Journal of Computing, Science &Technology
https://focjournal.unidel.edu.ng/
editorjcst@unidel.edu.ng
ISSN:......Print ISSN:........Online

Bio-Inspired Feature Extraction Techniques for Intrusion Detection Systems

Folasade Adeyemi[1], Samuel Adebayo Oluwadare[2], Boniface Kayode Alese[3]
[1,2,3] Department of Computer Science. Federal University of Technology, Akure, Nigeria.
adeyemif@futa.edu.ng[1], saoluwadare@futa.edu.ng[2], bkalese@futa.edu.ng[3]

Corresponding Author's Email: adeyemif@futa.edu.ng

ABSTRACT

Article Info

Date Received: 24-02-2024
Date Accepted: 03-04-2024

Computer networks and systems are protected against unauthorised access and harmful activity using intrusion detection systems (IDSs). IDS are very important to the security architecture of computer networks. They are at the forefront of detecting harmful activities and unauthorized access to the network. IDS usually manages large amounts of data traffic containing redundant and irrelevant features. Due to its potential to increase detection accuracy, decrease computing burden, and increase system scalability, feature selection has drawn significant attention as a crucial pre-processing stage in IDS design and was shown to significantly affect the classifiers' performance. To create appropriate detection algorithms, the set of features thought to be the most useful properties are retrieved during feature selection. An in-depth analysis of feature selection methods within the framework of IDS is presented in this paper, with a focus on bio-inspired strategies. Bio-inspired strategies offer creative answers to challenging issues by taking their cues from natural systems and processes. This paper explores and evaluates the effectiveness of bio-inspired algorithms for feature selection in intrusion detection systems (IDS) to enhance detection performance. This framework provides a thorough understanding of the current landscape of feature selection for IDS with bio-inspired techniques by critically evaluating the state-of-the-art research, challenges, and opportunities facilitating advancements in this important area of cyber security.

Keywords:
Intrusion Detection System, Pre-Processing, Feature Selection, Bio-inspired techniques, detection performance

1.0 INTRODUCTION

In recent times, the internet has evolved into a fundamental aspect of our daily lives encompassing activities like online bill payments, money transfers, booking tickets, and various other services. Furthermore, government institutions entrust valuable data to the network, with a steadfast commitment to safeguarding its security, integrity, accessibility, and dependability [1]. The expansion and extension of Internet networks have resulted in a growing user base. This surge in demand for online services is accompanied by a heightened need for secure data transmission across the network, especially sensitive information. To ensure secure connections and safeguard shared information over the internet, it is imperative to establish a dependable and adaptable safety system. No security measure can guarantee complete network protection, even with a variety of safeguards including firewalls, data encryption, access controls, and user authentication serving as the main lines of defense in computer security. Consequently, many researchers are actively engaged in creating novel security systems to address this challenge. One such pivotal solution is the Intrusion Detection System (IDS), introduced by Anderson in 1980, which plays a significant role in enhancing security within computer networks by detecting and thwarting attacks [2].

IDSs serve as a critical line of defence in the ever-evolving landscape of cyber security. These systems are built to recognise and take appropriate action against unauthorised access, suspicious activities and potential threats within computer networks and systems [3]. Effectively identifying intrusions is paramount to the security and integrity of digital infrastructures, making IDSs a fundamental component of modern cybersecurity strategies [4]. Statistical and knowledge-based techniques were used in conventional intrusion detection models. These methods had trouble identifying unknown assaults, as well as analyzing vast volumes of data on network traffic [5]. The creation of numerous efficient ways to improve the security of systems, like wireless networks, cloud computing [6], the Internet of Things, and many other network-based system is possible using machine learning, which has considerable potential in this regard. The development of IDS is a significant example of how machine learning is used. This system's function aims to examine network data, discern between normal and abnormal behavior, and appropriately categorize the findings [6]. The choice of features has a big impact on whether intrusion detection works or not [7]. Feature selection, a pre-processing step in IDS design, holds the key to optimizing the performance of these systems. It involves the careful selection of relevant features or attributes from the raw data, thereby reducing dimensionality, enhancing detection accuracy, and improving computational efficiency. As the volume and complexity of network data continue to grow, the significance of feature selection in IDS cannot be overstated.

This paper delves into the realm of feature selection for IDS, with a specific focus on the integration of bio-inspired techniques. The term “bio-inspired", short for “biologically inspired” refers to an area of science and technology that draws inspiration from biological systems and processes found in nature to develop innovative solutions to various complex problems. This approach seeks to replicate or adapt the mechanisms, strategies and structures found in living organisms and ecosystems to solve challenges in engineering, computing, robotics, materials science and other domains [8].

