Journal of Computer Science

Original Research Paper

Cybersecurity Mechanism for Automatic Detection of IoT Intrusions Using Machine Learning

[1]Cheikhane Seyed, [2]MBaye Kebe, [3]Mohamed El Moustapha El Arby, [3]El Benany Mohamed Mahmoud and [3]Cheikhne Mohamed Mahmoud Seyidi

[1]Department of Information Systems, University of Nouakchott Al Aasriya, Nouakchott, Mauritania
[2]Higher School of Polytechnic of Dakar, University Cheikh Anta Diop, Dakar, Senegal
[3]Department of Mathematics and Computer Science, Nouakchott University, Nouakchott, Mauritania

Article history
Received: 29-08-2023
Revised: 05-10-2023
Accepted: 09-10-2023

Corresponding Author:
Cheikhane Seyed
Department of Information Systems, University of Nouakchott Al Aasriya, Nouakchott, Mauritania
Email: ch.hamod@gmail.com

Abstract: This article proposes an ML-based cyber security mechanism to optimize intrusion detection that attacks internet objects (IoT). Our approach consists of bringing together several learning methods namely supervised learning, unsupervised learning and reinforcement learning within the same Canvas. The objective is to choose among them the most optimal for classifying and predicting attacks while minimizing the impact linked to the learning costs of these attacks. In our proposed model, we have used a modular design to facilitate the implementation of the intrusion detection engine. The first Meta-learning module is used to collect metadata related to existing algorithmic parameters and learning methods in ML. As for the second module, it allows the use of a cost-sensitive learning technique so that the model is informed of the cost of intrusion detection scenarios. Therefore, among the ML classification algorithms, we choose the one whose automatic learning of intrusions is the least expensive in terms of its speed and its quality in predicting reality. This will make it possible to control the level of acceptable risk in relation to the typology of cyber-attacks. We then simulated our solution using the Weka tool. This led to questionable results, which can be subject to the evaluation of model performance. These results show that the classification quality rate is 93.66% and the classification consistency rate is 0.882 (close to unit 1). This proves the accuracy and performance of the model.

Keywords: IA, IoT, Cyber Security, Machine Learning, Weka Tools, Performance Evaluation

Introduction

The internet of things is a technological innovation with inconceivable growth, impact and capabilities (Mazon-Olivo and Pan, 2021). It refers to the connection of a set of physical objects to the Internet, allowing them to collect and exchange data. However, the security of objects (IoT) is a major concern for their users given the proliferation of connected objects and the massive amount of data they collect (Yang *et al*., 2021).

What's more, the internet of things is one of the weakest links in the cyber security chain. It includes several technological failures. This applies to the application, internet and network layers. Which may lead to unnecessary or insecure network services (Iqbal *et al*., 2020). In particular, people exposed to the internet compromise security properties such as confidentiality, integrity, authenticity and availability of information. This can lead to unauthorized remote control from cyberspace. The attack surface of computer systems is expanding. The level and dynamics of vulnerabilities are increasing over time.

Faced with this new modus operandi in the attack chain, reactive network security methods are becoming inadequate. It is true that automatic intrusion detection systems have been deployed (Shi *et al*., 2022; Tran *et al*., 2017).

However, the problem is that these articles do not emphasize the impact linked to the learning cost, which can make attack prediction more difficult. Which leads us to optimize the detection of attacks by choosing the optimal method in terms of learning cost for fast, efficient attack prediction that reflects reality.

The contribution consists of proposing cyber security models using self-learning to optimize the automatic detection of intrusions into IoTs. Our new approach is based on an interactive modular design for the analytical segmentation of the learning process by splitting the modelling into two layers, one of which is dedicated to metadata management and the other is used for cost discrimination.

For the simulation of the model, we used the Weka tool, which is a set of machine learning algorithms for data mining. It encompasses tools for data preparation, regression, clustering, classification, statistical visualization and association rule mining.

This simulation of the model produced questionable results that can be used to evaluate the model. In this regard, we find a classification quality rate close to 93% and a classification consistency rate of 0.882 (close to unit 1), demonstrating a tendency towards convergence between observation and prediction. We deduce the precision and performance of the evaluated model.

