AI-based Learning Model for Socio-Cybernetics Systems in Web-of-Things

Priti Singh, Geetanjali Rathee, Chaker Abdelaziz Kerrache, Muhammad Bilal *Senior Member, IEEE*, Carlos T. Calafate *Senior Member, IEEE*, Huihui Wang *Senior Member, IEEE*

***Abstract*—Cybernetic threats have become a growing concern in recent years, highlighting the need for effective Intrusion Detection Systems (IDS) to detect and prevent social cyber-attacks. Socio-cybernetics are a significant p latform f or p roviding real-time mapping, or to enable information access across heterogeneous networks. However, ontology-based knowledge and web support for social cybernetics demands for massive warehouses that provide the required computational power for log applications and data processing mechanisms, in addition to effective decision-support solutions for business by extracting useful information in a very secure and intelligent way. In this work, we propose an IDS approach that combines a tree-based XGBoost algorithm and a bidirectional LSTM network to address the limitations of traditional approaches. The proposed approach includes multiple steps such as data preprocessing, feature selection using Infinite Feature Selection (IFS) algorithm, and the application of Principal Component Analysis (PCA) for dimensionality reduction. Furthermore, a direct trust-based scheme is used to strengthen the decision-making process by improving the overall accuracy in the network. The performance of the proposed approach is evaluated based on accuracy, precision, recall, and F1 score, being compared with the existing LBDMIDS method. Experimental results demonstrate that the proposed approach outperforms traditional methods by providing a higher accuracy, along with a slight improvement in terms of precision, recall, and F1 score. In particular, the proposed mechanism shows a 99% improvement in terms of accuracy as compare to existing schemes, while ensuring a secure communication in the network.**

***Index Terms*—Socio-cybernetics, Web-of-things, Onotology, cyber security, detection system, machine learning, security methods.**

I. Introduction

The Web-of-Things is considered as one of the emerging paradigms for connecting and controlling intelligent devices. In particular, it acts as an interface to monitor smart devices via web. Furthermore, socio-cybernetics is a platform that provides real-time, robust and massive data access to heterogenous sources by relying on automatic learning, pattern recognition, computational intelligence etc. Such a system enables individuals to generate, maintain, and update their social networks by using different languages and ideas that enable building ontology knowledge. Socio-cybernetics is further described as a theoretical framework that enables improving and understanding the cooperative behaviour while providing information in the context of the theory of evolution. The ontology knowledge in the Web of socio-cybernetics enhances the process of co-operative behavior understandability, in addition to decision-support in business, all by extracting useful information from the surroundings. In order to achieve this, the system requires massive warehouses, log applications, and data processing in a very secure and intelligent way. However, the huge amount of data generated attracts intruders who attempt to compromise or invade the system using malicious elements for their own benefit. Hence, the increasing frequency and sophistication of social cybernetics have made intrusion detection systems (IDSs) a critical component of network security [1].

IDSs are designed to identify and respond to unauthorized access or malicious activity on computer systems and networks, helping to prevent data breaches, theft of sensitive information, and other cyber-attacks. The need for IDSs arises from the limitations of traditional network security measures such as firewalls, antivirus software, and encryption, which can only do so much in terms of detecting and preventing cyber-attacks. These measures are primarily focused on keeping attackers away from the network, but they are not always effective at identifying and stopping attacks that have already succeeded at infiltrating the network.

The number of security concerns may arise while transmitting the information in the network without knowing their communicating behaviour. The concept of intrusion detection systems can be traced back to the early 1980s when researchers first started exploring ways to detect unauthorized access to computer systems. Since then, IDSs have evolved significantly, incorporating advanced machine learning algorithms, big data analytics, and other technologies to improve their accuracy and effectiveness.

Traditional IDSs typically rely on two main methods: signature-based detection, and anomaly-based detection. Signature-based detection works by comparing incoming traffic against a database of known attack signatures. If a match is found, the system generates an alert or takes other actions to stop the attack [2], [3]. Anomaly-based detection, on the other

Priti Singh is with the Department of Computer Science and Engineering, Netaji Subhas University of Technology, Dwarka Sector-3, New Delhi-110078, India. (e-mail: priti.singh@gmail.com)

Geetanjali Rathee is with the Department of Computer Science and Engineering, Netaji Subhas University of Technology, Dwarka Sector-3, New Delhi-110078, India. (e-mail: geetanjali.rathee123@gmail.com)

Chaker Abdelaziz Kerrache is with the Laboratoire d'Informatique et de Mathématiques, Université Amar Telidji de Laghouat, Laghouat, Algeria. (e-mail: ch.kerrache@lagh-univ.dz)

Muhammad Bilal is with the Department of Computer and Electron-ics Systems Engineering, Lancaster University, Lancaster, UK. (e-mail: m.bilal@ieee.org)

Carlos T. Calafate is with the Computer Engineering Department (DISCA), Universitat Politècnica de València, Valencia, Spain. (e-mail: calafate@disca.upv.es)