The concept of bio-inspiration stems from the recognition that nature has already evolved highly efficient and effective solutions to a wide range of problems through millions of years of adaptation and evolution. By studying and emulating these natural solutions, researchers and engineers aim to create novel technologies and designs to address human challenges more sustainably, efficiently and adaptively [9].

Many approaches to problem-solving show that bio-inspired optimisation solutions perform better than conventional solutions because of their advantages in terms of fault tolerance, speed, modularity, parallelism, scalability, and adaptability [10]. However, these algorithms possess several drawbacks, including algorithm framework, best performance measurements, the appropriate ratio of exploration to exploitation, the consequences of "free lunches" for choosing the best algorithm or algorithms, and automatic parameter tweaking, among others.

This paper provides a detailed synopsis of the current research in feature selection for IDS with bio-inspired techniques. Through a systematic analysis of existing literature, this research elucidates the strengths and weaknesses of various bio-inspired algorithms when used to pick features for intrusion detection. Additionally, challenges encountered in this domain are explored and Potential avenues are identified for future research and development. By synthesizing and critically evaluating the existing body of knowledge, this paper seeks to contribute to the advancement of feature selection methodologies in IDS, ultimately enhancing the ability of these systems to detect and mitigate cyber security threats effectively. In doing so, it underscores the critical role that bio-inspired techniques can play in shaping the future of intrusion detection and by extension, the broader field of cyber security.

The remainder of this essay is arranged as follows: The intrusion detection system is described in Section 2. Techniques for feature selection are the subject of Section 3. The bio-inspired algorithm for feature selection was covered in Section 4. Future directions for the research are provided in Section 5, while Section 6 discusses the findings.

2.0 INTRUSION DETECTION SYSTEM (IDS)

An IDS or intrusion detection system, is a security tool designed to track and examine system activity and network traffic to identify and respond to unauthorised access, suspicious behaviour, and potential security threats within a computer network or system [11]. IDSs are essential to improving the security of digital environments by delivering alerts in real-time or almost real-time [12] when suspicious or malicious activities are detected. IDS can be classified based on installation site or detection method. Regarding installation site, IDS are either host-based (HIDS) or network-based (NIDS). HIDS focused on monitoring activities on individual hosts within a network, analyzing host-specific logs, processes, and file systems for signs of intrusion or abnormal behavior[13]. On the other hand, NIDS is positioned before the firewall as the first line of defense[14]to monitor and analyze network traffic for indications of suspicious or malicious activities. NIDS is typically deployed at key junctures in the network to capture and inspect traffic flow, making them valuable for detecting attacks that target the network infrastructure[15].

IDS are categorized into two main types according to their detection methods: Signature-Based and Anomaly-Based [16]. These categories represent the fundamental approaches used by IDS systems to detect and respond to security threats. Knowledge-based or misuse detection commonly referred to as signature-based IDS, employs a database that contains established signature or attack patterns. the incoming network traffic or system activity is compared to these signatures to detect known threats; an alert is generated by the IDS upon finding a match. While signature-based IDS are efficient at identifying known attacks, they may struggle with novel or zero-day attacks for which no signature exists. In contrast, anomaly-based IDS establish a baseline profile of normal traffic or system behavior, the incoming packets are then continuously monitored and compared to the established baseline, and any deviation from the established pattern raises suspicions of malice or probable suspicion. Although anomaly-based IDS are good at spotting new attacks and zero-day exploits, but they can produce false positives if the baseline is improperly set up or if there are major changes in normal network behaviour[17]. To increase detection coverage and accuracy, the author in [14] mentioned some hybrid IDS systems that integrate signature-based and anomaly-based approaches. These IDS can effectively identify known assaults and also detect previously unknown threats based on aberrant behaviours.

3.0 FEATURE SELECTION TECHNIQUES

Intrusion Detection Systems (IDS) datasets consist of numerous features [18] that describe the characteristics of network traffic flows. Some of these features, some are redundant or irrelevant, potentially hindering the detection process and consuming excessive system resources [12]. Feature selection is employed to select an important subset of features, enhancing the performance of classifiers and producing a more accurate categorization of data compared to using all features. Understanding the data, reducing the amount of storage space needed, and cutting process costs are all advantages of feature selection [19]. Feature selection, as a preliminary step in designing Intrusion Detection Systems (IDS), aims to decrease complexity, enhance detection accuracy, and improve computational efficiency. It involves selecting the most relevant features or characteristics from the original data.The significance of feature selection in intrusion detection systems (IDS) cannot be overstated, given the increasing quantity and intricacy of network traffic[20]. Three main forms of feature selection techniques for intrusion detection systems (IDSs) can be distinguished based on their methodology and approach: filter, wrapper, and embedded method, which is commonly referred to as a traditional intrusion detection system.

