International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11
Article Received: *10 September 2023* ***Revised:*** *20 October 2023* ***Accepted:*** *30 October 2023*

An Efficient Fuzzy Based Multi Level Clustering Model Using Artificial Bee Colony For Intrusion Detection

Battini Sujatha[1], **Dr.Sammulal Porika**[2]
[1]Telangana Social Welfare Residential Degree College for women, Scholar of Computer science and Engg, JNTUH, battinisujata@gmail.com
[2]Professor Department of computer science and Engg, JNTUH College of Engineering, sam@jntuh.ac.in

Abstract: Network security is becoming increasingly important as computer technology advances. One of the most important components in maintaining a secure network is an Intrusion Detection System (IDS). An IDS is a collection of tools used to detect and report network anomalies. Threats to computer networks are increasing at an alarming rate. As a result, it is critical to create and maintain a safe computing environment. For network security, researchers employ a range of technologies, including anomaly-based intrusion detection systems (AIDS). These anomaly-based detections face a major challenge in the classification of data. Optimization algorithms that mimic the foraging behavior of bees in nature, such as the artificial bee colony algorithm, is a highly successful tool. A computer network's intrusion detection system (IDS) is an essential tool for keeping tabs on the activities taking place in the network. Artificial Bee Colony (ABC) algorithm is used in this research for effective intrusion detection. More and more intrusion detection systems are needed to keep up with the increasing number of attacks and the increase in Internet bandwidth. Detecting developing threats with high accuracy at line rates is the prerequisite for a good intrusion detection system. As traffic grows, current systems will be overwhelmed by the sheer volume of false positives and negatives they generate. In order to detect intrusions based on anomalies, this research employs an Efficient Fuzzy based Multi Level Clustering Model using Artificial Bee Colony (EFMLC-ABC). A semi-supervised intrusion detection method based on an artificial bee colony algorithm is proposed in this paper to optimize cluster centers and identify the best clustering options. In order to assess the effectiveness of the proposed method, various subsets of the KDD Cup 99 database were subjected to experimental testing. Analyses have shown that the proposed algorithm is suitable and efficient for intrusion detection system.

Keywords: Network Security, Intrusion Detection System, Optimization, Artificial Bee Colony, False Alarm Rate.

1. Introduction

After the failure of more traditional firewalls, access control, and encryption tactics to detect intrusion in the systems, the intrusion detection framework has become an essential component due to the hidden vulnerabilities included in programming applications [1]. Because of this, a separate barrier for assuring frameworks known as an Intellectual Intrusion Detection System (IIDS) is necessary regardless of the avoidance techniques used. Since James Anderson's 1980 presentation of intrusion discovery, intrusion location has played a crucial role in addition to the firewall [2]. Since the growth of attacks has inspired programmers to implement intrusion identification arrangement, this has led to the development of IDS. The parallel development of the hacking software specifically targeted the IDS [3]. Giving information a defense component isn't without its difficulties [4]. The development of expert IDS is a challenge, although researchers are making good efforts [5].

In contrast to IDS, firewalls cannot foresee an attack, but in the case of an intrusion, the administrator may identify the attack and take precautions to prevent it from happening again. An intrusion is indeed a method that compromises the security [6], integrity [7], flexibility, and availability of a system's resources [8]. Alerts are sent to the executive once the system has screened and analyzed customer and system traffic, verified framework arrangements, and identified vulnerabilities [9]. The purpose of an IDS is to protect a computer system from malicious intrusion [10]. Due to the numerous threats to the stability and safety of web-connected systems, IDSs are an essential component of the security architecture.

Profiles of client behavior over the system are incorporated into IDS, and using these examples, the system is able to identify intruders and react accordingly [11]. Due to the constant need to handle a large volume of data, which slows down the IDS's planning and testing phases and reduces its identification rate [12], the element selection process has emerged as a central topic in IDS. Abuse and abnormal recognition techniques are two forms of finding used in IDS [13]. Client-side information extraction [14], or information mining, entails sifting through a client's massive database or information warehouse in search of relevant information. Information can be described in terms of concept, law, and model [15]. The idea behind using data mining technique is to assist leaders in classifying data as either useful or irrelevant. Predicting unknown or future characteristics of another element can be done with the help of mining techniques including arrangement, regression, and divergence finding [16]. The host-based IDS monitors the host computer for suspicious activity. While effective, the abuse localization method has significant limitations [17].

An IDS's crucial function in network security is to identify malicious network intrusions. A synchronous optimization strategy is presented for multi level clustering and parameter calculation in the novel approach [18], which helps address the issue of redundant network data and inadequate model parameters. Moreover, the concept of clustering serves as a replacement for ABC's greedy selection property, and two distinct selection probability formulas are supplied for the algorithm's early and later iterations [19]. In terms of detection accuracy and false negative rate, the experimental findings reveal that the newly proposed method outperforms the state-of-the-art methods [20]. IDS defence policies are heavily influenced by how attackers tend to cluster. Finding the right number of clusters has proven to be a difficult problem in cluster analysis. As part of an IDS framework, an automatic clustering approach based on fuzzy set is proposed in this research. As a result of automated clustering methods, identification of groups of data can be quickly done that share the most characteristics with each other while having the fewest in common with other groups [21].

Statistical methodologies, neural networks, predictive pattern generation, association rules, and other methods are just some of the various ways IDS can be implemented [22]. Swarm intelligence techniques take into account one of the newly proposed methods to construct clustering and classification models to differentiate between normal and abnormal behaviour, allowing for the development of efficient and robust IDS. This research proposes a novel approach to the development of IDS, combining the fuzzy set with the Artificial Bee Colony (ABC) optimization technique. Using the NSL-KDD dataset, how well the suggested IDS can divide data into two categories is analyzed. The NSL-KDD dataset is a popular choice for evaluating intrusion detection systems. The ABC model is shown in Figure 1.