Huihui Wang is with the Cybersecurity Department, St. Bonaventure University, St. Bonaventure, NY, USA (e-mail: hwang@sbu.edu)

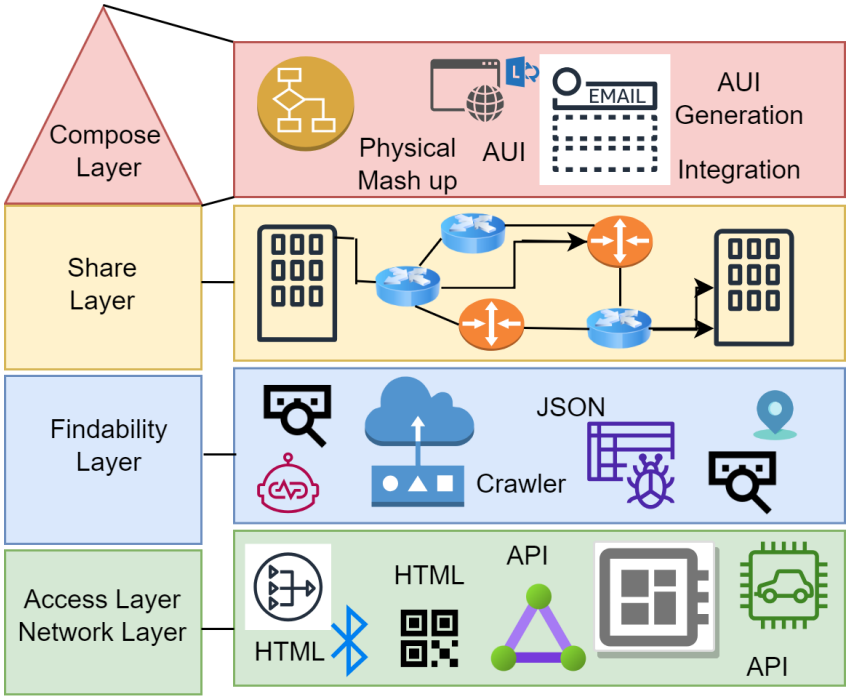

Fig. 1. General Architecture of the Web-of-Things.

hand, looks for activity that deviates from normal patterns or behavior. If a deviation is detected, the system generates an alert. Figure 1 presents the general architecture of the *Web of Things* having 4 different layers such as compose layer, sharing layer, findable layer, and network or access layer. The compose layer is used to provide the basic connection by providing basic infrastructure, where multiple routers are used to ensure efficient communication among devices. In addition, sharing layer provides the basic connection among devices by sharing the resources in the network.

A. Challenges in current IDS targetting socio-cybernetics systems

Currently existing methods have limitations when it comes to detecting new and sophisticated attacks. Signature-based detection, for example, can only identify attacks that match existing signatures in the database, and it is ineffective against zero-day attacks, which exploit previously unknown vulnerabilities. Anomaly-based detection can generate false positives or false negatives if the system's normal behavior model is not accurate, or if the attacker is skilled enough to mimic normal behavior. Additionally, Intrusion detection systems (IDSs) have become an essential component of modern network security infrastructure, playing a critical role in protecting organizations from cyber-attacks [4]. However, IDSs still face some challenges that need to be addressed in order to improve their accuracy and effectiveness. One of the major challenges faced by IDSs is the high rate of false positives and false negatives generated by traditional detection methods. False positives refer to alerts that are triggered when no actual attack is occurring, while false negatives refer to attacks that remain undetected. Both of these scenarios can be problematic for network administrators, as they can lead to unnecessary disruptions or potentially serious consequences for the network. Another challenge for IDSs is the constantly evolving nature of cyber-attacks. Attackers are constantly developing new techniques and exploiting new vulnerabilities, making it difficult for IDSs to keep up. This requires continuous updates and improvements to detection algorithms and techniques to ensure that they remain effective against the latest threats [5].

B. Motivation

To address these challenges, artificial intelligence (AI) based learning approaches have emerged as a promising area of research in the field of intrusion detection. Machine learning algorithms, such as neural networks and decision trees, can be trained on large datasets in terms of normal and abnormal network behavior to develop accurate models of network activity. These models can then be used to identify and respond to suspicious activity in real-time, reducing false positives and false negatives and improving the overall accuracy of IDSs. In addition, AI-based IDSs can adapt to new and emerging threats by continuously updating their models based on new data and feedback [6], [7]. However, there still some challenges that need to be addressed in order to fully realize the potential of AI-based intrusion detection systems. One challenge is the need for large and diverse datasets to train machine learning algorithms, which can be difficult to obtain in practice. Another challenge is the potential for attackers to bypass AI-based detection systems by manipulating their input data, or by exploiting weaknesses in the learning algorithms themselves. Despite these challenges, ongoing research on AI-based intrusion detection systems highlights the importance of continually improving and advancing network security technologies. In fact, as cyber-attacks continue to become more sophisticated and frequent, it has become imperative to design and develop effective intrusion

detection systems (IDSs) that can protect organizations and their data from potential harm [4], [8].