Filter methods assess the importance/relevance of features independently of any machine learning algorithm. They evaluate each feature's characteristics concerning the target variable (intrusion or normal). Among the popular filter techniques are the Chi-Square Test, Information Gain, Correlation-Based Feature Selection (CFS), and Gain Ratio. Wrapper methods evaluate feature subsets by directly employing a machine learning algorithm to determine the effectiveness of a particular feature combination. These methods involves an iterative search process and can be computationally intensive. Forward selection, backward elimination, and recursive feature elimination (RFE) are a few examples. Feature selection is integrated into a machine learning algorithm during the training phase in embedded approaches. They select features as the model is built, making them computationally more efficient than wrapper methods. Common embedded methods are Regularization (Lasso) and Tree-Based Feature Selection (e.g. decision trees, random forests). Apart from the aforementioned conventional categories, bio-inspired optimization methods, such as Genetic Algorithms (GA) [21], Particle Swarm Optimization (PSO), and Ant Colony Optimization (ACO), can also be employed for feature selection in IDS. Bio-inspired meta-heuristic algorithms are frequently employed in the wrapper-based method for the feature selection procedure in intrusion detection Systems.[22]. These algorithms are distinguished from traditional methods by their approach: bio-inspired techniques start with a set of simple rules and simple organisms that follow those rules.

Bio-inspired computing optimization algorithms (BIA) represent a novel approach based on the evolutionary principles found in nature, aiming to develop robust and innovative computational techniques [23]. These algorithms are inspired by the behaviour of natural organisms such as fish schools, ant colonies, bird flocks, and bee swarms. This natural inspiration has attracted the attention of computer science researchers who seek to apply these algorithms to solve a wide range of problems in science and engineering [23]. Among the most prominent classes of bio-inspired algorithms are evolutionary algorithms and swarm intelligence-based algorithms [24]. Evolutionary algorithms, such as genetic algorithms, are rooted in the study of natural evolution. They operate by evolving a population of potential solutions and returning a population of solutions simultaneously. Evolutionary algorithms are particularly well-suited for solving optimization problems that are multimodal, non-differentiable, or discontinuous, which cannot be effectively addressed using traditional methods. Swarm intelligence, on the other hand, is a subset of artificial intelligence focused on the collective behavior of biological swarms. It involves the interaction of individuals within these swarms to solve real-world problems by simulating biological behaviors [23]. Bio-inspired algorithms applied to feature selection (BIA-FS) offer a promising technique for handling non-linear, high-dimensional data.

4.0 BIO-INSPIRED ALGORITHM FOR FEATURE SELECTION

An array of bio-inspired feature selection methods is widely applied in the field of intrusion detection systems [23]. These algorithms provide adaptable and creative answers to the problems involved in locating pertinent features for intrusion detection by taking inspiration from natural processes and phenomena [25]. This study looks at three well-known bio-inspired algorithms and how they can be adapted be modified to meet the particular needs of intrusion detection:

4.1 Particle Swarm Optimization (PSO)

Kennedy and Eberhart introduced Particle Swarm Optimization (PSO) in 1995 [26]. PSO is a population-based stochastic search algorithm that uses a metaheuristic approach to identify the best solution in the search space. It was inspired by the social behavior of fish schools and bird flocking. PSO refers to the population as a swarm and each individual within it as a particle. A population of particles within the search space is first randomly initialized by the PSO algorithm. A 2D vector is used to represent the population, with the number of rows denoting the number of particles and the number of columns denoting the number of features in the dataset [20]. With a corresponding position and velocity vector, every particle in the population traverses the search space iteratively, modifying its location and velocity. With a corresponding position and velocity vector, every particle in the population traverses the search space iteratively, modifying its location and velocity using Eq. (1)

$$V_i^{t+1} = wV_i^t + c_1 r_1 \left(P_i^t - X_i^t\right) + c_2 r_2 \left(G^t - X_i^t\right) \quad (1)$$

the updated velocity of the i^{th} particle at time t+1 is represented by V_i^{t+1} where t denotes the generation counter, V_i^t is the current velocity of the particle at time t. The parameter w signifies inertial of the particle acting as a balancing weight to regulate the influence of forces governing exploration and exploitation within the swarm. Exploration involves searching the uncharted regions of the solution space, whereas exploitation involves exploring the vicinity of a promising region. The coefficients c_1 and c_2 are the acceleration coefficients, representing social and cognitive aspects respectively, which determine the impact of the particle's best position (P_i^t) and the global best position (G^t). The values r1 and r2 are random numbers ranging between 0 and 1, adding stochasticity to the algorithm. Finally, X_i^t denote the existing position of the particle at the time t, and it is updated using Eq. (2) by the particle after updating the velocity.

$$X_i^{t+1} = X_i^t + V_i^{t+1} \qquad (2)$$

PSO's main benefits over other Bio-Inspired Algorithms (BIAs) are its straightforward implementation and reduced number of regulating factors. But as [27] points out, premature convergence poses a serious problem for PSO. Specifically in complicated multimodal tasks, this problem is caused by a lack of population diversity [27]. PSO has just three parameters, however finding the right values for each iteration might be difficult to regulate. To overcome PSO constraints such as aggregation to local minima or suboptimal values and delayed integration, an intrusion detection system (IDS) needs to be carefully designed [24].