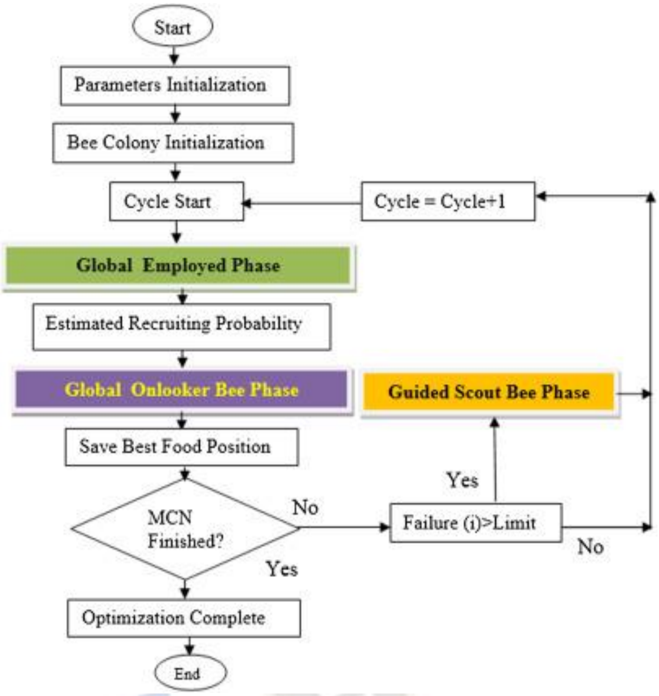

Fig 1: ABC Optimization Process

The use of fuzzy logic in IDS has the benefit of enabling the representation of concepts that may be assigned to more than one category, or of overlapping categories. Unlike traditional set theory, where objects are either fully or not at all members of a category, fuzzy set theory allows for partial membership. The foundation of fuzzy logic can be traced back to the study of probabilities. Fuzzy logic facilitates word-based computing and gives assistance in dealing with uncertainty during the development of expert IDS systems [23]. As it can deal with imprecise and uncertain situations, it has become an integral aspect of machine learning. Intruders can be spotted with the use of a wide variety of machine learning methods. Models of detection are constructed by the algorithms [24]. The concept of truth values ranging from true to false is handled in Fuzzy Logic, a superset of Boolean Logic. Fuzzy logic allows for intermediate values beyond just true and false, while boolean logic only allows for either.

Clustering is a useful technique for several fields, including data mining, statistical analysis, data compression, & vector quantization, since it helps organize large amounts of information into smaller, more manageable chunks called clusters. The function of clustering is really to organize data into groups called clusters in which similarities between cluster members are maximized while similarities between clusters are minimized. There are two main types of clustering algorithms, partitional clustering and hierarchical clustering [25]. Through a series of partitions, data objects can be moved from single clusters to a cluster containing all persons, or vice versa, using hierarchical clustering. Prototype-based clustering methods, in which the cluster centers serve as the representation for each cluster and the objective function is the total distance from the pattern to a cluster center, are among the most often used partitional clustering techniques. In order to detect intrusions based on anomalies, this research employs an Efficient Fuzzy based Multi Level Clustering Model using Artificial Bee Colony. The intrusion detection clustering process is shown in Figure 2.

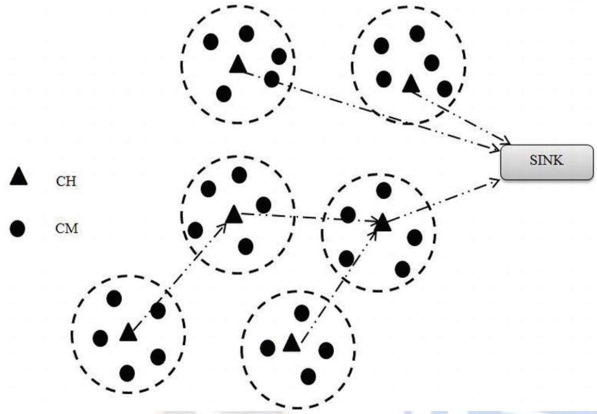

Fig 2: IDS Clustering Model

The rapid growth in internet users over the past few years underlines the importance of implementing a reliable security infrastructure. The three aspects of computer security Confidentiality, Integrity, and Availability (CIA) can be compromised by malicious penetration or attack on computer and information databases. In addition, standard security measures like firewalls, user authentication systems, and data encryption are deployed to keep computers and networks safe [26]. In addition to the aforementioned tools, Intrusion Prevention System (IPS) and Intrusion Detection Systems (IDS) as security instrumentation in the network layer [27] can take on the deterrence tasks to detect and prevent the harmful activities if the traditional firewall is unable to do so. Both hardware appliances and software products exist that can do automated network traffic analysis and send security alerts to the administration desk [28]. In addition, unlike firewalls [29], which only evaluate a small subset of a packet's fields, IDSs need to investigate everything from IP addresses and flags to the packet's optional fields.

2. Literature Review

By design, the controller area networks (CAN), which is still the most popular in-vehicle network today, does not offer any sort of security or authentication to its users. Modern vehicles are vulnerable to cyberattacks because they are constantly connected to the internet via Bluetooth, Wi-Fi, and mobile radio, all of which can be accessible from the outside world. For this reason, it is of the utmost importance to improve vehicle security by identifying and preventing cyberattacks. In this research, Freitas De Araujo-Filho et al. [2] introduced a new unsupervised intrusion prevention systems (IPS) for automotive CANs that can identify and prevent assaults without requiring access to information that is normally only available to car makers or modifying the architecture of the ECUs. The author examined the accuracy and efficiency of two machine learning techniques in detecting fuzzing and spoofing while using as little data as possible. The faster detection can begin and the earlier attacking frame can be recognised, the minimal data bytes are needed.