C. Contribution

While performing an accurate and effective decision making during transmission, or when extracting useful information, the intruders may generate several types of cyber or socio security threats in the network. By conducting a thorough literature survey, and by staying up-to-date on the latest research and technologies, we can gain a deeper understanding of the challenges faced by current IDSs, and develop new and innovative approaches that can improve their accuracy and effectiveness. In addition, in order to identify the intruders on a continuous basis, and therefore perform a continuous surveillance of network events, a trust-based computational method is required. To ensure a secure and efficient decision making process, while also extracting useful information in the context of the Web-of-Things, we introduce a direct trust-based method. The direct trust-based method is used to analyze the activities of communicating devices in the network, and to assign them a trust value based on their response. In particular, the contributions of this paper are the following:

- A tree-based XGBoost algorithm is integrated with a BiLSTM network to overcome the limitations of plain LSTM.
- An intruder detection system based on XGNoost is used to identify anomalies related to information transmission.
- The proposed mechanism is evaluated using the precision, recall and F1-score metrics.
- The proposed approach defines several steps that include data pre-processing, feature selection using the Infinite Feature Selection (IFS) algorithm, and application of Principal Component Analysis (PCA) for dimensionality reduction.
- A direct trust-based scheme is used to further strengthen the decision-making process by involving the trusted parties in the network.

The remaining structure of the paper is organized as follows. A literature survey providing an overview of several security approaches and methods to prevent cybernetic attacks, as proposed by various researchers/scientists, will be presented in section II. A secure and efficient communication model for the WoT, along with a proper algorithm, will be discussed in section III. In addition, section IV illustrates the performance analysis of the proposed framework in comparison with existing methods under various security measures. Finally, section V concludes the paper and discusses future directions.

II. Related Work

This section provides a brief literature review for current IDS systems detailing the findings, improvements, and achievements. All discussed works are also summarized in Table I.

Otoum, et. al. [9] suggested a scheme in which the spider monkey optimization algorithm and the stacked-deep polynomial network were combined to obtain the best detection and identification rates; SMO chooses the best features from the datasets, while SDPN categorizes the data as normal or anomalous. Denial of service, user-to-root attacks, probing attacks, and remote-to-local attacks are among the anomalous kinds that DL-IDS may identify. Analyses conducted in-depth show that the suggested DL-IDS performs better in terms of accuracy, precision, recall, and F-score. Baniasadi et al. [10] created a brand-new training technique to enhance the deep architecture's parameters. Eventually, in the interests of enhancing precision and efficiency, they enjoy the benefits of NSBPSO to appropriately train the deep learning model that serves as our network intrusion detector. They utilize two network breach detection datasets, UNSW-NB15 and Bot-IoT, to assess the effectiveness of the suggested classification. They score the correctness and effectiveness of the suggested classification. Jothi B., et. al. [11] performed detailed evaluations of suggested IDS with respect to validation techniques and visible metrics for the various IoT threat scenarios. Also, to assess both the suggested models and the other available learning methods, researchers conducted several in-depth tests. The WILS versions have shown superior accuracy, precision, and recall compared to other simulators now in use, demonstrating its suitability for an IoT network. Nasir et. al. [12] proposed a scheme named DF-IDS that was divided into two stages: in the first phase, the methodology used SpiderMonkey, principal component analysis, information gain, and correlation attribute assessment, to compare and pick the best features from the feature matrix. These characteristics are combined with given labels in the second step to train a deep neural network for intrusion detection. 99 percent precision can be achieved with DF-IDS. In comparison to previous systems resulting from earlier research, authors demonstrate improvements in terms of both accuracy and F1 score. S. T. Mehedi et. al. [13] suggested a reliable IDS model based on deep transfer learning that beats several current methods. The novel features comprise reliable deep transfer learning-based ResNet model construction, its evaluation taking into account real-world data, and effective attribute selection, which was best suited to detect normal and attack situations with a limited quantity of labeled data. The suggested model was reliable, more effective, and exhibited improved performance. S. T. Mehedi, et. al. [13] suggested a reliable IDS model based on deep transfer learning that outperforms several current methods. The novel features comprise reliable deep transfer learning-based ResNet model construction, evaluation taking into account real-world data, and effective attribute selection, which was best suited to detect normal and attack situations with a limited quantity of labeled data. The suggested model was reliable, more effective, and has exhibited enhanced results, according to the in-depth study and performance evaluation performed. Saba et al. [14] described a CNN-based method for anomaly-based intrusion detection systems that makes use of the potential of IoT, while offering capabilities to effectively investigate all IoT traffic. The suggested model demonstrates the capacity to recognize any potential incursion and unusual traffic patterns. The model's reliability was approximately 92% when relying on the NID Dataset and BoT-IoT datasets for training and testing, respectively. Moreover, Muthanna et al. [15] proposed a cutting-edge IoT-based collection; accepted assessment measures have been applied to accurately evaluate the suggested system. The suggested model got 98% accuracy. Designers contrast the findings of the suggested model with those of two of our own built models, along with the most recent benchmark algorithms