4.2 Cuckoo Search Algorithm (CSA)

The obligate brood parasitism behavior exhibted by some cuckoo species laying their eggs in the nests of other host birds or species inspired Yang and Deb [28] to develop the cuckoo search (CS) algorithm [29]. In such an instance, the host birds can discard the foreign eggs or abandon their nest to construct a new one to directly confront the invading cuckoos.

A key characteristic of the CS algorithm is its Levy-flight-style behavior, which mimics the flight patterns of many animals. This behavior involves individuals predominantly engaging in smaller-range activities but occasionally making long-range jumps with a small probability. Similarly, there is also a small probability of significant deviations from the mean value of activities, empowering the CS algorithm to escape local optima [30],

Three rules govern how the cuckoo search algorithm generates solutions.

1. At a time, a single egg is laid in a nest that is selected at random by each Cuckoo.
2. Nests containing high-quality eggs are preserved as elitism nests and are retained for the next generation.
3. The nest contains a fixed number of host nests, and each nest has a probability, Pa, that ranges from 0 to 1, associated with recognizing an alien egg. The host bird may dispose of an odd egg or abandon the nest to build a new one elsewhere if it notices something strange in the egg. Some elements of the existing solution need to be changed to produce a new and improved solution [21].

These three rules are followed while updating the nests iteratively using Eq. (3).

$$X_i^{t+1} = X_i^t + \propto \oplus Levy(\lambda)(\lambda),\ \mathrm{i} = 1, 2, \ldots, \qquad (3)$$

The new solution for Cuckoo i is denoted by X_i^{t+1} while X_i^t represents the current solution and the product $\oplus$ denotes entry-wise multiplication. The levy flight is computed using Eq. (4).

$$Levy \sim \mathrm{u} = \mathrm{t}^{\sim\lambda}\ (1 < \lambda < 3) \qquad (4)$$

Hence, the CS algorithm efficiently explores the solution space by adjusting its step size dynamically. It performs short-distance local searches and occasional long-distance jumps, with the length of the step gradually increasing over time[21].

4.3 Ant Colony Optimization (ACO)

Inspired by ants' foraging behaviour, Marco Dorigo developed Ant Colony Optimisation (ACO), a metaheuristic optimisation algorithm, in the early 1990s [31]. It is particularly effective in tackling combinatorial optimization problems. The algorithm founded on the principle that ants can use pheromone trails to locate the quickest route from their colony to a food source. ACO involves a population of artificial ants [32] that collaboratively searches for optimal solutions in a discrete search space. The algorithm models the behavior of real ants, which lay and follow pheromone trails to find the best path to a food source. ACO often uses a graph to represent the problem space, where each node represents a feature (selected or not selected). The quantity of pheromone along a path connecting two nodes influences the likelihood of an ant travelling through that path and is calculate using Eq. (5)

$$P_{i,j}^k = \begin{cases} \dfrac{(\tau_{i,j})^{\alpha}(\eta_{i,j})^{\beta}}{\sum_{m \in N_i^k}\left((\tau_{i,m})^{\alpha}(\eta_{i,m})^{\beta}\right)}, & j \in N_i^k \\ 0 & ,\ j \notin N_i^k \end{cases} \qquad (5)$$