The rest of this article is organized into the following sections: A section discusses the context of security in IoT and presents a detailed literature review on existing models for IoT intrusion detection. The following section presents the methods and materials used in the manuscript; then a section dedicated to the results and discussions and finally we present the conclusion of our work.

Background on IoT Security

IT security, as well as that of the internet of things, begins with a definition of the target perimeter. However, a critical look at the composition and environment of IoTs leads us to conceptualize the vulnerabilities associated with IoTs (Lin *et al*., 2017). These vulnerabilities relate to the physical integrity of the objects and their functional coherence. On the other hand, the different layers of integration of the IoT ecosystem are all affected by the presence of threats.

IoT security is even broader given the variety of methodologies it offers in this context. It includes hashing algorithms for data integrity, public Private Key Infrastructure (PKI) cryptography for confidentiality. Data availability and access controls for data availability and network security are just some of the methods IT departments can use to combat the growing threat of cybercrime rooted in vulnerable IoT devices.

All the resources forming part of the IoT asset must be identified. The prevention and protection measures for these resources vary according to the security models chosen from the recognized standards. Figure 1 illustrates how to integrate security into IoT connected smart cities.

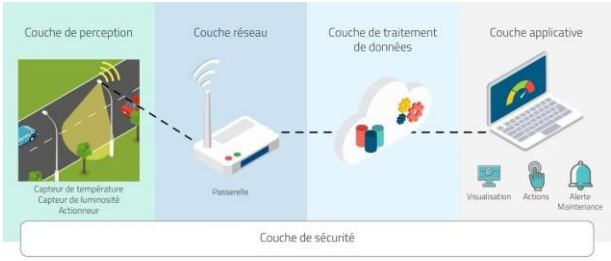

Fig. 1: Smart city connected lighting system

Each organization defines its own security strategy. It is important that its implementation complies with cybersecurity standards (NIST, ISO, COBIT, etc.). Depending on the objectives selected, the applicable security measures may be derived from several established standards. These standards help in the development of information and IT system security programmed (Ammar *et al*., 2018).

This helps to reduce the risk of attacks and to manage vulnerabilities. In the case of the "NIST cyber security framework", fundamental cyber security functions must be fulfilled.

These functions are not intended to follow a serial path or lead to a static end state. However, they must be executed simultaneously and continuously to form an operational culture to respond to dynamic cyber security risks (Restuccia *et al*., 2018).

Machine Learning: A Solution for IoT Security

Machine learning refers to intelligent approaches used to improve system performance based on dataset or experience. It also allows machines to learn without being explicitly programmed. These models are used to make future predictions using training and testing data.

It is interdisciplinary in nature and draws its advantages in many technical and scientific fields such as artificial intelligence, information theory, optimization theory and cognitive science (Qiu *et al*., 2016).

It is also used in situations where the solution to a problem may evolve over time (routing in a computer network or search for malicious code in software or an application).

In addition, several practical systems for intelligent security use ML for example, Google uses ML to predict attacks that hit mobile devices and apps running Android. It is also used to identify and remove malware in infected phones. Similarly, Amazon launched a service that uses ML to sort and classify data stored on its cloud storage service.

However, the use of advanced security techniques in IoT applications brings new challenges. These are diverse and varied. For example, it seems difficult to develop an

appropriate model to classify data from various IoT objects. Similarly, effective labeling of input data is also a tedious task.

Further challenges arise from deploying these models on IoT devices in the face of limited resources (Yao *et al*., 2018). In the above context, it is imperative to leverage the capabilities offered by ML to systematically improve security solutions in IoT.

Literature Review of Existing Models

Intrusion detection systems are built on the basis of data collected and trained using supervised, semi-supervised and unsupervised learning methods (Yao *et al*., 2018). This article offers solutions for evaluating the performance of intrusion detection systems over the long term. The aim is to be able to detect as yet unknown zero-day attacks.

On the other hand, a synthesis on the analysis of threats, problems and security solutions for cloud computing uses machine learning algorithms (Butt *et al*., 2020). They are used to solve security problems in IoT objects using supervised, unsupervised, semi-supervised and reinforcement learning.