TABLE I
SUMMARY OF EXISTING SECURITY MECHAISMS

Authors	Proposed Framework	Method Used	Limitations
Otoum et al. [9]	Spider monkey optimization algorithm.	Denial of service, user-to-root attacks, probing attacks, and remote-to-local attacks are among the anomalous behaviors.	Provides untrusted scenarios.
Baniasadi, Sahba et al. [10]	Created a brand-new training technique to enhance the deep architecture's parameters.	Train the deep learning model that serves as our network intrusion detector.	Increases communication overhead during transmission.
Jothi B. et al. [11]	Suggested IDS.	Performed detailed evaluations on various IoT threat scenarios	Introduces delay in the analysis.
Nasir et al. [12]	Intrusion detection using deep learning	Methodology used combines SpiderMonkey, principal component analysis, information gain, and correlation attribute assessment.	Increases computational overhead.
S. T. Mehedi et al. [13]	CNN-based method for anomaly-based intrusion detection systems.	Suggested model demonstrates the capacity to recognize any potential incursion and unusual traffic patterns.	Has large communication and storage overhead.
Tanzila Saba et al. [14]	Deep learning model.	Intrusion samples were used to categorize the benign and malicious connections using the LSTM and FCN.	The communication and computation delay while transmitting the information.
M. S. A. Muthanna et al. [15]	Cutting-edge IoT-based collection.	Designers contrast the findings of the suggested model with those of two of our own models.	Incurs long delays while recognizing the behaviour of communicating devices.

to provide further assurance. The suggested model outperformed the competition in terms of assessment parameters including speed efficiency, detection accuracy, and sensitivity.

A. Problem Statement

After conducting a literature survey on intrusion detection systems (IDSs), several research gaps and areas for improvement have emerged. One major research gap is the need for more accurate and effective anomaly detection techniques, as traditional methods are often prone to generate false positives and false negatives. Another research gap is the need for IDSs to be more resilient and adaptable to changing attack patterns, as cyber attacks are constantly evolving and becoming more sophisticated. Recurrent neural networks (RNNs) such as Long Short-Term Memory (LSTM) networks have emerged as a promising approach for detecting anomalies in network traffic. However, the use of LSTMs for intrusion detection is not without its challenges. Issues such as overfitting and computational complexity must be addressed to fully realize the potential of LSTMs for IDSs. Addressing these research gaps will be crucial in the development of more robust and reliable IDSs that can provide adequate protection against the growing threat of cyber attacks.

III. PROPOSED WORK

This section describes an efficient and accurate decision making procedure by addressing the issue of overfitting while avoiding a high computational complexity. To address the challenges associated with using Long Short-Term Memory (LSTM) networks for intrusion detection, we propose an approach that combines a tree-based XGBoost algorithm with a bidirectional variant of LSTM. This hybrid approach aims to address the issues of overfitting and computational complexity that can arise with the traditional use of LSTM networks for intrusion detection. The reason for combining XGBoost and a bidirectional LSTM (BiLSTM) network is to address some of the limitations of traditional intrusion detection systems (IDSs) based on single machine learning models. XGBoost is a powerful tree-based algorithm that is widely used in various machine learning tasks, including anomaly detection. It has been shown to outperform many other traditional machine learning algorithms in terms of accuracy and computational efficiency. XGBoost can handle missing values, outliers, and noisy data, making it a robust and reliable method for intrusion detection. On the other hand, a BiLSTM network is a type of recurrent neural network that can effectively capture temporal dependencies in sequential data. In the context of intrusion detection, this means that a BiLSTM can learn to detect subtle patterns and anomalies in network traffic over time, which is crucial for identifying advanced persistent threats (APTs) and other sophisticated attacks. By combining XGBoost and a BiLSTM network, the proposed hybrid approach can leverage the strengths of both methods while mitigating their limitations. Specifically, XGBoost can be used to extract informative features from the raw network traffic data, and to provide an initial classification, while the BiLSTM can further refine the classification by taking into account the temporal dynamics of the data. This collaboration can help to improve the detection rate and reduce false positives, making the proposed approach more effective and efficient than traditional IDSs based on single machine learning models. Furthermore, in order to make an accurate decision-making, the intermediate devices must be fully trusted by the entities involved. To this aim, we have further used a direct trust-based scheme, along with BiLSTM and XGBoost methods, for overcoming the mentioned issues. The direct trust-based method is used to analyze the operation and activities of communicating devices in the network, assigning them some trust value based on their response.

A. Methodology

Let us now discuss about the process methodology for our proposed scheme.

- **Data Reading**: The first step in our proposed approach is to read the raw network traffic data. This data is obtained

from www.kaggle.com (a repository for research datasets).