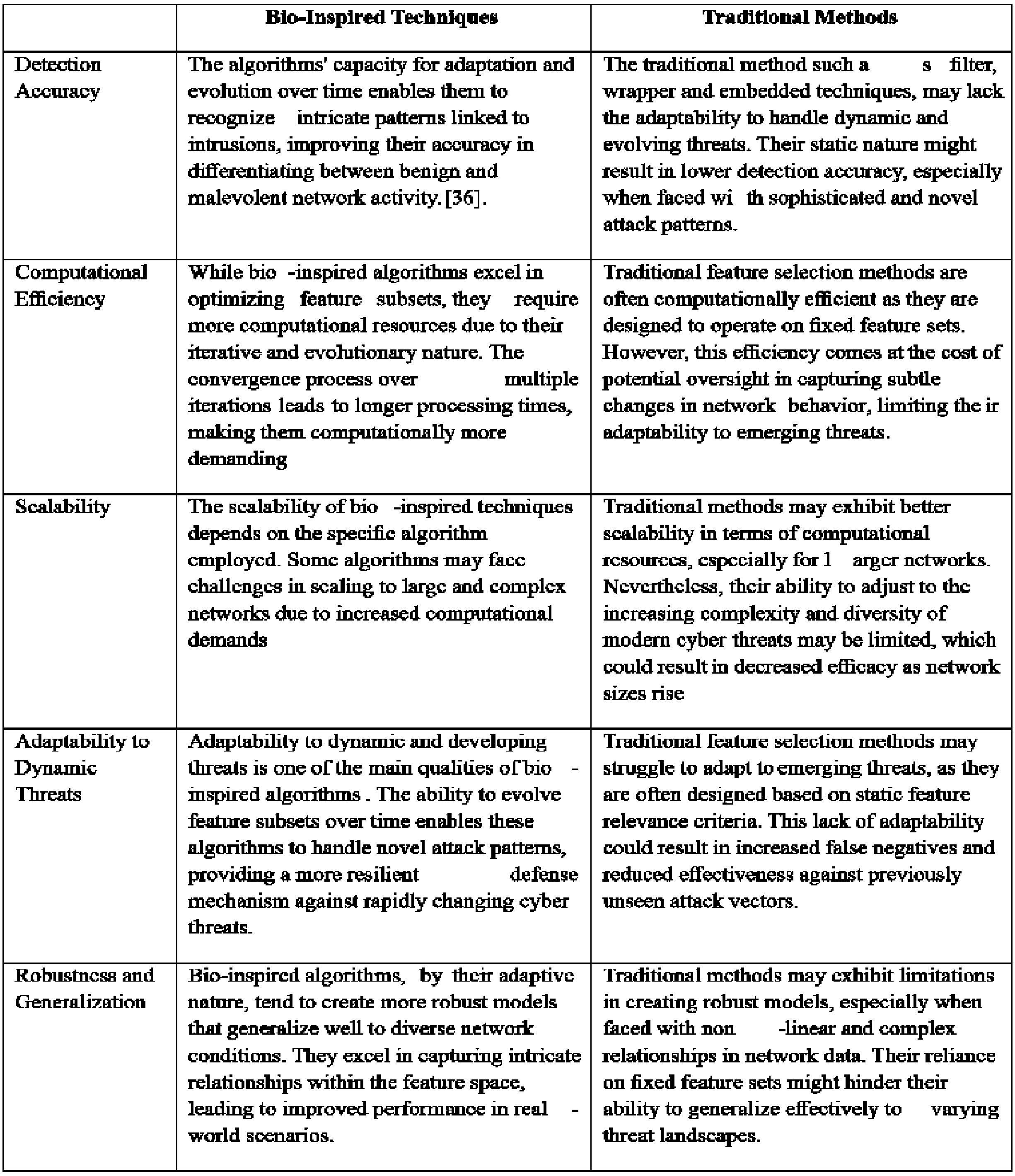
where η represents heuristic information, indicating the number of times a feature has been visited. The variable j represents the set of neighboring nodes that have not yet been visited by ant k. The parameter ∝ and β determine the influence of pheromone versus heuristic information [33]. The pheromone value is updated using the equation 6.

$$\tau_{i,j}^k = (1 - \rho)\tau_{i,j} + \sum_{k=1}^{m} \triangle\, \tau_{i,j}^k \qquad (6)$$

4.4 Applications of Bio-Inspired Techniques in Feature Selection:

Numerous research projects and real-world implementations have demonstrated the usefulness of bio-inspired techniques for feature selection in intrusion detection systems (IDS) to increase overall efficacy, reduce false positives, and improve detection precision. A study in [27], [34] applied a genetic algorithm to feature selection for IDS in a large-scale corporate network. The genetic algorithm efficiently evolved feature subsets that significantly improved the IDS accuracy. By evolving over multiple generations, it adapted to diverse attack scenarios and reduced false positives by prioritizing the most relevant features. Researchers implement PSO for feature selection in an academic network environment. The PSO algorithm facilitated the exploration of the feature space, converging towards a feature subset that enhanced the IDS's accuracy. By considering both global and local information, it demonstrated effectiveness in adapting to different network conditions and reducing false positives. Artificial Immune System (AIS) was employed in the research work of [35] to enhance feature selection for an IDS. The AIS algorithm, inspired by immune system principles, demonstrated adaptability to

Table 1: Explains a Summary of the Comparison Between Bio-Inspired Techniques and Traditional Methods for Intrusion Detection:



