Adaptive Ensemble Learning for Intrusion Detection Systems

Vincenzo Agate, Federico Concone*, Alessandra De Paola, Pierluca Ferraro, Salvatore Gaglio, Giuseppe Lo Re and Marco Morana

Università degli Studi di Palermo, Dipartimento di Ingegneria, Palermo, Italy

Abstract

For years, the European Commission has highlighted the need to invest in cybersecurity as a means of protecting institutions and citizens from the many threats in cyberspace. Attacks perpetrated through the network are extremely dangerous, also because their mitigation is complex, making it difficult to ensure an adequate level of security. One of the crucial elements in building an overall system of protection against network-based cyber attacks are Intrusion Detection Systems (IDSs), whose goal is to detect and identify such attacks and misuse of computer networks in a timely manner. Nowadays, the most effective IDSs are based on Machine Learning (ML) and are able to combine and analyze information from heterogeneous sources, such as network traffic, user activity patterns, and data extracted from system logs. However, these tools commonly exploit specific classifiers, whose performance is highly dependent on the attacks being considered, and are unable to generalize adequately enough to be applied in different contexts. The research laboratories of *Networking and Distributed Systems* and *Artificial Intelligence* at the University of Palermo are carrying out research activities in order to address these issues, with the main goal of designing a new generation of IDSs that, by dynamically and adaptively combining multiple classifiers, are able to overcome the limitations of state-of-the-art solutions.

Keywords

Cybersecurity, Artificial Intelligence, Intrusion Detection Systems

1. Introduction

Today, with the increasingly pervasive use of ICT technologies, cyber attacks pose a serious risk to the infrastructural, productive and economic aspects of our society. One of the most critical threats to today's hyperconnected world are attacks that come from the network. In fact, all social and productive realities are closely dependent on the ability to exchange data through the network. This dependence can be exploited by the malicious parties to gain unauthorized access to the resources of institutions and organizations. One of the most effective solutions to such attacks are Intrusion Detection Systems (IDSs), whose main goal is to timely detect and identify misuse of resources early enough to enable timely responses that stop any malicious behavior and ensure normal operation of systems.

Currently, the most promising approach to designing IDSs capable of dealing with the threats our systems will face in the near future is the adoption of Machine Learning (ML) and, more generally, Artificial Intelligence (AI) methods.

However, a thorough study of the literature shows that the adoption of machine learning methods to design IDSs involves several critical issues. One of the most noticeable concerns is that, due to the high heterogeneity of network traffic generated by different attacks, specific classifiers are characterized by performance that is highly dependent on the attacks considered. This means that there is no single universal ML approach that can detect any kind of attack in different scenarios. In addition, different classes of ML approaches have very different capabilities: for example, supervised methods can achieve excellent performance but are unable to handle unknown attacks, while unsupervised methods can detect anomalies and unknown attacks but generally achieve poor performance with already known intrusions [1].

The adoption of ensemble machine learning techniques, which leverage multiple machine learning algorithms, promises to be a very effective approach to achieve higher overall performance than single methods. However, in the current literature, the ensemble of classifiers is often designed through trial-and-error procedures, and there is no evidence that an approach suitable for a specific scenario can be general enough to be adopted in different scenarios.

Our research group, through scientific activities funded by various projects, seeks to contribute to this research area by designing new methodologies and adap-

Ital-IA 2024: 4th National Conference on Artificial Intelligence, organized by CINI, May 29-30, 2024, Naples, Italy

*Corresponding author.

vincenzo.agate@unipa.it (V. Agate); federico.concone@unipa.it (F. Concone); alessandra.depaola@unipa.it (A. De Paola); pierluca.ferraro@unipa.it (P. Ferraro); salvatore.gaglio@unipa.it (S. Gaglio); giuseppe.lore@unipa.it (G. Lo Re); marco.morana@unipa.it (M. Morana)

0000-0002-3326-8500 (V. Agate); 0000-0001-7638-3624 (F. Concone); 0000-0002-7340-1847 (A. De Paola); 0000-0003-1574-1111 (P. Ferraro); 0000-0002-8217-2230 (G. Lo Re); 0000-0002-5963-6236 (M. Morana)

tive solutions aiming to improve the robustness of existing approaches in the field of AI- and ML-based intrusion detection systems (IDS).