The internet of connected objects in the industrial domain (I-IoT) is also an active area of research and is the subject of several studies. The problem of low detection rates and high proportions of false alarms is addressed in this article (Butt *et al*., 2020). The sole objective of this study is to detect and stop cyber-attacks. Concerns about the costs and impact of this detection are not the focus of attention.

The article (Khan *et al*., 2022) makes an important contribution to solving the problem of the security of connected objects. A thorough analysis of the literature is attributed to them. The articles cited in this study certainly differ in their aims and objectives. Some of them approach the issue from the reasonable angle of the technical constraints intrinsic to IoT, in particular storage, memory and energy.

Other authors (Williams *et al*., 2022; Kumar and Bansal, 2019) introduce the notions of layered architecture with or without the integration of techniques such as machine learning, artificial intelligence and cryptography. The contextualization of the problem of security for connected objects remains reactive and corrective. However, the solutions proposed do not seem to be part of an innovative and proactive methodology.

In the article (Sifat *et al*., 2022), the authors have access to a review of the typology of anomalies, the detection layers, the context and the methodology. What emerges is an oversimplified view of anomaly classification. All attacks are classified in a single anomaly category, resulting in only four anomaly types. Moreover, this represents more than half of the population. In addition, the type of attack is not well specified. Over 90% of the articles do not take context into account. This further weakens the robustness of the proposed solutions.

In the article, (Rodríguez *et al*., 2023) show the need to focus on learning methods, the quality of the data used and the importance of safety issues in free decision-making. This last point is crucial in terms of the cognitive dimension of the proposed solution.

The authors (Naji and Zougagh, 2023) have proposed several approaches based on Recurrent Neural Networks (RNN) using Long Term Memory (LSTM), auto encoders and multi-layer perceptron's.

However, these articles do not seem to draw attention to the innovative approach of modular analysis and segregation of security of learning costs. Which pushed us to focus in this research study on the impact of the learning cost of ML methods thus influencing the effectiveness of attack predication in the IoT.

Materials and Methods

Weka Tools

The physical implementation of the IoT cyber security model requires the use of hardware and software resources. In order to produce an Optimized Physical Cyber security Model (OPCM), we opted for open source software that is well known in the scientific research community. This is Weka and its applications.

Weka provides implementations of machine learning algorithms that can be easily applied to a database. It also includes a variety of tools for data transformation and classification. These include discretization and sampling algorithms. It can also be used to pre-process a data set, integrate it into a learning system and analyze the resulting classifier and its performance (Lal *et al*., 2023).

Weka software is a data analysis and modeling platform. This software adopts both supervised and unsupervised data learning methods. A method is said to be supervised if the values of the variable of interest are fixed before learning (Fig. 2).

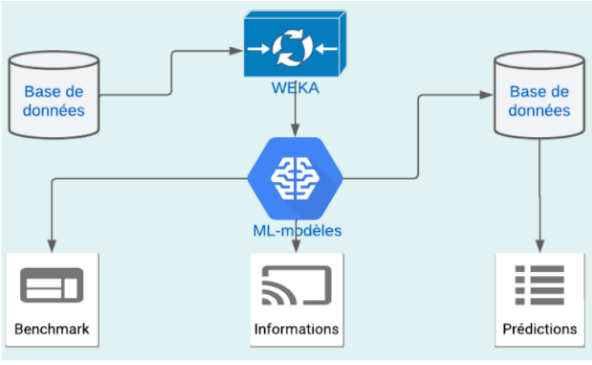

Fig. 2: Weka learning options

Weka contains tools to provide solutions to data mining problems such as regression, classification, association rule mining, attribute clustering and selection.

There are several uses of Weka, as shown in Fig. 2:

- One method is to apply a learning model to perform analyses using a dataset to learn more about the data
- Another is to use the learned models to generate predictions on new instances
- A third is to apply several different models and compare their performance in order to select one for prediction

In this article, we build on the first option by applying the One-R algorithm to verify the performance of our cyber security mechanism in effectively predicting attacks hitting the IoT.