Multiple clustering techniques have been proposed in recent years to mine information from rapidly-growing data streams produced by diverse sets of hardware and software. The ability to deal with outliers and capture clusters of arbitrary forms make density-based techniques stand out as a particularly appealing type of algorithm. Streaming environments present their own unique set of difficulties that must be overcome; for example, data streams may be infinite in size and be subject to idea drift, which is defined as a change in the fundamental data creation process over time. Bechini et al. [3] presented TSF-DBSCAN, an unique fuzzy clustering approach for streaming data, which uses density-based clustering of application with noise. One of the most widely used density-based clustering methods, TSF-DBSCAN is a refinement of the original DBSCAN algorithm. The distance boundary that defines a neighbourhood of an object is fuzzy, and TSF-DBSCAN introduces fuzzy logic to reflect this uncertainty. Thus, TSF-DBSCAN finds clusters with ambiguous, overlapping borders. TSF-DBSCAN is equipped to change with the times to a fading model that reduces the importance of objects as they go farther back in time. The model is integrated in a two-stage technique to ensure computational and memory efficiency; in the first stage, continuously arriving items are arranged in correct data structures that are then utilised in the second stage to identify a fine-grained partition.

In recent years, fuzzy clustering has emerged to be one of the most discussed methods for organising data. Unfortunately, its computational complexity cost has created a bottleneck, preventing it from being used in large-scale situations. Noise also has an impact on the majority of fuzzy clustering algorithms. A new fuzzy clustering algorithm, fast fuzzy grouping based on anchoring graph (FFCAG), is presented by Nie et al. [6] to handle these complications. So that prior information of users may be further leveraged to increase clustering performance, the FFCAG algorithm combines the two processes of user-based similarity graph creation and membership matrix training into a unified framework. To be more specific, FFCAG first builds a user-based analyze similarities and differences using a neighbour assignment technique that does not require any parameters. And then, in a radical departure from standard fuzzy clustering methods, it creates a quadratic model to discover the users membership matrices. To further improve the accuracy of the clustering results, a novel balanced regularization term is included into the goal function. When it comes to solving the proposed approach, users finally turn to an alternative optimization algorithm that is guaranteed to converge.

The use of a FRBM, or fuzzy rule-based model, to characterise a system is a valid method of doing so. However, in practise, there may be cases where the user of a scheme only owns the output or input information of that framework, and where, out of respect for the privacy of users data, users unable to access the additional information required to construct the FRBMs. Hu et al.[7] devised a method for dealing with this difficulty so as to satisfy the particular privacy considerations during the modelling process, as this circumstance has not been completely realized and explored previously. This research applies the idea and algorithm of collaborating fuzzy clustering (CFC) to the task of identifying FRBMs that characterise MIMO and MISO systems, respectively. When inputs were unable to be gathered and used together, FRBMs may still be constructed with the cooperation between output and input areas based on its structural data. Surprisingly, on top of this main goal, the collaboration mechanism provides a novel way for the output and input spaces of a system to exhaustively share, exchange, and utilise the information about the structure between each other, resulting in their more pertinent frameworks that ensure good model performance compared to higher intrinsic motivation by some state-of-the-art modelling strategies.

Due to the exponential growth of intelligent mechanism, it is necessary to continuously modify and personalise network traffic patterns. Software-defined networking is a viable networking solution for intelligent network monitoring because of its high agility and fully programmable, which allow it to dynamically control industrial networks. However, network attacks can reduce manufacturing output and potentially cause accidents in a software-defined industrial network architecture. To this end, Hu et al.[11] proposed a deep learning-based, single-class intrusion detection technique (DO-IDS) for protecting industrial networks. To begin with, DO-IDS periodically gathers flow data from industrial network traffic in order to generate network information features. Then it employs a dimension reduction technique based on deep learning to eliminate unnecessary data points. In addition, the unusual numbers of the network information features are determined with the help of a one-class detector built on deep learning.

In this research, Dutt et al.[12] examined the immunological concept and apply it to the field of network intrusion detection. The primary focus of this research is on utilising network traffic monitoring, logging, and detection methods to identify potential network intrusions. The suggested model takes into account both the innate and adaptive layers of the real Immune System (IS), making it a close approximation of the real thing. As the first line of defence in an IDS, the current work suggests Statistical Modelling based Detection Methods (SMAD). In its role as an interface to the Innate Immunity System (IIS), it records a network's early traffic in order to identify potential security flaws. As a second layer of defence, Adaptation Immune-based Anomalies Detection (AIAD) has been suggested for identifying anomalous characteristics in network packets. It mimics the immune system's adaptive response by accounting for T-cell and B-cell activation. Effective intrusion detection is achieved by capturing features from the header and payload. Both real-time internet traffic and the gold-standard intrusion detection dataset KDD99 and UNSW-NB15 have been the subject of experimental research. Using historical data and live traffic, the SMAD model achieves a 97.04% TPV and an approximate 97% TPV. When highly suspect traffic is identified by the SMAD model, it is sent to the AIAD model for additional vulnerability testing.

Due to the proliferation of the cloud and the internet storage, the IDS in Software-Defined Networks is gaining more and more acclaim. Institutions with many users who rely on cloud services need this technology desperately. There are a number of security features offered by the Intrusion Detection System, but its performance is not up to par with those of competing systems in large industrial networks. Most current methods for implementing a security system rely on a central processing unit and a plethora of features. As a result, the controller and the OpenFlow switches become overloaded, resulting in subpar performance. So, there are concerns that arise from using existing methods, especially on huge networks. Furthermore, strengthened security applications increase the network's dependability. Janabi et al.[13] introduced a new model for Intrusion Detection Systems, based on a review of the current literature, which addresses problems with system overload and poor performance by using decentralised processing and exchanging data via a separate channel. To decrease the amount of data being sent over the channels, the model makes use of a suitable method for selecting features to filter down the extracted features. Since it is a quick classifier, the Naive Bayes method has also been used for flow categorization. Using the Network simulation emulator, which simulates a functional network, authors successfully, constructed the framework.