The following of this paper introduces the current state of the art of IDS and discusses the main limitations of current solutions, followed by a summary description of our research group's contribution. Finally, a description of the challenges and goals we intend to address in the near future is provided.

2. Related Work

In the dynamic domain of cybersecurity, the arms race between intrusion detection mechanisms and cyber-attack methodologies has accelerated, highlighting an urgent need for innovative detection techniques. Several IDSs have been proposed in the literature, exploiting both signature-based and anomaly-based approaches [2, 3]. The former are reliable in recognizing known attacks but are ineffective against those not previously seen. Conversely, the latter show a more flexible behavior and are better suited to detect constantly evolving attacks, especially by using Machine Learning (ML) techniques.

Nevertheless, the design of ML-based IDSs faces several challenges, such as the difficulty of ensuring fast responses when dealing with high-dimensional data, as in the case of network traffic, or providing consistently good performance for all types of intrusions. Moreover, in modern network environments with heterogeneous devices, the input data distributions are subject to unpredictable fluctuations over time. This phenomenon, referred to as concept drift, poses a significant challenge in the fields of machine learning and cybersecurity, as noted in [4]. One of the most promising directions to achieve overall good performance is the adoption of ensemble learning techniques [5], which exploit multiple ML algorithms to obtain better results than those of individual methods.

The IDS presented in [6], for instance, combines a two-stage meta classifier ensemble (i.e., rotation forest and bagging) with hybrid feature selection (particle swarm optimization, ant colony algorithm, and genetic algorithm) to better distinguish regular and anomalous traffic. However, such a solution is tailored on single attacks instances and not suitable for dealing with multi-class problems. The IDS introduced in [7] adopts an ensemble approach that combines decision trees, Random Forest, and Forest by Penalizing Attributes algorithms, and a voting technique to combine their probability distributions. Although the system achieves good performance with popular attacks, this drops in the case of rare ones. Multi-class intrusion detection is also addressed in [8], where an ensemble approach is designed to detect different attacks. Such IDS also exploits a hybrid feature selection method and a ranking technique that evaluates the ability of different base classifiers to detect different attacks. Results are promising, but only for a subset of the considered attack classes. The authors of [9] propose a model based on sustainable ensemble learning and on incremental learning. Such a system exploits multiclass regression models so that the ensemble is adapted to recognize different types of attacks; moreover, by means of an iterative update method the parameters and the decision results of the historical model are included into the training process of the final ensemble model.

The performances of the solutions described above, as well as many other existing ensemble frameworks, are severely limited as many different classes of attacks can occur. Moreover, the combination of multiple ML-based classifiers generally increases the computational load, thus limiting the IDS's ability to operate timely. This issue is particularly critical, given the need to promptly identify incoming threats and immediately apply appropriate countermeasures.

3. Research Contribution

In this perspective, a first contribution of our research unit is discussed in [10], where we introduced a system which addresses critical limitations in existing frameworks, achieving the right trade-off between number of recognized classes and prediction speed, in contrast to other multi-class IDSs in the literature.

In particular, we presented a multi-layered architecture for a behavior-based Intrusion Detection System that uses machine learning and ensemble learning techniques to distinguish between benign and malicious traffic and categorize detected malicious activities into one of nine possible attack classes. The architecture of the system is shown in Figure 1.

The experimental evaluation was performed on the CIC-IDS2017 public dataset, showing that the proposed IDS exhibits good performance in detecting all attack classes according to well-established metrics.