Learning Method

We use a dataset to verify the performance of our proposed approach to IoT cyber security. To do this, we use the One-R (Miao *et al*., 2021) basic classification method.

The advantage of the basic One-R method lies in its algorithmic simplicity. In other words, the One Rule (One-R) method uses a single prediction rule (variable) to design the cyber-attack detection model. Consequently, this method simply ignores the information from the other variables in the optimal selection described above (Mashhadi and Garousi, 2020).

However, like Zero-R, the One-R method is one of the basic methods. This means that it allows us to create an initial state. This is a basic state (zero state) of the system to be improved towards a more reasonable and acceptable optimum.

In short, the use of the zero state is for comparative purposes in the search for an optimal classification model. Other learning methods will have to be tested and compared with the zero state in order to better optimise the detection model in terms of hardware resources and algorithmic performance.

Confusion Matrix

The confusion matrix is a tool for predictive analysis in machine learning. It is also called error matrix and is used to evaluate the performance of a classification-based machine-learning model. This type of model aims to predict a categorical label for each input instance (Schmelzer, 2020).

Table 1: Confusion matrix basic metrics

	Observation	
Techniques	----------	----------
Learning (Predicted Class)	True Positives (TP) False Negative (FN)	False Positive (FP) True Negative (TN)

We can also say that the confusion matrix is a summer table of the number of correct and incorrect predictions produced by a classifier for binary classification tasks (García-Balboa *et al*., 2018).

The matrix illustrates the number of True Positives (TP), True Negatives (TN), False Positives (FP) and False Negatives (FN) produced by the model on the test data (Table 1):

- True Positive (TP): When the actual value is positive and predicted is positive
- True Negative (TN): When the actual value is negative and prediction is negative
- False Positive (FP): When the actual is negative but prediction is positive
- False Negative (FN): When the actual is positive but the prediction is negative. Also known as the type two errors

The confusion matrix is used to identify the various errors made, particularly those made by a prediction algorithm. By analyzing them, it is possible to determine the results that indicate how these errors occurred. Knowledge of the type of error is a major advantage of the confusion matrix.

Evaluation Criteria

In this validity assessment study of the proposed model, we focus on several metrics such as Classification quality; error rate and Kappa coefficient are monomeric. Sensitivity, specificity and accuracy are multimeric.

Classification quality: This is the proportion of well classified individuals relative to the total population examined. It is obtained using the following equation:

$$Accuracy = (TP + TN) / \left(\left(TP + FN\right) + \left(TN + FP\right)\right)$$

Classification consistency: This can be evaluated using the Kappa coefficient. The purpose of this indicator is to show the convergence and divergence between predictions and observations.

Model sensitivity: This is the actual proportion of packets included in this form of cyber-attack. It is obtained using the following equation:

$$Sensibilite' = TPR = TP / (TP + FN)$$

Model specificity: This is the actual proportion of misclassified packets excluded from this form of cyber-attack. It is obtained using the following equation:

$$FPR = FP / (FP + TN)$$

$$Spe'cificite' = 1 - FPR = NT / (TN + FP)$$

Accuracy: This is an indicator of false alarms. It provides answers to the following question. What

proportion of positive identifications are correct? It is obtained using the following equation:

$$Pre'cision = TP / (TP + FP)$$

Recall (true positive rate): This is a metric that answers the following question. What proportion of true positive results have been correctly identified? It is obtained using the following equation:

$$Recell = TP / (TP + FN)$$

Results and Discussion

Optimizing IoT Cyber-Security

The motivation for proposing a new model for optimizing intrusion detectors lies in the fact that NIDSs attempt to apply the same intrusion filters regardless of the risk policies in place. As zero risk is unrealistic, it is essential to control its assessment and acceptance level.

The ultimate goal of this approach is to develop a model allowing you to choose among machine learning methods (supervised, unsupervised and reinforcement), the optimal method to effectively detect intrusions in IoT objects. This choice is essentially based on the cost-sensitive technique thus minimizing the learning impact of attacks and intrusions in IoT objects.