In this paper, Bulajoul et al.[14] described experimental research that demonstrates current NIDPSs that have significant limitations in detecting or blocking growing unwanted traffic and various hazards in high-speed situations. It demonstrates how the NIDPS's performance might suffer in the face of fast and heavy malicious traffic, in terms of dropped packets, unanalyzed packets that remain in the queue, and undetected or prevented traffic. In order to improve the efficiency of intrusion prevention and detection systems, a brand-new quality-of-service (QoS) framework has been developed. A unique QoSconfig in a multi-layer switching has been suggested and analysed in the study to organise packets/traffic and to speed up packet processing utilising parallel approaches. The author put the new architecture through its paces with varying volumes of traffic and a variety of workloads.

3. Proposed Method

There is always the risk of unauthorised access and misuse of private data transmitted over the internet or any other network. The integrity and confidentiality of data in a network context are particularly vulnerable to intrusions [29]. Any attempt to jeopardise the safety of a resource is known as an incursion. Network destabilisation, privileged file access abuse, and improper programme usage are all examples. When it comes to protecting data and networks, IDS are now indispensable. Intrusion detection's primary objectives are the automatic monitoring of network activity, the identification of malicious attacks [30], and the establishment of a sound security architecture for digital networks. Because of its capacity to express imperfect forms of thinking in contexts where decisive actions must be taken in uncertain settings, fuzzy logic is a good fit for intrusion detection systems.

Intrusions increased and evolved steadily in tandem with the speed at which technology was advancing. Day after day, successful attacks lead to massive sums of money being lost, privacy being compromised, and information being transferred illegally. Intruders provide a variety of threats to computer systems, data, and other networked resources. There is the user-to-root intrusion, the purpose of which is to get complete control over a system or network by gaining administrative privileges. Another sort of attack is called a probing intrusion, and its purpose is to scan computer and network systems to find vulnerabilities. The preceding forms of incursion could be made into pre-requisites for the variety of attacks. This form of attack is likely to use up a lot of system resources, which will force the suspension of some services for legitimate users. About 60% of all cyber attacks are distributed as denial-of-service assaults. When it comes to stopping and thwarting intrusions, IDS is one of the best tools at the disposal. Since attacks can cause significant monetary loss and privacy violations, IDS has become an essential part of any network security strategy. Different obstacles must be overcome on the route to a fully functional IDS. The use of either/or choices in detection methods is one such difficulty. There was a limit to the effectiveness of the usual IDS detection procedures. When dealing with boundary issues, the fuzzy system has a number of advantages. It also provides the security engineer with a more comprehensible presentation of the intrusion detection degree level.

As the internet has evolved, the importance of keeping data safe has also increased. Computer security, as most internet users agree, is around measures taken to keep the internet safe. The term intrusion detection is used to describe the various methods employed to uncover malicious activity on a network or within an information system. This can be done with the help of specialised software designed for the sole aim of identifying suspicious online behavior. Many scholars have previously put forth frameworks for creating effective intrusion detection systems.

Integral Detection Systems are increasingly important components of system defense, and are used to detect malicious or suspicious behavior in a computer network. Because it is technically impossible to put up a system with no weaknesses, research into methods of intrusion detection has grown in importance. As a branch of many-valued logic or probabilistic logic, fuzzy logic is concerned with approximate reasoning as opposed to strict logical proofs. In contrast to the fixed truth values of true and false that characterize binary sets, variables in fuzzy logic can take on values between zero and one. Since the truth value might be anywhere from entirely true to absolutely untrue, fuzzy logic has been expanded to account for this grey area. The proposed model framework is shown in Figure 3.

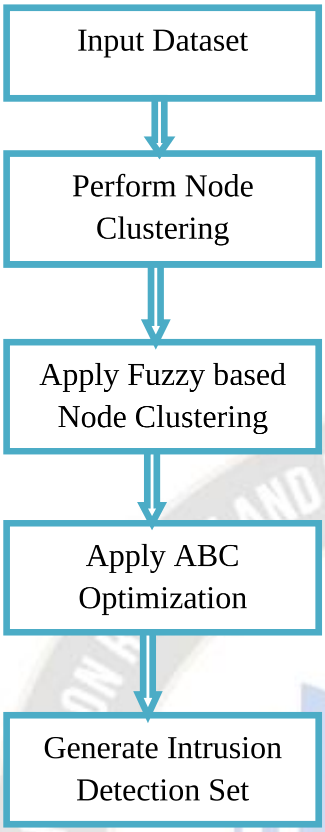
Input Dataset
Perform Node Clustering
Apply Fuzzy based Node Clustering
Apply ABC Optimization
Generate Intrusion Detection Set

Fig 3: Proposed Model Framework

In addition, when linguistic variables are employed, certain operations can be applied to control these levels. The foundation of fuzzy logic is a user-supplied set of rules written in natural language. Fuzzy systems translate these guidelines into their numerical counterparts. The work of the system designer and the computer is made easier, and more accurate models of the behaviour of systems in the real world are produced as a result. Fuzzy logic's ease of use and adaptability are two further advantages. Fuzzy logic is capable of modelling nonlinear functions of arbitrary complexity and solving issues with inaccurate or missing data. When compared to systems relying solely on signature matching or the more traditional method of pattern deviation detection, fuzzy logic has shown promise in the field of intrusion detection. In situations when precise forms of reasoning are not appropriate, such as intrusion detection, fuzzy logic can be used to indicate a degree of uncertainty. In order to detect intrusions based on anomalies, this research employs an Efficient Fuzzy based Multi Level Clustering Model using Artificial Bee Colony. A semi-supervised intrusion detection method based on an artificial bee colony algorithm is proposed in this paper to optimize cluster centers and identify the best clustering options.

Algorithm EFMLC-ABC

{

Input: Intrusion Detection Set {IDset}

Output: Intrusion Prediction Set {IPSet}

Step-1: Initially nodes in the network will be registered in the network and the network node analysis will be performed to perform clustering. The node analysis and clustering is performed as

$$NodeRegCls[N] = \sum_{i=1}^{N} \frac{getNode(i)}{addr(i)} + addr(i+1) + Tinst + Th$$

Here addr is used to access the node address and Tinst is the time instant and Th is threshold value maintained for the entire network.