	Bio-Inspired Techniques	Traditional Methods
Detection Accuracy	The algorithms' capacity for adaptation and evolution over time enables them to recognize intricate patterns linked to intrusions, improving their accuracy in differentiating between benign and malevolent network activity. [36].	The traditional method such a s filter, wrapper and embedded techniques, may lack the adaptability to handle dynamic and evolving threats. Their static nature might result in lower detection accuracy, especially when faced wi th sophisticated and novel attack patterns.
Computational Efficiency	While bio -inspired algorithms excel in optimizing feature subsets, they require more computational resources due to their iterative and evolutionary nature. The convergence process over multiple iterations leads to longer processing times, making them computationally more demanding	Traditional feature selection methods are often computationally efficient as they are designed to operate on fixed feature sets. However, this efficiency comes at the cost of potential oversight in capturing subtle changes in network behavior, limiting the ir adaptability to emerging threats.
Scalability	The scalability of bio -inspired techniques depends on the specific algorithm employed. Some algorithms may face challenges in scaling to large and complex networks due to increased computational demands	Traditional methods may exhibit better scalability in terms of computational resources, especially for l arger networks. Nevertheless, their ability to adjust to the increasing complexity and diversity of modern cyber threats may be limited, which could result in decreased efficacy as network sizes rise
Adaptability to Dynamic Threats	Adaptability to dynamic and developing threats is one of the main qualities of bio -inspired algorithms . The ability to evolve feature subsets over time enables these algorithms to handle novel attack patterns, providing a more resilient defense mechanism against rapidly changing cyber threats.	Traditional feature selection methods may struggle to adapt to emerging threats, as they are often designed based on static feature relevance criteria. This lack of adaptability could result in increased false negatives and reduced effectiveness against previously unseen attack vectors.
Robustness and Generalization	Bio-inspired algorithms, by their adaptive nature, tend to create more robust models that generalize well to diverse network conditions. They excel in capturing intricate relationships within the feature space, leading to improved performance in real -world scenarios.	Traditional methods may exhibit limitations in creating robust models, especially when faced with non -linear and complex relationships in network data. Their reliance on fixed feature sets might hinder their ability to generalize effectively to varying threat landscapes.

changing network dynamics. The proposed methodology implemented two layers of defense: the innate system and the adaptive system. The innate system emulates the innate immune system found in nature, serving as the initial line of defense. Meanwhile, the adaptive system replicates the functionality of the Adaptive Immune System, integrating defensive mechanisms akin to T-cells and B-cells. The results demonstrated the efficacy of the proposed methodology in efficiently detecting intrusions following the induction of malicious attacks on the network system. Ant Colony Optimization (ACO) was utilized in a practical scenario for feature selection in an industrial control system network. The ACO algorithm effectively navigated the feature space, optimizing the IDS for the unique characteristics of industrial network traffic. This resulted in increased accuracy in detecting anomalous patterns and a notable decrease in false positives. Researchers applied the Bee Algorithm to feature selection in critical infrastructure networks. The Bee Algorithm demonstrated efficiency in foraging the optimal feature subsets in the feature space. This adaptation resulted in improved detection accuracy for critical infrastructure-related anomalies and a reduction in false positives.

These applications demonstrated the bio-inspired approaches' versatility to various network settings and attack situations, and they were essential in improving feature selection for IDS. These algorithms' evolutionary and iterative design enabled them to adapt dynamically to network changes, which increased overall efficiency. Moreover, a reduction in false positives was consistently observed during these trials, indicating that the selected features were more relevant and tailored to malicious activity, leading to a more dependable and precise intrusion detection system

It is critical to modify the parameters, operators, and fitness functions of these bio-inspired intrusion detection algorithms to meet the unique demands of the IDS environment. With network monitoring occurring in real-time, the difficulty is striking a compromise between computational efficiency and the requirement for precise intrusion detection. To make sure that these algorithms successfully capture the features of network traffic suggestive of both harmful and legitimate activity, parameter adjustment and customization are crucial. Because bio-inspired algorithms are adaptive, they are a good fit for dealing with the constantly changing and dynamic nature of cyber threats in intrusion detection systems.

4.5 Bio-Inspired Techniques vs. Traditional Feature Selection Methods in IDS

In summary, a trade-off between detection accuracy, computing efficiency, and flexibility is revealed when bio-inspired strategies for intrusion detection are compared with conventional feature selection methods. Although bio-inspired approaches may require more processing power, they excel at improving detection accuracy and adapting to changing threats. On the other hand, traditional approaches provide computational efficiency but could find it difficult to stay up with the constantly changing landscape of cyber threats. The network environment's unique requirements and the kinds of threats it faces should be taken into account while selecting between these options. Research is also being done on hybrid systems that combine the best features of classical and bio-inspired techniques to create an intrusion detection system that is both balanced and efficient.

4.6 Challenges of Integrating Bio-Inspired Techniques for Feature Selection in IDS

Parameter Tuning: Bio-inspired algorithms frequently have a set of parameters that need to be carefully adjusted to work at their best. It can be difficult to discover the ideal combination for these algorithms because their efficiency can be sensitive to parameter adjustments. Inadequate parameter switching could produce less-than-ideal outcomes, which would prevent the algorithm from converging to useful solutions. It requires expertise and extensive experimentation to identify parameter configurations suitable for specific IDS environments.