A key aspect of our proposed system is its two-layer architecture. To prevent the system from being overloaded with all the network traffic, and consequently to prevent delayed detections, traffic filtering is preliminarily performed in order to distinguish "normal" and "abnormal" traffic, ensuring that only potentially malicious traffic is advanced to the next stage for further analysis. This layer thus acts as a filter, improving the efficiency of the whole system. Accurate classification at this stage is crucial, as traffic deemed benign is not subject to subsequent scrutiny, highlighting the importance of minimizing false negatives to safeguard network integrity. For the design of the first layer, we decided to adopt a Decision Tree (DT), since experimental evaluation showed its better per-

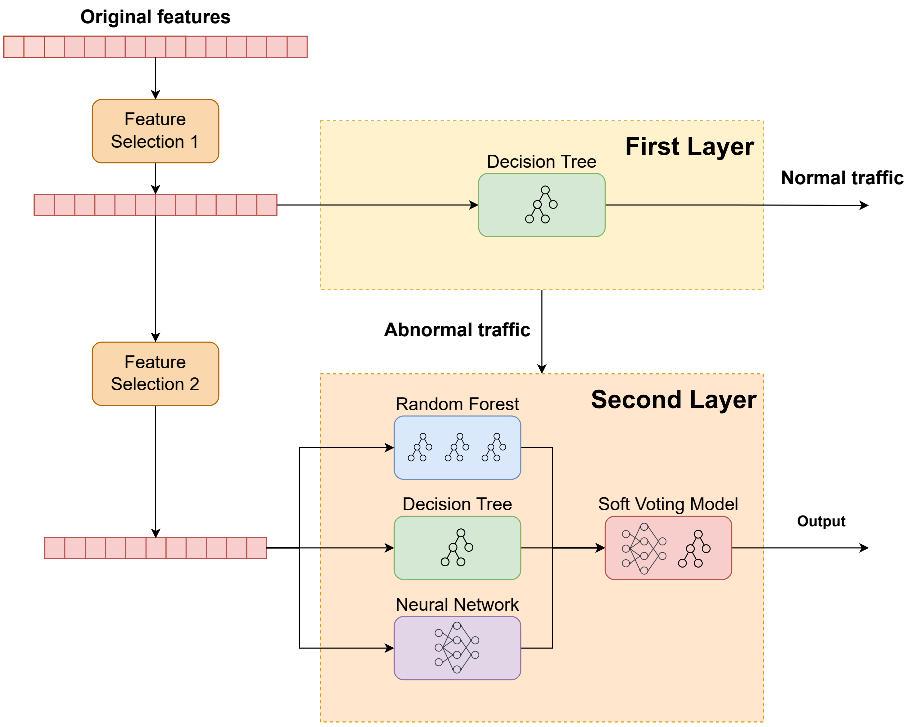

Figure 1: Architecture of the multi-layered IDS proposed in [10].

formance for binary classification, compared to Neural Networks, Random Forest, and Gaussian Naive Bayes.

In the second layer, a detailed analysis of malicious traffic is performed so thus the system generates alerts more accurately. These alerts provide network administrators with the information they need to quickly and effectively respond to threats [11], allowing them to neutralize ongoing attacks quickly and efficiently.

Our solution proposes the adoption of ensemble learning techniques, incorporating a combination of different learning models, such as Neural Networks (NNs), Random Forests (RFs), and additional DTs as weak learners.

The results of the predictions of the single models are aggregated using appropriate ensemble techniques that yield better classification performances than those of the single weak learners. Specifically, we adopt a weighted voting technique that assigns higher weights to the predictions of classifiers with low uncertainty in order to determine the ensemble's final verdict.

The adoption of this weighted voting strategy for aggregating classifier outputs, integrating the confidence values from neural network predictions with those of Decision Trees and Random Forests, notably improves the performance of the whole IDS. Finally, it is worth noticing that our system's architecture facilitates parallelization in the training and testing of weak learners, thereby enhancing efficiency in both training and prediction phases, a critical feature for IDS systems where timely threat detection is paramount.

This work is partially funded by the European Union - FESR o FSE, PON Ricerca e Innovazione 2014-2020 - DM 1062/2021.

4. Preliminary Evaluation

To conduct a preliminary evaluation of the proposed solution, the CIC-IDS2017 dataset was used [12]. This dataset perfectly fits the goals of our study as it includes various attacks encompassing SQL-Injection, Brute Force, XSS, DoS GoldenEye, DoS Hulk, DoS Slowhttptest, and DoS Slowloris. These attacks were grouped under two categories, i.e., Web and DOS Attacks, to streamline computation while maintaining detailed and accurate identification of malicious events.