- **Data Preprocessing**: Once the data has been read, it needs to be preprocessed and cleaned. This involves removing any irrelevant or redundant information, as well as handling missing values and outliers. The cleaned data is then normalized to ensure consistency across all features.
- **Feature Selection**: The next step is to select relevant features from the preprocessed data. To achieve this, we will use the Infinite Feature Selection (IFS) method. IFS iteratively selects the most informative features until a stop criterion is met, providing a subset of the most relevant features for the intrusion detection task.

The IFS algorithm is an iterative feature selection method that selects a subset of the most relevant features for the given task. It works by ranking the features based on their individual contribution to the classification performance, and then selecting the features with the highest rankings. The IFS Algorithm 1 starts with an empty feature set and iteratively adds features until a stopping criterion is met. At each iteration, the algorithm evaluates the performance of the classification model using a subset of the available features. The feature subset is then expanded by adding the next highest-ranked feature and the classification model is retrained.

Algorithm 1 Mathematical process of the IFS algorithm.

Require: Start with an empty feature set F = .
Input: Set i =0
Output: Final feature set F
Repeat until stopping criterion is met
Step 1: Select the i-th highest ranked feature f_i
Step 2: Add f_i to the feature set: $F = Ff_i$
Step 3: Train a classification model M on the feature set F
Step 4: Evaluate the performance of M using a validation set
Step 5: Increment i by 1

The general diagram of proposed approach is depicted in Figure 2 starting from information reading till evaluation of training model. The information reading, data repository, feature selection, and ranked feature components, are used to retrieve the information from their respective devices, where the raw information is further structured by handling the missing or empty values, including outliers. In addition, the infinite feature selection method is used over structured information that is further ranked from high to low based upon its behavioral reports. The Principle Component Analysis (PCA) is used to further process the information, along with the integration of the bidirectional variant of LSTM, along with XGBoost, that is further analyzed or evaluated using our training model.

This process continues until a predefined stop criterion is met, such as reaching a maximum number of features, or achieving a desired level of performance. The ranking of the features is determined by their individual contribution to the classification performance, which can be measured by a variety of metrics, such as the mutual information, correlation coefficient, or Gini index. For example, the mutual information between a feature and the class label measures how much information the feature provides about the class label.

- **Processing Dimensionality**: After the feature selection step using the Infinite Feature Selection (IFS) algorithm, the next step in the proposed approach is to apply Principal Component Analysis (PCA) to the selected features. PCA is a widely used technique for dimensionality reduction which aims to reduce the number of features in the dataset while retaining as much information as possible. Reducing the dimensionality of the feature space can also help to alleviate the problem of overfitting, which is a common issue in machine learning tasks, including intrusion detection. By reducing the number of features, PCA can help to prevent the model from memorizing noise or irrelevant patterns in the data, and improve its generalization performance on unseen data.
- **Model Development**: Once the relevant features have been selected, the proposed model is developed. This involves using a hybrid approach that combines a tree-based XGBoost algorithm with a bidirectional variant of LSTM. The XGBoost algorithm is used to provide a more robust and interpretable feature selection method, while the bidirectional LSTM helps to capture the temporal dependencies in the data.
- **Intrusion Detection:** Finally, the proposed model is used for the intrusion detection task. The trained model is evaluated on a test set of network traffic data, and its performance is measured in terms of accuracy, precision, recall, and F1-score. The results are then compared to those of existing intrusion detection methods to demonstrate the effectiveness of the proposed approach.

B. Direct Trust-based Scheme

In order to identify the intruders on a continuous basis, and to perform surveillance in the network for every time interval, a trust-based computational method must be introduced. For ensuring a secure and efficient decision-making process, while extracting useful information in the Web-of-Things context, we have used a direct trust-based method. The direct trust-based method is used to analyze the working and activity of communicating devices in the network, and assign them some trust value based on their response. The devices having a higher trust value are considered as trustful, and will always be included in the network for accurate decision making. The number of attributes involved when measuring the direct trust value are the following:

- Information Activeness: It is necessary to analyze the correlation of information sent by neighbouring nodes. In case a device is too actively involved in repeatedly sending the requests, or in transmitting the same information in the network, it can be considered as a threat.

$$IA = \frac{NI_{ij}(t)}{DI_{ij}(t) + NI_{ij}(t)} \tag{1}$$

Where, NI is defined as the amount information transmitted, while DI stands for duplicate information sent by the device.