Step-2: Node monitoring in a fuzzy set can be thought of as having a continuous scale of transmission. A Node monitoring function characterizing such a set gives each item a score among zero and one based on its degree of data transmission in the set. The cluster node monitoring using fuzzy feature set is performed as

$$Nfuzzy\big(Node(N)\big) = \frac{\max\left(NodeRegCls(i)\right)}{\mathrm{len(NodeRegClas)}} + \sum_{i=1} getFeat(IDset(i)) - \frac{maxVal(i+1) - minVal(i)}{count(IDset)}$$

Here λ is the mode used to extract min attribute based nodes.

Step-3: The data from the various nodes is gathered by the cluster head. When it has acquired enough information, it passes it on to the base station via an intermediary cluster-head. The cluster head will monitor the malicious actions in the network and the cluster head selection and analysis is performed as

$$\mathrm{ChNode(Nfuzzy(N))} = \sum_{\mathrm{i}=1}^{\mathrm{N}} \frac{2\pi}{\mathrm{maxVal(i)}} + (\max\left(\mathrm{Nfuzzy(i+1)}\right) - \max\left(NodeRegCls(i)\right))^2 + \delta + \frac{\tau}{\max\left(\mu\right)}$$

Here δ is the model that considers node having maximum computational power and μ is the node used that has maximum energy level.

***Article Received:** 10 September 2023 **Revised:** 20 October 2023 **Accepted:** 30 October 2023*

Step-4: The Artificial Bee Colony (ABC) algorithm, a method of optimization that mimics the foraging activity of honey bees, has found widespread use in solving a wide range of real-world issues. In the intrusion detection process, ABC optimization is applied. The ABC is initialized using the search probability of the population bees. The process is performed as

$$PopSr(L) = \sum_{i=1}^{N} \frac{\lambda(i+1_2)*\lambda(i+1|i)}{\sum_{i=1}^{N} \frac{\lambda(\delta|\mu)}{\lambda(Th)}}$$

$$FitV\big(PopSr(L)\big) = \frac{\lambda(i+1_2)*\lambda(i+1|i)}{1+\frac{1}{\lambda}+\sum_{i=1}^{N} \frac{\lambda(\delta|\mu)}{\lambda(Th)}}$$

$$FitVal(L) = \begin{cases} \frac{1}{1+FitV(L)} & if\ FitV \geq PopSr \\ FitV & if\ FitV < PopSr \end{cases}$$

Step-5: An IDS is a software designed to keep monitoring computer network for any signs of infiltration or policy breaches. The standard procedure for dealing with any kind of intrusion activities or violation is to notify an administrator or use an event and security information system to centralize collect the relevant data. The final intrusion prediction set is performed as

$$IDPred[M] = \sum_{i=1}^{N} \frac{\max\big(Chnode(i)\big) + \max\,(FitVal(i))}{\max\,(Nfuzzy(i))} - \min\Big(Chnode\big(NodeRegCls(i)\big)\Big) + \frac{\max\,(\mu)}{\delta}$$

}

4. Results

Recently, machine learning techniques have become increasingly relevant in the context of solving network security problems, such as network intrusion detection. A high detection rate and low false alarm rate are two primary goals of intrusion detection systems. As intrusion patterns and features are constantly evolving, learning techniques that rely on classification are generally inefficient for detecting intrusions. As a result, unsupervised learning strategies have been scrutinised more closely for use in network intrusion detection. The network security professional can use clustering results to better categorise network traffic data as either normal or suspicious. Due to the sheer volume of available network traffic audit data, the expert-based labelling procedure of all records is extremely laborious, time-consuming, and costly. Further, mistakes can be introduced while classifying a big amount of network traffic information. When data is grouped according to similarity, experts can more easily assign labels. If the clusters obtained are relatively clean, the expert can label all data examples in a cluster at once by examining the common properties of the cluster, without having to evaluate each data instance individually. To be 100% pure, a cluster must only ever have data instances that fall under the normal or attack category. By monitoring data in real time, intrusion detection systems can pinpoint and identify suspicious behavior on a network in an effort to prevent or mitigate further damage. In order to detect intrusions based on anomalies, this research employs an Efficient Fuzzy based Multi Level Clustering Model using Artificial Bee Colony (EFMLC-ABC). The proposed model is compared with the traditional network intrusion-detection algorithm based on Dynamic Intuitionistic Fuzzy Sets (DIFSs) model.

In its most basic form, a network cluster is a network of connected SG nodes working together to complete a single task. The term node clustering is used to describe both the computational challenge of extracting fully connected but relatively separated SG networks from a system and the corresponding set of strategies and methods for doing so. The node clustering is performed that groups similar kind of node attributes. The node clustering time levels of the proposed and existing model are shown in Figure 4.

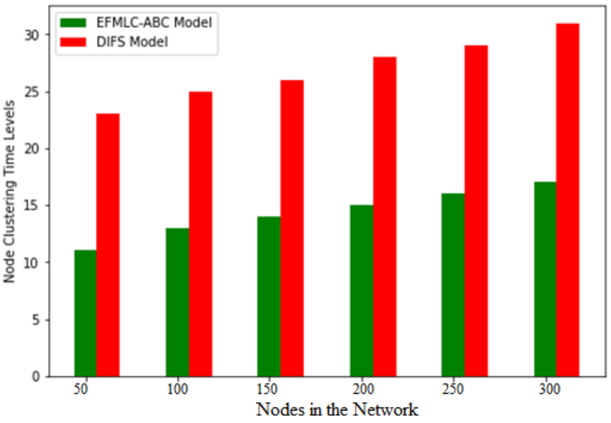

Fig 4: Node Clustering Time Levels

Clustering is the process of splitting nodes into subsets wherein the nodes in clusters are considered. When applied to a SG that has been hierarchically categorized, multilevel clustering can concurrently divide node within each group and reveal grouping patterns within categories. Groups, or clusters, are created by the application of node cluster analysis, with the goal of ensuring that the information within a cluster are as possible similar while the observation belonging to other clusters are as unlike as possible. The node multi level clustering accuracy levels are shown in Figure 5.