All tests have been performed on off-the-shelf laptops equipped with Intel 3805U 1.9GHz CPU and 4GB RAM. Moreover, all the models that constitute the proposed IDS have been run 1000 times using different train and test sets at every execution.

The numerous tests performed on the system have demonstrated its reliability and accuracy in detecting malicious traffic, as well as its time efficiency. The IDS is able to recognize and identify 9 different types of attack in real-time, promptly alerting administrators to minimize serious consequences. In fact, on average, the system misses attacks in very small percentages (close to 1%), while it requires extremely low execution time for both the first and second levels: some slight difference is appreciated in dependence on the model used in the ensemble.

Besides the good performance achieved, numerous improvements are needed to address other important limitations, that are common to many IDSs in the literature.

First of all, the solutions proposed in the literature (as well as [10]) select the set of classifiers to be adopted through a trial-and-error process and lack a formalized methodology that can drive the design process in different scenarios. Moreover, many of the existing solutions have been designed ignoring the outbreak of unknown attacks. Such a "closed-world" approach makes IDSs unsuitable for recognizing special types of attacks known as "zero-day".

5. Challenges and Goals

The main goal of the research unit is the design and development of a novel class of IDSs based on the combination of several dynamically orchestrated classifiers (both supervised and unsupervised), with the aim of recognizing a large set of different threats, also detecting the occurrence of zero-day attacks.

Given the strong characterization of the many application scenarios in which IDSs are needed, the design of the system architecture will be guided by a formalized, rigorous, and replicable approach that can steer the realization of specific IDS instances. The goal is to design a scalable and modular architecture, capable of maintaining a low computing load while guaranteeing high detection performance and responsiveness, even in the presence of huge amounts of data.

The main challenge will be the definition of adaptive orchestration techniques, which will be crucial for the design of IDSs capable of dynamically adjusting their ensemble strategies based on the observed context. This will include the integration of both supervised and unsupervised learning approaches, allowing an adaptive response to emerging threats.

To reach this ambitious goal, the system will also have to address the phenomenon of concept drift, which is the continuous shift of the statistical distribution of network data over time. This poses a big challenge for current IDSs, often necessitating manual retraining of their machine learning models. Indeed, ignoring the phenomenon of concept drift, like many current IDSs do, inevitably lead to performance degradation over time.

Our future approach will try to overcome these challenges by orchestrating supervised and unsupervised systems to exploit the benefits of both approaches. The detection of unknown attacks can rely on online unsupervised anomaly detection systems that are adept at recognizing signs of zero-day attacks, all the while automatically adapting to concept drift without the constant need for manual intervention. This, in turn, can also reduce the frequency of model re-training and enhance system efficiency. Such systems will be used in conjunction with supervised ones to improve the overall accuracy for known attacks.

The efficacy of our methodologies will be validated through extensive experimental evaluation, showcasing our system's capability of real-time threat detection compared to traditional models. This will provide the research community with valuable insights into the effectiveness of different ML methods and ensemble strategies against a wide range of security attacks.

Looking forward, we envision further enriching our IDS framework to improve its resilience against unknown attacks and concept drift, offering robust defenses against the ever-evolving landscape of cyber threats.

6. Research Unit

The *Networks and Distributed Systems* and *Artificial Intelligence* research laboratories at the University of Palermo, directed by Prof. Giuseppe Lo Re and Salvatore Gaglio, have experience in several research fields such as distributed systems, cybersecurity, artificial intelligence, and machine learning. In particular, the research unit has developed deep expertise in several topics related to the cybersecurity domain that mainly concern the adoption of artificial intelligence to assist the detection and identification of potential threats in cyberspace. The identified methodologies and proposed solutions have been applied in different scenarios, such as intrusion detection systems [10], malware detection systems [13, 14], social network security [15, 16], privacy-preserving distributed systems [17, 18], adversarial machine learning [19] and secure crowdsensing [20].