Upgrading the generic classification and detection model involves redefining the methodology. To do this, we have imagined the creation of two functional layers at the conceptual level.

This leads us to the modular programming of the detection engine. On the one hand, the first module is used to design the algorithmic component of the methodology in order to integrate a wide range of learning methods. On the other hand, the second module models the security-cost component of the methodology. This enables the security manager to control the acceptable level of risk in relation to the typology of cyber-attacks. In this way, the most optimal classification method will be chosen. This is the least costly method in terms of negative impact.

We find ourselves in a process of optimization automation with the possibility of acting on the algorithmic and security parameters.

Achieving the optimization objectives set previously will be a prerequisite for arriving at the optimized cybersecurity logic model (Fig. 3). This process is carried out in two successive phases. First, we introduce the concept of Meta-learning on which the algorithmic policy is based. Next, we present the Cost-Sensitive-Learning technique. This will make it possible to implement the entity's security policy in terms of intrusion detection and cost control.

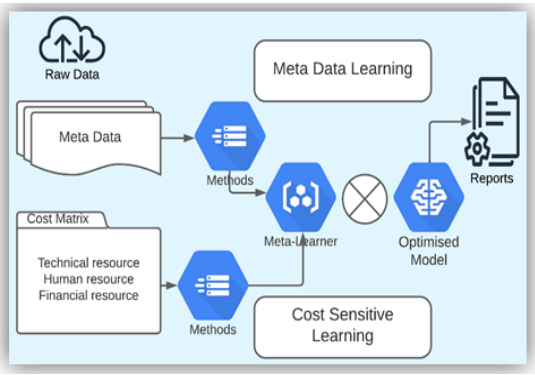

Fig. 3: Optimization of the learning process by meta-learning and cost sensitive learning

Meta-learning corresponds to what might be called macro learning. It involves understanding the behavior of several learning methods. The aim is to collect metadata made up of performance values and algorithmic parameters associated with the methods.

This approach allows several learning methods to be nested or encompassed within a single Caneva. We know that the quality of an algorithm includes not only its effectiveness in predicting reality, but also the speed with which it is executed. This is the basic principle of the meta-learning process.

In the second block, Safety Policy, we propose to introduce the notion of learning costs (impacts). The aim of classical learning is to minimize the errors generated by the difference between prediction and observation. Since not all errors have the same cost or impact, we will use the cost-sensitive learning technique. The fundamental principle of cost-sensitive learning is that the learning engine is informed of the cost or impact of the intrusion detection scenarios.

Evaluation of Model Optimization

To evaluate the relevance and quality of our model, we used the dataset from the data-sharing platform made available to computer security researchers, by the Hacking and Countermeasure Research (HCR) laboratory specializing in research into cyber-attacks and countermeasures.

The evaluation study of the proposed model is based on a set of tests obtained after simulating the model in the Weka tool. These data will be provided as input values to the prediction function. The results of this operation will be compared with the corresponding observation values.

As these results show in Fig. 4, all 2863 packets are predicted to be Denial of Service (DoS category). This prediction is confirmed by the observations of the test set. There are also 1928 packets recognized as benevolent (Normal category) in both prediction and observation. We find the errors made by the model for a column of observations outside the first diagonal. There were 148 predictions, including 147 botnets and 1 ARP spoofing. This prediction is contradicted by the reality of the test data.

a	b	c	d	e	<-- classified as
19970	147	44	548	110	a = Mirai
54	2863	3	18	5	b = DoS
143	0	2877	649	29	c = Scan
41	0	1	1928	18	d = Normal
141	1	17	14	1668	e = MITM ARP Spoofing

Fig. 4: Method One-R method confusion matrix

To assess the validity of the model, these objectives can be measured by several indicators at the same time. We have the holistic statistical estimates, which evaluate the overall performance of the model. These indicators reflect the quality, shortcomings and consistency of the learning process.

Figure 5, the results obtained illustrate the following behaviors.

The classification quality rate is 93.66%. This rate shows a high level of conformity between predictions and observations. This result is supported by a classification consistency rate of 0.882 (close to unity 1), demonstrating a trend towards convergence between observation and prediction. This deduces the accuracy and performance of the model evaluated.