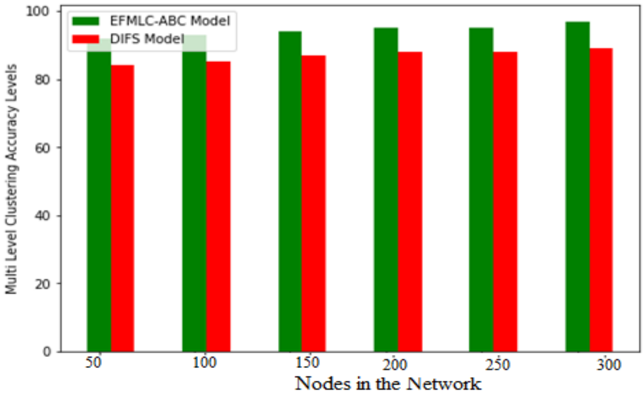

Fig 5: Multi Level Clustering Accuracy Levels

When using automated fuzzy clustering, a single node may be assigned to monitor two or more different groups. The method determines how much each node is trusted and the how the adjacent node is trusted. The fuzzy based multi level clustering time levels of the existing and proposed models are shown in Figure 6.

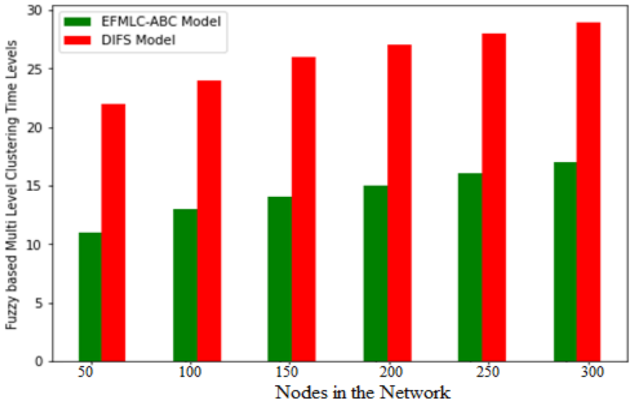

Fig 6: Fuzzy based Multi Level Clustering Time Levels

An instance of a degree of node trust is found in the fuzzy set. Each member of the set is given a score between zero and one using a membership function. It has been demonstrated that fuzzy set theory is a helpful tool for describing situations where precise or clear data is lacking. Fuzzy sets are able to deal with this kind of issue by assigning a degree of trust to a SG node. The fuzzy set generation accuracy levels of the traditional and proposed models are shown in Figure 7.

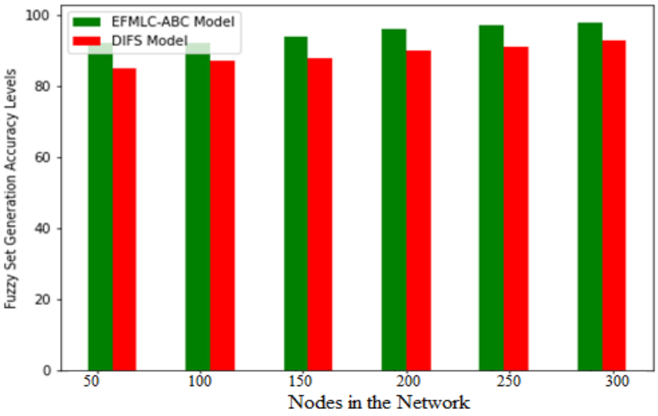

Fig 7: Fuzzy Set Generation Accuracy Levels

When using ABC Optimization, a population-based algorithm, the location of a food source stands in for a potential solution to the optimization issue, and the quantity of nectar that source produces stands in for the quality of intrusion detection of the solution. There are as many working bees as there are total population solutions in the process of intrusion detection. The ABC optimization process accuracy levels in the proposed and existing models are shown in Figure 8.

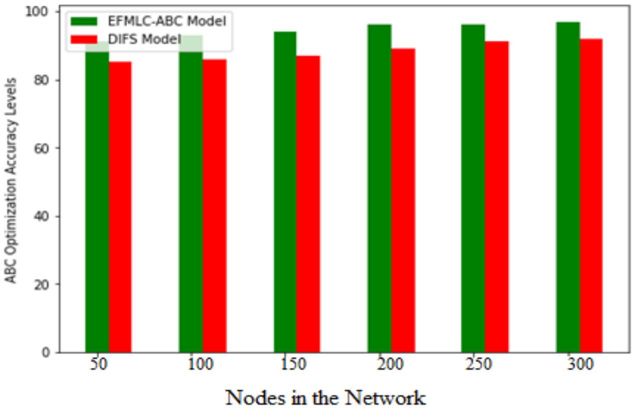

Fig 8: ABC Optimization Accuracy Levels

IDS will keep tabs on network's activity, examine it for indicators of infiltration, and sound the alarm if it detects anything out of the ordinary. Meanwhile, traffic is unaffected. Additionally, traffic is monitored by an intrusion detection system. The proposed model verifies the intrusions in the network with optimization models and fuzzy set. The intrusion detection rate of the proposed model is high than the traditional models that are depicted in Figure 9.