Furthermore, it is worth noting that the research group's experience in applying artificial intelligence approaches and methods to distributed systems and cybersecurity challenges has been leveraged in several funded research projects, such as FRASI - FRamework for Agent-based Semantic- aware In-teroperability (FAR MIUR D.M. 8 agosto 2000), Bigger Data (D.D. MIUR n. 2690 dell'11.12.2013, Piano di Azione e Coesione), SeNSori - SEnsor Node as a Service for hOme and buildings

eneRgy savIng (Industria 2015: Bando Nuove Tecnologie per il Made in Italy), Smart Buildings - An Ambient Intelligence system for optimizing energy resources in building complexes (PO FESR Sicilia 2007-2013), OnSicily.com - a Web 3.0 platform with intelligent virtual A.V.I. assistance (PO FESR Sicilia 2007-2013), VASARI - VAlorizzazione Smart del patrimonio ARtistico delle città Italiane (PNR 2015-2020), CrowdSense (PO FESR Sicilia 2014-2020), Smart Wave (PO FESR Sicilia 2014-2020), S6 Project - A Smart, Social and SDN-based Surveillance System for Smart-cities (PO FESR Sicilia 2014-2020), S3 Campus - SHARING, SMART AND SUSTAINABLE CAMPUS (POC Sicilia 2014-2020), Smart Venues for Agrotech Ecosystem (POC Sicilia 2014-2020).

References

[1] T. Zoppi, A. Ceccarelli, T. Puccetti, A. Bondavalli, Which algorithm can detect unknown attacks? comparison of supervised, unsupervised and meta-learning algorithms for intrusion detection, Computers & Security 127 (2023) 103107.

[2] A. Khraisat, I. Gondal, P. Vamplew, J. Kamruzzaman, Survey of intrusion detection systems: techniques, datasets and challenges, Cybersecurity 2 (2019) 20. URL: https://doi.org/10.1186/s42400-019-0038-7. doi:10.1186/s42400-019-0038-7.

[3] A. L. Buczak, E. Guven, A survey of data mining and machine learning methods for cyber security intrusion detection, IEEE Communications Surveys & Tutorials 18 (2016) 1153–1176. doi:10.1109/COMST.2015.2494502.

[4] J. Lu, A. Liu, F. Dong, F. Gu, J. Gama, G. Zhang, Learning under concept drift: A review, IEEE Transactions on Knowledge and Data Engineering 31 (2018) 2346–2363.

[5] A. A. Aburomman, M. B. I. Reaz, A survey of intrusion detection systems based on ensemble and hybrid classifiers, Computers & Security 65 (2017) 135–152. URL: https://www.sciencedirect.com/science/article/pii/S0167404816301572. doi:https://doi.org/10.1016/j.cose.2016.11.004.

[6] B. A. Tama, M. Comuzzi, K.-H. Rhee, Tse-ids: A two-stage classifier ensemble for intelligent anomaly-based intrusion detection system, IEEE Access 7 (2019) 94497–94507. doi:10.1109/ACCESS.2019.2928048.

[7] Y. Zhou, G. Cheng, S. Jiang, M. Dai, Building an efficient intrusion detection system based on feature selection and ensemble classifier, Computer Networks 174 (2020) 107247. URL: https://www.sciencedirect.com/science/article/pii/S1389128619314203. doi:https://doi.org/10.1016/j.comnet.2020.107247.

[8] S. Seth, K. K. Chahal, G. Singh, A novel ensemble framework for an intelligent intrusion detection system, IEEE Access 9 (2021) 138451–138467. doi:10.1109/ACCESS.2021.3116219.

[9] X. Li, M. Zhu, L. T. Yang, M. Xu, Z. Ma, C. Zhong, H. Li, Y. Xiang, Sustainable ensemble learning driving intrusion detection model, IEEE Transactions on Dependable and Secure Computing 18 (2021) 1591–1604. doi:10.1109/TDSC.2021.3066202.

[10] V. Agate, D. Felice Maria, A. De Paola, P. Ferraro, G. Lo Re, M. Morana, A behavior-based intrusion detection system using ensemble learning techniques., in: ITASEC, 2022, pp. 207–218.

[11] A. De Paola, P. Ferraro, S. Gaglio, G. Lo Re, M. Morana, M. Ortolani, D. Peri, A context-aware system for ambient assisted living, in: S. F. Ochoa, P. Singh, J. Bravo (Eds.), Ubiquitous Computing and Ambient Intelligence, Springer International Publishing, Cham, 2017, pp. 426–438.

