CS797V AI for Cybersecurity

Literature Review

10/02/2023

Submitted by,

Rebecca Soza

Abstract

I reviewed several research papers and reports on using effective deep learning strategies to protect satellite networks by detecting data security and privacy intrusion. The focus of this project is to develop an understanding of the findings found in the main article [1] and seek papers that are relevant to the case. There are four contributions for a hybrid intrusion detection system which all contain a base model of the machine learning concept Random Forest (RF) applied with the sequential forward feature selection (RF-SFS) and one of the four machine or deep learning methods: Random Forest (RF-SFS-RF), Artificial Neural Network (RF-SFS-ANN), Long-Short-Term Memory (RF-SFS-LSTM), and Gated Recurrent Unit (RF-SFS-GRU). These four resulted in different accuracies after using 10 selected features from the simulated satellite networks (STIN) or terrestrial networks (UNSW-NB15) dataset. For my project it is unclear as to which model I will be using, a decision will be made shortly.

I. Introduction

Large amounts of data can be transmitted and received from self-contained communication devices called satellites. In society, we have the ability to send a message from Earth to a satellite that is then retransmitted, giving us the capabilities to connect worldwide. The most common human interaction is driven by the Internet of things (IoT) [2]. IoT applications are connected to satellite networks such as cellphones, smartwatches, televisions, and more making it susceptible to security vulnerabilities. These smart devices interconnect causing one security breach to spread to another just by contact. The approach to protecting or defending satellite networks routes from the idea of having no privacy with our personal devices. Infiltration among common devices have been reportedly increased.

As hardware, software, and network topologies advance in satellite development a reflective spike in data security and privacy intrusion rises. Cyberattacks on satellite-terrestrial networks are protected using machine and deep learning algorithms to mitigate reliable intrusion detection system (IDS). Once an attack has been set the resources are more than likely to have been exhausted making it close to impossible for defending, therefore creating an efficient IDS can keep the network healthy.

The rest of this paper is organized as follows: literature review is covered in section 2, proposed model in section 3, and conclusion in section 4.

II. Literature Review

A majority of IDSs use machine or deep learning techniques to decrease the attacks on networks. I have examined three additional papers to refine the idea of using machine and deep learning models to protect the integrity of SAT networks. To support the main paper on its methods “Deep Learning Approach for Intelligent Intrusion Detection System” [3] contributes to a similar approach. The benefit of using a hybrid deep learning model to construct an IDS is to detect and classify unpredictable cyberattacks through network-based intrusion detection system (NIDS) and host-based intrusion detection system (HIDS).

These two previously mentioned papers use STIN and deep learning while another paper uses a different method approach called the federated learning (FL) [4] which sways me to believe that deep learning is the best choice for protecting future cyberattacks. The perspective is that with FL the goal is to minimize time cost rather than having better

accuracy, disregarding hyperparameters since they require fine tuning. Although in theory it sounds best to not worry about hyperparameters I believe this approach fails to consider other datasets that need to have an automatic model that can learn complex patterns. It is a powerful tool to use the learning rate to the model's advantage.

To tie all ideas together, identifying the right feature selection [5] delves in deep to the importance of enhancing IDSs specifically with machine or deep learning models. IDS can use significant features to identify anomalies and decrease execution time. Ideally these features can improve the performance making them scalable since the subset represents the entire dataset. Allowing IDS to detect cyberattacks through Anomaly-Based and Signature-Based detection. Anomaly-Based detection learns how the dataset network traffic flows behavior is and then compares new traffic to detect malicious attacks. Then Signature-Based detection detects attacks that are known. This combination is favored but can fail if there aren't enough signature attacks in the dataset.

III. Proposed Models

Most of the papers read preferred IDSs that have machine learning (ML) and deep learning (DL) models. These models abstract complexity by using a supervised approach for detecting network intrusions. The four main hybrid models use data from UNSW-NB15 and STIN. Both have a role in evaluating NIDS by distinguishing normal and anomalous traffic. The performance on discovering attacks depends on the dataset's application, UNSW-NB15 identifies terrestrial networks while STIN identifies satellite networks. For my project I will be using one of the four models and the STIN dataset since I will be simulating satellite networks. This dataset has two files that include 32 features with labels and 9 types of attacks. The SAT-IDS approaches will be described in detail. As stated in the abstract, the four models have a model base of RF-SFS along with one other combination. To recap, the RF-SFS model is used to choose the important features within a dataset.

3.1 RF-SFS: Random Forest

The Random Forest model is the only hybris machine learning approach used in this paper. In an RF there are decision trees that make up the model based on input values and make predictions based on their branches. This method is used for regression and classification to avoid overfitting by adding randomness in various decision trees. The accuracy produced was the highest at 90.5%.

IV. Conclusion

In conclusion, the reviewed literature and research papers have provided valuable insights into intrusion detection systems (IDS) for satellite networks, with a specific focus on deep learning strategies. The research explored the development of a hybrid intrusion detection system (IDS) comprising a base model using Random Forest (RF) with sequential forward feature selection (RF-SFS) and various deep learning methods such as Artificial Neural Network (ANN), Long-Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU).

The interconnected nature of Internet of Things (IoT) devices with satellite networks introduces security vulnerabilities, emphasizing the need for effective IDS. As technology advances in satellite development, the threat of data security and privacy intrusion escalates. Leveraging machine and deep learning algorithms becomes crucial for mitigating cyber threats and maintaining a healthy network.

Proposed models centered around four hybrid approaches, each combining the RF-SFS base model with a different deep learning method. The Random Forest (RF-SFS-RF) model demonstrated high accuracy at 90.5%, showcasing the effectiveness of the hybrid approach. The decision to choose one of these models for the project remained pending, with a commitment to using the STIN dataset. The research contributes to the growing body of knowledge on securing satellite networks through intelligent IDS. The findings emphasize the significance of selecting appropriate features and leveraging the strengths of both machine and deep learning models for robust intrusion detection. As this field continues to evolve, the proposed hybrid models offer promising avenues for enhancing the security posture of satellite-terrestrial networks.

V. References

[1] Azar, Ahmad Taher, et al. *Deep Learning Based Hybrid Intrusion Detection Systems to Protect Satellite Networks*, vol. 31, no. 4, 4 Sept. 2023, pp. 1573–7705. *82*, https://doi.org/10.1007/s10922-023-09767-8.

[2] Magdy, Mina & Matter, Ahmed & Hussin, Saleh & Hassan, Doaa & Elsaid, Shaimaa. (2023). *Anomaly-based intrusion detection system based on feature selection and majority voting*. Indonesian Journal of Electrical Engineering and Computer Science. 30. 1699. 10.11591/ijeecs.v30.i3.pp1699-1706.

[3] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. Al-Nemrat and S. Venkatraman, *Deep Learning Approach for Intelligent Intrusion Detection System*, in IEEE Access, vol. 7, pp. 41525-41550, 2019, doi: 10.1109/ACCESS.2019.2895334.

[4] K. Li, H. Zhou, Z. Tu, W. Wang and H. Zhang, *Distributed Network Intrusion Detection System in Satellite-Terrestrial Integrated Networks Using Federated Learning*, in IEEE Access, vol. 8, pp. 214852-214865, 2020, doi: 10.1109/ACCESS.2020.3041641.

[5] Kurniabudi, D. Stiawan, Darmawijoyo, M. Y. Bin Idris, A. M. Bamhdi and R. Budiarto, *CICIDS-2017 Dataset Feature Analysis With Information Gain for Anomaly Detection*, in IEEE Access, vol. 8, pp. 132911-132921, 2020, doi: 10.1109/ACCESS.2020.3009843.