***Article Received:** 10 September 2023 **Revised:** 20 October 2023 **Accepted:** 30 October 2023*

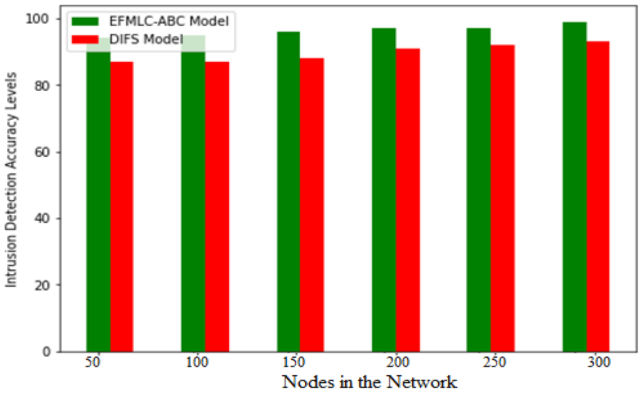

Fig 9: Intrusion Detection Accuracy Levels

5. Conclusion

Through this research efforts, an IDS that makes use of an original classification technique is proposed. The detection accuracy in the suggested study is enhanced by using a novel clustering approach. The number of clusters is initially formed using hierarchical clustering, and the ideal clusters are subsequently identified using the ABC optimization technique. After the optimal clusters have been chosen, the features necessary to develop the classifier are extracted from them. Finally, the best features are chosen using the ABC optimization technique and introduced to the classifier, which determines whether a given set of data is normal or abnormal based on how closely the training and testing sets of data match. Clustering the benchmarking classification problems is employed for classification in this work, with the help of Artificial Bee Colony method, which is a novel, straightforward, and resilient optimization technique. In this research, a novel IDS that is the product of combining fuzzy set and ABC optimization is propsoed. The fuzzy clustering technique generates many training subsets. The information in each clusters share a significant amount of similarity while also being significantly distinct to those of other clusters, making clustering a useful categorization tool. The studies validate the feasibility of applying the Artificial-Bee-Colony method to clustering for classification. Several metrics, including detection accuracy, false alarm rate, and accuracy, are used to evaluate the proposed model's performance based on experimental results, including reliability evaluation utilising functions and the behavior of the anomaly detector. In order to compare the proposed method to those that have come before it, various empirical evaluations of these parameters are conducted for various classical algorithms on various datasets. The validity of the proposed method in all of these respects is attested to by the high calibre of the solutions it produces. Future research could focus on a number of other questions, such as comparing the outcomes of clustering using ABC algorithm to those of other algorithms. Also future work can compare the performance of the proposed algorithm to that of another evolutionary algorithm, also using a hybridization scenario. In addition, a case study would be examined to see if the suggested method outperforms alternative hybridizations.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest

References

[1]. J. Xie, H. Wang, J. M. Garibaldi and D. Wu, "Network Intrusion Detection Based on Dynamic Intuitionistic Fuzzy Sets," in IEEE Transactions on Fuzzy Systems, vol. 30, no. 9, pp. 3460-3472, Sept. 2022, doi: 10.1109/TFUZZ.2021.3117441.

[2]. P. Freitas De Araujo-Filho, A. J. Pinheiro, G. Kaddoum, D. R. Campelo and F. L. Soares, "An Efficient Intrusion Prevention System for CAN: Hindering Cyber-Attacks With a Low-Cost Platform," in IEEE Access, vol. 9, pp. 166855-166869, 2021, doi: 10.1109/ACCESS.2021.3136147.

[3]. A. Bechini, F. Marcelloni and A. Renda, "TSF-DBSCAN: A Novel Fuzzy Density-Based Approach for Clustering Unbounded Data Streams," in IEEE Transactions on Fuzzy Systems, vol. 30, no. 3, pp. 623-637, March 2022, doi: 10.1109/TFUZZ.2020.3042645.

[4]. F. Nie, C. Liu, R. Wang, Z. Wang and X. Li, "Fast Fuzzy Clustering Based on Anchor Graph," in IEEE Transactions on Fuzzy Systems, vol. 30, no. 7, pp. 2375-2387, July 2022, doi: 10.1109/TFUZZ.2021.3081990.

[5]. X. Hu, Y. Shen, W. Pedrycz, X. Wang, A. Gacek and B. Liu, "Identification of Fuzzy Rule-Based Models With Collaborative Fuzzy Clustering," in IEEE Transactions on Cybernetics, vol. 52, no. 7, pp. 6406-6419, July 2022, doi: 10.1109/TCYB.2021.3069783.

[6]. B. Hu et al., "A Deep One-Class Intrusion Detection Scheme in Software-Defined Industrial Networks," in IEEE Transactions on Industrial Informatics, vol. 18, no. 6, pp. 4286-4296, June 2022, doi: 10.1109/TII.2021.3133300.

[7]. I. Dutt, S. Borah and I. K. Maitra, "Immune System Based Intrusion Detection System (IS-IDS): A Proposed Model," in IEEE Access, vol. 8, pp. 34929-34941, 2020, doi: 10.1109/ACCESS.2020.2973608.

[8]. A. H. Janabi, T. Kanakis and M. Johnson, "Overhead Reduction Technique for Software-Defined Network Based Intrusion Detection Systems," in IEEE Access, vol. 10, pp. 66481-66491, 2022, doi: 10.1109/ACCESS.2022.3184722.

[9]. W. Bul'ajoul, A. James and S. Shaikh, "A New Architecture for Network Intrusion Detection and Prevention," in IEEE Access, vol. 7, pp. 18558-18573, 2019, doi: 10.1109/ACCESS.2019.2895898.

[10]. M. Nadeem, A. Arshad, S. Riaz, S. S. Band and A. Mosavi, "Intercept the Cloud Network From Brute Force and DDoS Attacks via Intrusion Detection and Prevention System," in

IEEE Access, vol. 9, pp. 152300-152309, 2021, doi: 10.1109/ACCESS.2021.3126535.

[11]. T. D. Ramotsoela, G. P. Hancke and A. M. Abu-Mahfouz, "Behavioural Intrusion Detection in Water Distribution Systems Using Neural Networks," in IEEE Access, vol. 8, pp. 190403-190416, 2020, doi: 10.1109/ACCESS.2020.3032251.