We then proceed to deepen our assessment of the model's validity using other metrics such as the model's sensitivity, specificity and precision.

From these test results (Fig. 6), we deduce that the sensitivity varies between 77.8 and 97.3%, while the specificity of the model is between 95.8 and 99.8%. The false alarm rate (false positives) is therefore between 0.2 and 4.2%.

Correctly Classified Instances	29306	93.6623 %
Incorrectly Classified Instances	1983	6.3377 %
Kappa statistic	0.882	
Mean absolute error	0.0254	
Root mean squared error	0.1592	
Relative absolute error	12.0258 %	
Root relative squared error	49.0447 %	
Total Number of Instances	31289	

Fig. 5: Performance statistics and One-R

	TP Rate	FP Rate	Precision	Recall	F-Measure	Class
	0,959	0,036	0,981	0,959	0,970	Mirai
	0,973	0,005	0,951	0,973	0,962	DoS
	0,778	0,002	0,978	0,778	0,867	Scan
	0,970	0,042	0,611	0,970	0,749	Normal
	0,906	0,006	0,911	0,906	0,909	MITM ARP
Weighted Avg.	0,937	0,028	0,950	0,937	0,939	

Fig. 6: Learning assessment metrics

The sensitivity and specificity of a model are crucial to detection. A sensitive model must therefore sense attacks and block them at the right moment. If, in addition, this model only detects real attacks, without adding false alarms, it is then qualified as a specific model.

We then generated two curves, the Receiver Operating Characteristic Curve (ROC) and the Precision-Recall Curve (PRC). These curves are graphical metrics for assessing detection sensitivity and prediction accuracy.

When faced with negative individuals, a specific model must minimize the False Positive Rate (FPR) = (1 Specificity). This introduces the notion of the Area Under ROC Curve (AUC), which must tend towards 1 to be optimal.

As the result shows in Fig. 7, this AUC is 0.8879 for the model evaluated. This is very close to 1, so we can deduce that the model is optimal.

The ROC curve is based on the majority class. It therefore becomes ineffective for samples unbalanced in favour of the majority class. For this reason, the PRC curve is used to complete the analysis.

For the latter, the proportion of individuals confirmed positive out of all the predictions for a criterion is given by the Precision. The Precision-Recall Curve (PRC) must be convex towards unity, which is the optimum. It can be seen from the Fig. 8 that the curve covers almost a unit area for the basic model.

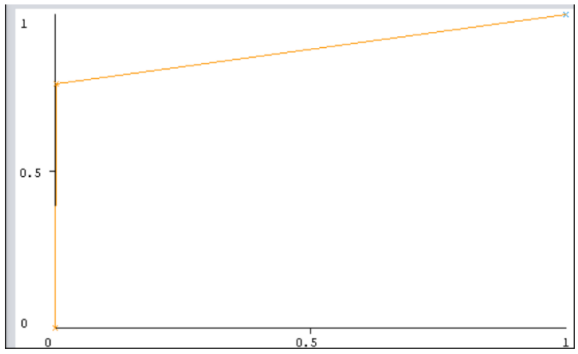

Fig. 7: Sensitivity and ROC curve

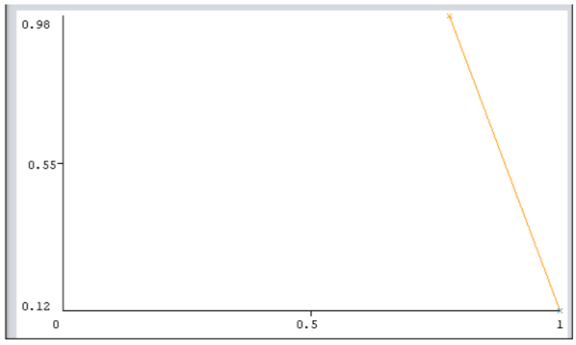

Fig. 8: Specificity and PRC curve

Conclusion

IoT security plays a central role in the commercialization of IoT technology. The study revealed major IoT security risks. The physical security of the IoT is questioned on the remote sites vis-à-vis their hardware and software upgrades and updates which are essential.