Information Forwarding Delay: It is defined as the amount of time required by a device from the time it receives to the time it forwards the information in the network. The device having

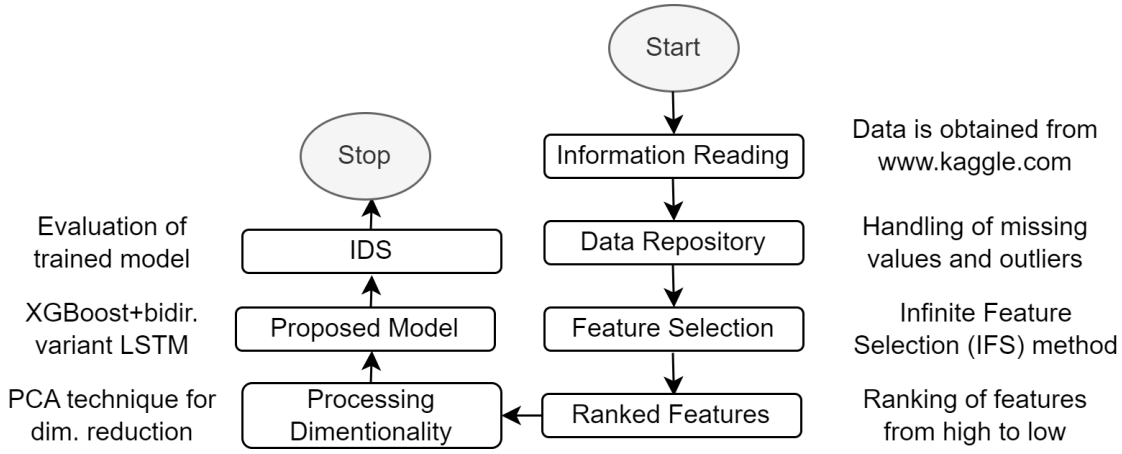

Fig. 2. Overall process of the the proposed framework.

malicious activity will drop or never forward the incoming packets to forwarding nodes.

$$IFD = \frac{TD_{ij}(t)}{AD_{ij}(t)} \tag{2}$$

Where, TD is defined as time delay for information forwarding for a specific device, and AD is average delay time to forward the information when considering all devices.

The direct trust is determined while computing these matrices in order to analyze the malicious behavior of network.

$$DTV = w1 \times IA + w2 \times IFD \tag{3}$$

Where $w1$ and $w2$ are the weight coefficients that are needed for transferring the information from one location to another.

Algorithm 2 Process of Accurate Decision-making in the Web of Things.

Require: A network N consists of D number of devices as $D = d_1, d_2, ...d_n$.
Input: Set i =0
Output: D is either legitimate or malicious
Repeat until stopping criterion is met
Step 1: Apply **XGBoost and IFS mechanism** for identifying the malicious activity in the network
Step 2: Train a classification model M on the feature set F
Step 3: Use **Direct Trust method** for improving the efficiency in the network
Step 4: Evaluate the performance of M using a validation set with computed trust values
Step 5: Increment i by 1

Algorithm 1 illustrates the mathematical process of IFS based upon a ranking system, from highest to lowest, by training its models. The trained models are further evaluated for validating the entire communication process. Furthermore, Algorithm 2 illustrates the complete process of proposed framework. Overall, our proposed approach aims to address the limitations of traditional LSTM-based intrusion detection methods, while providing a more accurate and efficient approach for detecting network intrusions. The simulation is conducted in *python*, and

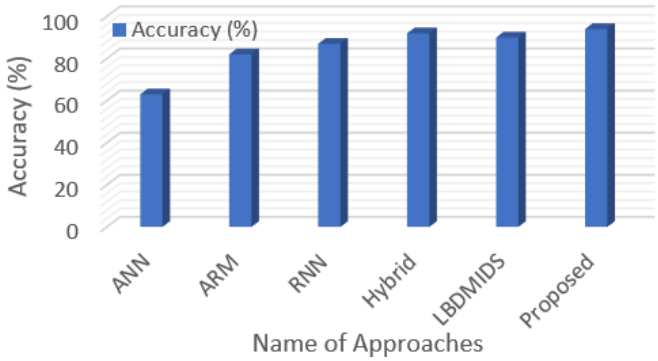

Fig. 3. Accuracy comparison of different algorithms.

results are obtained in terms of accuracy, precision, recall, etc.; these performance metrics are discussed next.

IV. Results and Discussion

This section provides a comprehensive analysis of the simulation results, and discusses the strengths and limitations of the proposed intrusion detection system. Such performance is evaluated based on several metrics, including accuracy, precision, recall, and F1 score. Simulation results are then compared with the existing intrusion detection approaches to assess the effectiveness of the proposed approach in detecting network intrusions.

The proposed mechanism is simulated using the CICIDS2017 dataset [16] by splitting it into 70% and 30% of training and testing, respectively; this is achieved by categorizing them into selected features.