[12] I. Sharafaldin, A. H. Lashkari, A. A. Ghorbani, Toward generating a new intrusion detection dataset and intrusion traffic characterization, in: Proceedings of the 4th International Conference on Information Systems Security and Privacy - Volume 1: ICISSP,, INSTICC, SciTePress, 2018, pp. 108–116. doi:10.5220/0006639801080116.

[13] A. De Paola, S. Gaglio, G. Lo Re, M. Morana, A hybrid system for malware detection on big data, in: IEEE INFOCOM 2018 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), 2018, pp. 45–50. doi:10.1109/INFCOMW.2018.8406963.

[14] F. Concone, A. De Paola, G. Lo Re, M. Morana, Twitter analysis for real-time malware discovery, in: 2017 AEIT International Annual Conference (2017 AEIT), Cagliari, Italy, 2017.

[15] F. Concone, G. Lo Re, M. Morana, S. K. Das, Spade: Multi-stage spam account detection for online social networks, IEEE Transactions on Dependable and Secure Computing (2022) 1–16. doi:10.1109/TDSC.2022.3198830.

[16] F. Concone, G. Lo Re, M. Morana, C. Ruocco, Twitter spam account detection by effective labeling, in: 3rd Italian Conference on Cyber Security, ITASEC 2019, volume 2315, IT, 2019.

[17] V. Agate, P. Ferraro, G. Lo Re, S. K. Das, Blind: A privacy preserving truth discovery system for mobile crowdsensing, Journal of Network and Computer Applications (2023) 103811. URL: https://www.sciencedirect.com/science/article/pii/S1084804523002308. doi:https://doi.org/10.1016/j.jnca.2023.103811.

[18] V. Agate, A. De Paola, P. Ferraro, G. Lo Re, M. Morana, Secureballot: A secure open source e-voting system, Journal of Network

and Computer Applications 191 (2021) 103165. URL: https://www.sciencedirect.com/science/article/pii/S1084804521001776. doi:https://doi.org/10.1016/j.jnca.2021.103165.

[19] S. Gaglio, A. Giammanco, G. Lo Re, M. Morana, Adversarial machine learning in e-health: attacking a smart prescription system, in: International Conference of the Italian Association for Artificial Intelligence (2021 AI*IA), Milan, Italy, 2021.

[20] F. Concone, G. Lo Re, M. Morana, Smcp: a secure mobile crowdsensing protocol for fog-based applications, Human-centric Computing and Information Sciences 10 (2020) 1–23. URL: https://doi.org/10.1186/s13673-020-00232-y. doi:10.1186/s13673-020-00232-y.

Vol-3762
urn:nbn:de:0074-3762-7

Ital-IA 2024
Ital-IA 2024 Thematic Workshops

Proceedings of the Ital-IA Intelligenza Artificiale - Thematic Workshops co-located with the 4th CINI National Lab AIIS Conference on Artificial Intelligence (Ital-IA 2024)

Naples, Italy, May 29-30, 2024.

Edited by

Sergio Di Martino *
Carlo Sansone *
Elio Masciari *
Silvia Rossi *
Michela Gravina *

* University of Naples Federico II, Department of Electrical Engineering and Information Technology (DIETI), Via Claudio 21, 80125, Naples, Italy

Table of Contents

- Summary: There were 128 papers submitted for peer-review to Ital-IA 2024 Thematic Workshops. Out of these, 93 papers were accepted for this volume, as short papers.

Thematic Workshop: Generative AI

Thematic Workshop: Responsible and Trustworthy AI

Thematic Workshop: AI for Cybersecurity

Thematic Workshop: AI for Industry

Thematic Workshop: AI for Finance and Marketing

Thematic Workshop: AI for Health and Medicine

Thematic Workshop: AI for the Public Administration

Thematic Workshop: AI and Sustainability

2024-09-12: submitted by Michela Gravina, metadata incl. bibliographic data published under Creative Commons CC0
2024-09-21: published on CEUR Workshop Proceedings (CEUR-WS.org, ISSN 1613-0073) |valid HTML5|