[12]. R. Bitton and A. Shabtai, "A Machine Learning-Based Intrusion Detection System for Securing Remote Desktop Connections to Electronic Flight Bag Servers," in IEEE Transactions on Dependable and Secure Computing, vol. 18, no. 3, pp. 1164-1181, 1 May-June 2021, doi: 10.1109/TDSC.2019.2914035.

[13]. V. K. Kukkala, S. V. Thiruloga and S. Pasricha, "INDRA: Intrusion Detection Using Recurrent Autoencoders in Automotive Embedded Systems," in IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 39, no. 11, pp. 3698-3710, Nov. 2020, doi: 10.1109/TCAD.2020.3012749.

[14]. B. Gao, B. Bu, W. Zhang and X. Li, "An Intrusion Detection Method Based on Machine Learning and State Observer for Train-Ground Communication Systems," in IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 7, pp. 6608-6620, July 2022, doi: 10.1109/TITS.2021.3058553.

[15]. H. Hindy et al., "A Taxonomy of Network Threats and the Effect of Current Datasets on Intrusion Detection Systems," in IEEE Access, vol. 8, pp. 104650-104675, 2020, doi: 10.1109/ACCESS.2020.3000179.

[16]. C. F. T. Pontes, M. M. C. de Souza, J. J. C. Gondim, M. Bishop and M. A. Marotta, "A New Method for Flow-Based Network Intrusion Detection Using the Inverse Potts Model," in IEEE Transactions on Network and Service Management, vol. 18, no. 2, pp. 1125-1136, June 2021, doi: 10.1109/TNSM.2021.3075503.

[17]. G. De CarvalhoBertoli et al., "An End-to-End Framework for Machine Learning-Based Network Intrusion Detection System," in IEEE Access, vol. 9, pp. 106790-106805, 2021, doi: 10.1109/ACCESS.2021.3101188.

[18]. A. Kim, M. Park and D. H. Lee, "AI-IDS: Application of Deep Learning to Real-Time Web Intrusion Detection," in IEEE Access, vol. 8, pp. 70245-70261, 2020, doi: 10.1109/ACCESS.2020.2986882.

[19]. T. Wisanwanichthan and M. Thammawichai, "A Double-Layered Hybrid Approach for Network Intrusion Detection System Using Combined Naive Bayes and SVM," in IEEE Access, vol. 9, pp. 138432-138450, 2021, doi: 10.1109/ACCESS.2021.3118573.

[20]. M. H. Haghighat and J. Li, "Intrusion detection system using voting-based neural network," in Tsinghua Science and Technology, vol. 26, no. 4, pp. 484-495, Aug. 2021, doi: 10.26599/TST.2020.9010022.

[21]. H. Yang and F. Wang, "Wireless Network Intrusion Detection Based on Improved Convolutional Neural Network," in IEEE Access, vol. 7, pp. 64366-64374, 2019, doi: 10.1109/ACCESS.2019.2917299.

[22]. X. -X. Zhang, L. -R. Zhao, H. -X. Li and S. -W. Ma, "A Novel Three-Dimensional Fuzzy Modeling Method for Nonlinear Distributed Parameter Systems," in IEEE Transactions on Fuzzy Systems, vol. 27, no. 3, pp. 489-501, March 2019, doi: 10.1109/TFUZZ.2018.2861726.

[23]. M. Peng and J. Fan, "Analysis of Green Economy Development in Pearl River Delta based on Fuzzy Clustering and Entropy Weight Comprehensive Evaluation Model," 2021 IEEE 3rd Eurasia Conference on IOT, Communication and Engineering (ECICE), 2021, pp. 413-416, doi: 10.1109/ECICE52819.2021.9645722.

[24]. J. Hu, M. Wu, L. Chen and W. Pedrycz, "A Novel Modeling Framework Based on Customized Kernel-Based Fuzzy C-Means Clustering in Iron Ore Sintering Process," in IEEE/ASME Transactions on Mechatronics, vol. 27, no. 2, pp. 950-961, April 2022, doi: 10.1109/TMECH.2021.3076208.

[25]. S. Blažič and I. Škrjanc, "Incremental Fuzzy C-Regression Clustering From Streaming Data for Local-Model-Network Identification," in IEEE Transactions on Fuzzy Systems, vol. 28, no. 4, pp. 758-767, April 2020, doi: 10.1109/TFUZZ.2019.2916036.

[26]. P. Xu et al., "Concise Fuzzy System Modeling Integrating Soft Subspace Clustering and Sparse Learning," in IEEE Transactions on Fuzzy Systems, vol. 27, no. 11, pp. 2176-2189, Nov. 2019, doi: 10.1109/TFUZZ.2019.2895572.

[27]. S. Zeng, X. Wang, X. Duan, S. Zeng, Z. Xiao and D. Feng, "Kernelized Mahalanobis Distance for Fuzzy Clustering," in IEEE Transactions on Fuzzy Systems, vol. 29, no. 10, pp. 3103-3117, Oct. 2021, doi: 10.1109/TFUZZ.2020.3012765.

[28]. J. Zuniga-Mejia, R. Villalpando-Hernandez, C. Vargas-Rosales and A. Spanias, "A Linear Systems Perspective on Intrusion Detection for Routing in Reconfigurable Wireless Networks," in IEEE Access, vol. 7, pp. 60486-60500, 2019, doi: 10.1109/ACCESS.2019.2915936.

[29]. Y. Tao, Y. J. Chen, L. Xue, C. Xie, B. Jiang and Y. Zhang, "An Ensemble Model With Clustering Assumption for Warfarin Dose Prediction in Chinese Patients," in IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 6, pp. 2642-2654, Nov. 2019, doi: 10.1109/JBHI.2019.2891164.

[30]. Y. Shen, W. Pedrycz, X. Jing, A. Gacek, X. Wang and B. Liu, "Identification of Fuzzy Rule-Based Models With Output Space Knowledge Guidance," in IEEE Transactions on Fuzzy Systems, vol. 29, no. 11, pp. 3504-3518, Nov. 2021, doi: 10.1109/TFUZZ.2020.3024804.