We assess our proposed solution by comparing its performance against different existing solutions, such as artificial neural network, ARM, RNN, Hybrid model, and LBDMIDS algorithms. In particular, an analysis in terms of accuracy is presented in Figure 3; notice that the results for the proposed solution obtained via simulation achieve an accuracy of 99.11%. Such accuracy outperforms existing methods because of our improved mechanism for identifying malicious devices using BiLSTM. In addition, the dimensionality is further reduced due to the XGBoost mechanism. The individual algorithms accuracy

values are given below in Table II. The proposed approach performs better than existing approaches because the Principle Component Analysis (PCA) is used to further process the information along with the integration of the bidirectional variant of LSTM and XGBoost, that is further analyzed or evaluated using the training model. In addition, PCA is a widely used technique for dimensionality reduction which aims to reduce the number of features in the dataset while retaining as much information as possible. Reducing the dimensionality of the feature space can also help to alleviate the problem of overfitting, which is a common issue in machine learning tasks, including intrusion detection. By reducing the number of features, PCA can help to prevent the model from memorizing noise or irrelevant patterns in the data, and improve its generalization performance on unseen data. In addition, t proposed approach includes a process of data pre-processing, a feature selection step using the Infinite Feature Selection (IFS) algorithm, and the application of Principal Component Analysis (PCA) for dimensionality reduction. Finally, the proposed mechanism uses a direct trust-based method in order to accurately identify malicious devices in the network.

TABLE II
COMPARISON OF DIFFERENT SCHEMES IN TERMS OF ACCURACY AND COMPLEXITY.

S. No.	Algorithm	Accuracy	Complexity
1	ANN	63.97000	$n \times n$
2	ARM	86.45000	$n \times n$
3	RNN	95.70000	$n \times n$
4	Hybrid	98.70000	$nlogn$
5	LBDMIDS	96.60000	$nlogn$
6	Proposed	99.11336	n

As shown in table II, the traditional approaches ANN, ARM, RN, Hybrid and LBDMIDS achieve an accuracy of 63.9, 86.45, 95.7, 98.7, 96.6, respectively; these values are clearly below the proposed scheme, which achieves an improvement of 1.6% compared to the highest accuracy system (Hybrid). Then, in order to evaluate the effectiveness of the proposed intrusion detection approach, it is tested and compared with one of the existing methods, LBDMIDS, using several evaluation metrics such as precision, recall, and F1 score, as presented in Figure 4. The proposed mechanism outperforms existing methods because of adopting the direct trust-based scheme, which is effective at identifying malicious behaviour by any communicating device in the network.

These metrics are commonly used to measure the performance of machine learning models in the field of intrusion detection. By comparing the results obtained from the proposed approach with the results obtained from LBDMIDS, it is possible to determine whether the proposed approach is superior in terms of detecting network intrusions. This analysis provides a comprehensive understanding of the proposed approach and its effectiveness in comparison to existing methods. Simulation results demonstrate that the proposed intrusion detection approach outperforms the traditional scheme in terms of precision, recall, and F1 score. In particular, such scheme has a precision, recall, and F1 score of 0.96, while our proposed scheme shows a slight improvement of 0.3% in all these metrics. This indicates that the proposed approach is better able to detect network intrusions

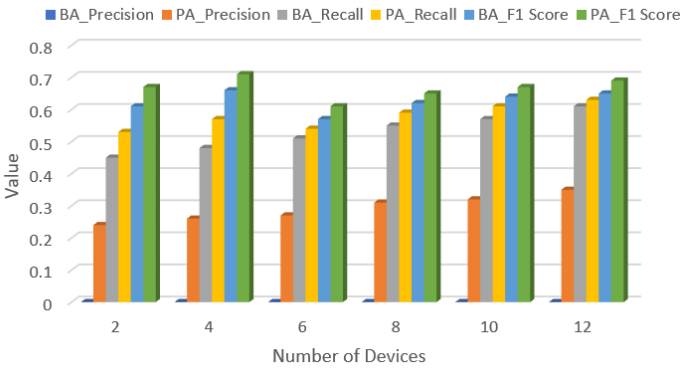

Fig. 4. Performance factors comparison.

compared to the traditional approach.

TABLE III
COMPARISON OF DIFFERENT PERFORMANCE FACTORS.

S. No.	Parameters	LBDMIDS	Proposed
1	Precision	0.96	0.991174
2	Recall	0.96	0.991134
3	F1-Score	0.96	0.991098

Finally, Table III details the individual parameter values obtained from the proposed scheme in comparison with traditional LBDM-based approach. Despite this improvement may seem small, it is quite significant in the context of intrusion detection, since even a small increase in performance can have a significant impact on network security. Therefore, the proposed approach seems promising at improving the overall effectiveness of intrusion detection systems. In addition, reducing the dimensionality of the feature space can also help to alleviate the problem of overfitting, which is a common issue in machine learning tasks, including intrusion detection.

The overall complexity and computational power of the proposed mechanism, along with other existing schemes, is presented in Table II.

Notice that the time complexity of the proposed mechanism outperforms others as it integrated two highly secure mechanisms that may further help at identifying, and thus reducing, the number of malicious devices in the network.