Traditional security and privacy solutions suffer from a number of concerns related to the dynamic nature of IoT networks. This threat is heightened by the availability of tools to find and exploit vulnerabilities in the IoT system. In this context, we have proposed a cyber-security model for the IoT whose main objective is to optimize the detection of intrusions thanks to learning algorithms. This optimization is based on algorithmic and security policies by integrating the potential of algorithmic methods and the reduction of learning costs.

To verify the validity of the proposed approach, we proceeded to the evaluation of the model. This evaluation leads to prove the validity of the model in terms of quality, consistency of classification and sensitivity and specification of the model.

In terms of perspective, we intend to use other learning methods and compare the results in order to analyze their impacts on the security of IoT and to deduce the best in terms of performance and quality of intrusion detection in the connected objects.

Acknowledgment

First, we would like to thank God for giving us courage and overcoming all difficulties.

Secondly, I would like to warmly greet and thank the members of the team of two Labs (LIRT and CSIDS) who contributed directly or indirectly to the development of this manuscript.

Finally, I would also like to particularly thank Prof. Jeanne Roux Bilong who supervised this study.

Funding Information

The authors have not received any financial support or funding to report.

Author's Contributions

Cheikhane Seyed: Developed the methodology, designed the proposed approach, evaluated the model, and interpreted the results.

MBaye Kebe: Conducted data collection, assisted in the analysis and contributed to the interpretation of the source data.

Mohamed El Moustapha El Arby: Focused on model simulation, managed the Weka tool, executed tests, and produced results.

El Benany Mohamed Mahmoud: Provided supervision and guidance throughout the semantic research component of the project.

Cheikhne Mohamed Mahmoud Seyidi: Contributed to annotation and interpretation; performed revisions and corrections to the manuscript prior to submission.

Ethics

The authors confirm that this manuscript has not been published elsewhere and no ethical issues are involved because the article conforms to all scientifically known ethical principles.

References

Ammar, M., Russello, G., & Crispo, B. (2018). Internet of things: A survey on the security of IoT frameworks. *Journal of Information Security and Applications*, *38*, 8-27. https://doi.org/10.1016/j.jisa.2017.11.002

Butt, U. A., Mehmood, M., Shah, S. B. H., Amin, R., Shaukat, M. W., Raza, S. M., ... & Piran, M. J. (2020). A review of machine learning algorithms for cloud computing security. *Electronics*, *9*(9), 1379. https://doi.org/10.3390/electronics9091379

García-Balboa, J. L., Alba-Fernández, M. V., Ariza-López, F. J., & Rodriguez-Avi, J. (2018). Homogeneity test for confusion matrices: A method and an example. *In IGARSS 2018 IEEE International Geoscience and Remote Sensing Symposium*, 1203-1205. IEEE. https://doi.org/10.1109/IGARSS.2018.8517924

Iqbal, W., Abbas, H., Daneshmand, M., Rauf, B., & Bangash, Y. A. (2020). An in-depth analysis of IoT security requirements, challenges, and their countermeasures via software-defined security. *IEEE Internet of Things Journal*, *7*(10), 10250-10276. https://doi.org/10.1109/JIOT.2020.2997651

Khan, I. A., Keshk, M., Pi, D., Khan, N., Hussain, Y., & Soliman, H. (2022). Enhancing IIoT networks protection: A robust security model for attack detection in internet industrial control systems. *Ad Hoc Networks*, *134*, 102930. https://doi.org/10.1016/j.adhoc.2022.102930

Kumar, A., & Bansal, A. (2019). Software fault proneness prediction using genetic based machine learning techniques. *In 2019 4th International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU)*, 1-5. IEEE. https://doi.org/10.1109/IoT-SIU.2019.8777494

Lal, B., Ravichandran, S., Kavin, R., Kumar, N. A., Bordoloi, D., & Kumar, R. G. (2023). IoT-based cyber security identification model through machine learning technique. *Measurement: Sensors*, *27*, 100791. https://doi.org/10.1016/j.measen.2023.100791