Algorithm Complexity: The inherent complexity of some bio-inspired algorithms, such as particle swarm optimization or genetic algorithms, can pose challenges in terms of understanding, implanting, and optimizing these algorithms for feature selection in IDS. Increased algorithmic complexity may lead to higher computational demands, making real-time intrusion detection a challenge. Additionally, complex algorithms may be more susceptible to overfitting reducing their generalization capabilities.

Interpretability and Explainability: Bio-inspired techniques often produce results that are not easily interpretable. Understanding the rationale behind the selected features can be adoption crucial for security analysts in IDS, and the lack of interpretability may hinder the adoption of these techniques in practical settings. Security analysts may face difficulties in trusting and interpreting the decisions made by bio-inspired algorithms. This lack of transparency can be a barrier to widespread adoption in security-critical applications.

Adaptation to Evolving Threats: Although bio-inspired algorithms are recognized for their flexibility, ensuring their continuous effectiveness in the face of evolving and sophisticated threats remains a challenge. Failure to adapt quickly to new attack patterns may result in delayed detection or increased false negatives. Maintaining the algorithm’s currency with the latest threat intelligence is essential for robust intrusion detection.

5.0 DISCUSSION

Combining the strength of bio-inspired techniques with traditional methods, such as statistical or information-theoretic approaches, offers a promising avenue. With hybrid models, feature selection in IDS can be enhanced by combining the effectiveness of conventional techniques with the flexibility of bio-inspired algorithms. Integrating diverse feature selection strategies can address the limitations of individual approaches, providing a more comprehensive and resilient solution for intrusion detection such as exploring the integration of deep learning with bio-inspired algorithms for feature selection in IDS.

Deep learning models have shown success in learning complex patterns from large-scale data. Integrating them with bio-inspired algorithms can leverage the representation learning capabilities of deep models while maintaining the adaptability of bio-inspired approaches Research should focus on adapting and evaluating emerging bio-inspired algorithms for feature selection [23] in IDS. This includes exploring algorithms inspired by swarm intelligence, nature-inspired optimization, or unconventional bio-inspired paradigms. Emerging bio-inspired algorithms may offer novel perspectives and improved performance over traditional approaches. Evaluating their applicability in the context of intrusion detection can support developments in the area of enhancing interpretability and explainability of results generated by bio-inspired algorithm for feture selection. Developing techniques to increase the clarity and transparency of bio-inspired algorithms' decision-making processes guarantees the reliability of the intrusion detection system and makes working with security analysts easier. Investigating real-time optimization techniques to address computational demands of bio-inspired algorithms, ensuing their practical applicability for real-time intrusion detection. Achieving feature selection in real-time is crucial for the success of intrusion detection systems. Research in this direction can lead to more practical and scalable solutions.

As the field of bio-inspired feature selection for intrusion detection systems advances, it will be necessary to address these problems and look into these possible future research directions to create more resilient, effective, and adaptable intrusion detection systems that can successfully combat the constantly shifting landscape of cyber threats.

6.0 CONCLUSION

In conclusion, this paper highlights how feature selection for intrusion detection systems (IDS) has been transformed by bio-inspired methodologies. These algorithms, which draw inspiration from natural processes and offer a dynamic and adaptive feature selection technique, are successful in raising the overall performance of IDS. Bio-inspired techniques such as Genetic Algorithms, Particle Swarm Optimization, Artificial Immune Systems, and Ant Colony Optimization have shown their ability to evolve feature subsets that significantly improve detection accuracy. The adaptive nature of these algorithms allows them to capture complex patterns indicative of both known and novel cyber threats. The integration of bio-inspired algorithms in feature selection has led to a notable reduction in false positives. By dynamically selecting features that are more indicative of malicious activities, these algorithms contribute to the precision and reliability of intrusion detection systems. One of the paramount strengths of bio-inspired techniques lies in their adaptability to evolving cyber threats. As the threat landscape constantly changes, the ability of these algorithms to dynamically adjust and optimize feature subsets ensure that the IDS remains resilient against emerging attack vectors. Bio-inspired techniques hold significant potential for contributing to the development of robust and adaptive IDS. Their ability to adapt to changing network conditions, learn from evolving threats, and optimize feature subsets positions them as key components in the arsenal against cyber threats.

REFERENCES

[1] G. K. Ijemaru, I. Adeyanju, K. Olusuyi, and T. J. Ofusori, “Security Challenges of Wireless Communications Networks : A Survey Security Challenges of Wireless Communications Networks : A Survey,” no. July, 2018.