V. DISCUSSION

Socio-cybernetics is termed as the theoretical framework for understanding the cooperative and communicating behavior of devices in the context of evolution theory. By incorporating advanced technologies, the number of cyber threats encountered in the network can be reduced. In particular, to identify the intruders on a continuous basis, and to perform surveillance in the network following every time slot, a trust-based computational method is an adequate solution. For ensuring a secure and efficient decision-making process while extracting the useful information in the context of the Web of Things, we have used a direct trust-based method. In addition, the proposed AI-based learning model is able to successfully improve the accuracy and legitimacy of the decision-making process by completely understanding the behavior of each communicating device in

the network. In addition, the scientific impact of this paper is to provide an accurate decision-making framework that enables secure information transmission of intelligent devices in the network.

VI. Conclusion

Cybersecurity threats have become a growing concern in recent years, highlighting the need for effective Intrusion Detection Systems (IDS) to detect and prevent cyber attacks. In this work, we propose an IDS approach that combines a tree-based XGBoost algorithm and a bidirectional LSTM network to address the limitations of traditional approaches. The proposed approach includes a data preprocessing step, feature selection using an Infinite Feature Selection (IFS) algorithm, and the application of Principal Component Analysis (PCA) for dimensionality reduction. The performance of the proposed approach is evaluated using standard metrics - accuracy, precision, recall, and F1 score - surpassing the 99% threshold, which means it outperforms the existing method we used as reference for comparison: LBDMIDS. In fact, performance results demonstrate that the proposed approach outperforms such approach by offering a higher accuracy, and a slight improvement in terms of precision, recall, and F1 score.

Acknowledgments

This work is derived from R&D project PID2021-122580NB-I00, funded by MCIN/AEI/10.13039/501100011033 and "ERDF A way of making Europe".

References

[1] M. Alrowaily, F. Alenezi, and Z. Lu, "Effectiveness of machine learning based intrusion detection systems," in *Security, Privacy, and Anonymity in Computation, Communication, and Storage: 12th International Conference, SpaCCS 2019, Atlanta, GA, USA, July 14–17, 2019, Proceedings 12*. Springer, 2019, pp. 277–288.

[2] B. Alberti, "Archaeologies of ontology," *Annual review of anthropology*, vol. 45, pp. 163–179, 2016.

[3] B. Nour, M. Pourzandi, and M. Debbabi, "A survey on threat hunting in enterprise networks," *IEEE Communications Surveys & Tutorials*, 2023.

[4] A. Khraisat, I. Gondal, P. Vamplew, and J. Kamruzzaman, "Survey of intrusion detection systems: techniques, datasets and challenges," *Cybersecurity*, vol. 2, no. 1, pp. 1–22, 2019.

[5] A. Grob, P. Hagmann-von Arx, A. Jaworowska, A. Matczak, and D. Fecenec, *ids 2*. Hogrefe, 2018.

[6] R. F. Geyer *et al.*, "Sociocybernetics," in *Cybernetics and applied systems*. CRC Press, 2018, pp. 95–124.

[7] R. F. Geyer and J. Van der Zouwen, *Sociocybernetics: An actor-oriented social systems approach Vol. 1*. Springer, 2014, vol. 1.

[8] B. B. Zarpelão, R. S. Miani, C. T. Kawakani, and S. C. de Alvarenga, "A survey of intrusion detection in internet of things," *Journal of Network and Computer Applications*, vol. 84, pp. 25–37, 2017.

[9] Y. Otoum, D. Liu, and A. Nayak, "Dl-ids: a deep learning–based intrusion detection framework for securing iot," *Transactions on Emerging Telecommunications Technologies*, vol. 33, no. 3, p. e3803, 2022.

[10] S. Baniasadi, O. Rostami, D. Martín, and M. Kaveh, "A novel deep supervised learning-based approach for intrusion detection in iot systems," *Sensors*, vol. 22, no. 12, p. 4459, 2022.

[11] B. Jothi and M. Pushpalatha, "Wils-trs—a novel optimized deep learning based intrusion detection framework for iot networks," *Personal and Ubiquitous Computing*, vol. 27, no. 3, pp. 1285–1301, 2023.

[12] M. Nasir, A. R. Javed, M. A. Tariq, M. Asim, and T. Baker, "Feature engineering and deep learning-based intrusion detection framework for securing edge iot," *The Journal of Supercomputing*, pp. 1–15, 2022.

[13] S. T. Mehedi, A. Anwar, Z. Rahman, K. Ahmed, and R. Islam, "Dependable intrusion detection system for iot: A deep transfer learning based approach," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 1, pp. 1006–1017, 2022.

[14] T. Saba, A. Rehman, T. Sadad, H. Kolivand, and S. A. Bahaj, "Anomaly-based intrusion detection system for iot networks through deep learning model," *Computers and Electrical Engineering*, vol. 99, p. 107810, 2022.

[15] M. S. A. Muthanna, R. Alkanhel, A. Muthanna, A. Rafiq, and W. A. M. Abdullah, "Towards sdn-enabled, intelligent intrusion detection system for internet of things (iot)," *IEEE Access*, vol. 10, pp. 22 756–22 768, 2022.

[16] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization." *ICISSp*, vol. 1, pp. 108–116, 2018.