moore, "Cybersecurity of Autonomous Vehicles: A Systematic Literature Review of Adversarial Attacks and Defense Models," IEEE Open Journal of Vehicular Technology, vol. 4, pp. 417-437, 2023.

[9] N. Martins, J. M. Cruz, T. Cruz, and P. H. Abreu, "Adversarial Machine Learning Applied to Intrusion and Malware Scenarios: A Systematic Review," IEEE Access, vol. 8, pp. 35403-35419, 2020.

[10] J. Chen, X. Gao, R. Deng, Y. He, C. Fang, and P. Cheng, "Generating Adversarial Examples Against Machine Learning-Based Intrusion Detector in Industrial Control Systems," IEEE Transactions on Dependable and Secure Computing, vol. 19, no. 3, pp. 1810-1825, 1 May-June 2022.

[11] U. Verma, Y. Huang, C. Woodward, C. Schmugar, P. P. Ramagopal, and C. Fralick, "Attacking Malware Detection using Adversarial Machine Learning," in 2022 4th International Conference on Data Intelligence and Security (ICDIS), Shenzhen, China, 2022, pp. 40-49.

[12] N. Moustafa, N. Koroniotis, M. Keshk, A. Y. Zomaya, and Z. Tari, "Explainable Intrusion Detection for Cyber Defences in the Internet of Things: Opportunities and Solutions," IEEE Communications Surveys & Tutorials, vol. 25, no. 3, pp. 1775-1807, thirdquarter 2023.

[13] M. A. Al-Garadi, A. Mohamed, A. K. Al-Ali, X. Du, I. Ali, and M. Guizani, "A survey of machine and deep learning methods for Internet of Things (IoT) security," IEEE Commun. Surveys Tuts., vol. 22, no. 3, pp. 1646-1685, 3rd Quart. 2020.

[14] S. S. Satam, A. A. Patil, D. B. Narkhede, S. Singh, and N. Pulgam, "Zero-Day Attack Detection and Prevention," in 2023 7th International Conference On Computing, Communication, Control And Automation (ICCUBEA), Pune, India, 2023, pp. 1-6.

[15] S. Chen, M. Xue, L. Fan, S. Hao, L. Xu, H. Zhu, and B. Li, "Automated Poisoning Attacks and Defenses in Malware Detection Systems: An Adversarial Machine Learning Approach," Comput. Secur., vol. 73, pp. 326–344, 2018.