Lin, J., Yu, W., Zhang, N., Yang, X., Zhang, H., & Zhao, W. (2017). A survey on internet of things: Architecture, enabling technologies, security and privacy, and applications. *IEEE Internet of Things Journal*, *4*(5), 1125-1142. https://doi.org/10.1109/JIOT.2017.2683200

Mashhadi, M., & Garousi, M. R. (2020). O-plane couplings at order α′ 2: One RR field strength. *Journal of High Energy Physics*, *2020*(6), 1-30. https://doi.org/10.1007/JHEP06(2020)171

Mazon-Olivo, B., & Pan, A. (2021). Internet of things: State-of-the-art, computing paradigms and reference architectures. *IEEE Latin America Transactions*, *20*(1), 49-63. https://doi.org/10.1109/TLA.2022.9662173

Miao, Y., Chen, C., Pan, L., Han, Q. L., Zhang, J., & Xiang, Y. (2021). Machine learning-based cyber-attacks targeting on controlled information: A survey. *ACM Computing Surveys (CSUR)*, *54*(7), 1-36. https://doi.org/10.1145/3465171

Naji, M., & Zougagh, H. (2023). Deep learning models for cybersecurity in IoT networks. *In International Conference on Business Intelligence*, 29-43. Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-37872-0_3

Qiu, J., Wu, Q., Ding, G., Xu, Y., & Feng, S. (2016). A survey of machine learning for big data processing. *EURASIP Journal on Advances in Signal Processing*, *2016*, 1-16. https://doi.org/10.1186/s13634-016-0355-x

Restuccia, F., D'Oro, S., & Melodia, T. (2018). Securing the internet of things in the age of machine learning and software-defined networking. *IEEE Internet of Things Journal*, *5*(6), 4829-4842. https://doi.org/10.1109/JIOT.2018.2846040

Rodríguez, M., Tobón, D. P., & Múnera, D. (2023). Anomaly classification in industrial internet of things: A review. *Intelligent Systems with Applications*, 200232. https://doi.org/10.1016/j.iswa.2023.200232

Schmelzer, R. (2020). How to build a machine learning model in 7 steps. *LeMagIT*. https://www.lemagit.fr/conseil/Comment-construire-un-modele-de-Machine-Learning-en-7-etapes

Shi, Z., He, S., Sun, J., Chen, T., Chen, J., & Dong, H. (2022). An efficient multi-task network for pedestrian intrusion detection. *IEEE Transactions on Intelligent Vehicles*, *8*(1), 649-660. https://doi.org/10.1109/TIV.2022.3166911

Sifat, F. H., Mahzabin, R., Anjum, S., Nayan, A. A., & Kibria, M. G. (2022). IoT and machine learning-based hypoglycemia detection System. *In 2022 International Conference on Innovations in Science, Engineering and Technology (ICISET)*, 222-226. IEEE. https://doi.org/10.1109/ICISET54810.2022.9775890

Tran, B., Picek, S., & Xue, B. (2017). Automatic feature construction for network intrusion detection. In *Simulated Evolution and Learning: 11th International Conference, SEAL 2017, Shenzhen, China, Proceedings 11* 569-580. Springer International Publishing. https://doi.org/10.1007/978-3-319-68759-9_46

Williams, P., Dutta, I. K., Daoud, H., & Bayoumi, M. (2022). A survey on security in internet of things with a focus on the impact of emerging technologies. *Internet of Things*, *19*, 100564. https://doi.org/10.1016/j.iot.2022.100564

Yang, K., Zhang, Y., Lin, X., Li, Z., & Sun, L. (2021). Characterizing heterogeneous internet of things devices at internet scale using semantic extraction. *IEEE Internet of Things Journal*, *9*(7), 5434-5446. https://doi.org/10.1109/JIOT.2021.3110757

Yao, S., Zhao, Y., Zhang, A., Hu, S., Shao, H., Zhang, C., ... & Abdelzaher, T. (2018). Deep learning for the internet of things. *Computer*, *51*(5), 32-41. https://doi.org/10.3390/sym15061251