[2] T. A. Alamiedy, M. Anbar, and A. K. Al-ani, *Review on Feature Selection Algorithms for Anomaly-Based Intrusion Detection System*, vol. 1. Springer International Publishing. doi: 10.1007/978-3-319-99007-1.

[3] R. R. Chaudhari and S. P. Patil, “INTRUSION DETECTION SYSTEM: CLASSIFICATION, TECHNIQUES AND DATASETS TO IMPLEMENT,” *Int. Res. J. Eng. Technol.*, 2017, [Online]. Available: www.irjet.net

[4] B. Selvakumar and K. Muneeswaran, “Firefly algorithm based feature selection for network intrusion detection,” *Comput. Secur.*, vol. 81, pp. 148–155, Mar. 2019, doi: 10.1016/j.cose.2018.11.005.

[5] R. Panigrahi and S. Borah, “A detailed analysis of CICIDS2017 dataset for designing Intrusion Detection Systems,” *Int. J. Eng. Technol.*, vol. 7, no. 3.24 Special Issue 24, pp. 479–482, 2018.

[6] M. Bakro *et al.*, “Efficient Intrusion Detection System in the Cloud Using Fusion Feature Selection Approaches and an Ensemble Classifier,” pp. 1–27, 2023.

[7] A. H. Mohammad, T. Alwada, O. Almomani, S. Smadi, and N. Elomari, “Bio-inspired Hybrid Feature Selection Model for Intrusion Detection,” no. August, 2022, doi: 10.32604/cmc.2022.027475.

[8] L. B. Jivanadham, W. H. Hassan, and O. Zakaria, “Biologically Inspired Intrusion Detection (Biid): a Review,” *ARPN J. Eng. Appl. Sci.*, vol. 10, no. 16, pp. 7142–7152, 2015.

[9] L. Kalabarige and H. B. Maringanti, “A Survey on Identity-Based Security in Wireless Sensor Networks,” *Lect. Notes Networks Syst.*, vol. 134, pp. 487–498, 2021, doi: 10.1007/978-981-15-5397-4_50.

[10] S. U. Otor, B. O. Akinyemi, T. A. Aladesanmi, G. A. Aderounmu, and B. H. Kamagaté, “An adaptive bio-inspired optimisation model based on the foraging behaviour of a social spider,” *Cogent Eng.*, vol. 6, no. 1, Jan. 2019, doi: 10.1080/23311916.2019.1588681.

[11] M. Aljanabi, M. A. Ismail, and A. H. Ali, “Intrusion detection systems, issues, challenges, and needs,” *Int. J. Comput. Intell. Syst.*, vol. 14, no. 1, pp. 560–571,

2021, doi: 10.2991/ijcis.d.210105.001.

[12] "Recent Trends in Data Science and Soft Computing..pdf.crdownload."

[13] R. Ghanbarzadeh, A. Hosseinalipour, and A. Ghaffari, "A novel network intrusion detection method based on metaheuristic optimisation algorithms," *J. Ambient Intell. Humaniz. Comput.*, vol. 14, no. 6, pp. 7575–7592, 2023, doi: 10.1007/s12652-023-04571-3.

[14] A. Alaba, S. Maitanmi, and O. Ajayi, "An Ensemble of classification techniques for Intrusion Detection Systems." [Online]. Available: https://sites.google.com/site/ijcsis/

[15] R. Vijayanand and D. Devaraj, "A Novel Feature Selection Method Using Whale Optimization Algorithm and Genetic Operators for Intrusion Detection System in Wireless Mesh Network," *IEEE Access*, vol. 8, pp. 56847–56854, 2020, doi: 10.1109/ACCESS.2020.2978035.

[16] A. Hashmi, "BIE IDS : Bio-Inspired Ensemble Method for Intrusion Detection System," vol. 15, no. 1, pp. 56–68, 2022.

[17] B. F. Balogun, "Feature Selection based on Bat Algorithm and Residue Number System for Feature Selection based on Bat Algorithm and Residue Number System for Intrusion Detection System," no. June, 2022, doi: 10.5120/ijais2022451929.

[18] S. Sharma, V. Kumar, and K. Dutta, "Multi-objective optimization algorithms for intrusion detection in IoT networks: A systematic review," *Internet Things Cyber-Physical Syst.*, vol. 4, no. February, pp. 258–267, 2024, doi: 10.1016/j.iotcps.2024.01.003.

[19] M. Samadi Bonab, A. Ghaffari, F. Soleimanian Gharehchopogh, and P. Alemi, "A wrapper-based feature selection for improving performance of intrusion detection systems," *Int. J. Commun. Syst.*, vol. 33, no. 12, Aug. 2020, doi: 10.1002/dac.4434.

[20] A. Shanbhag, S. Vincent, B. G. S. B, and O. M. Prakash, "Leveraging Metaheuristics for Feature Selection with Machine Learning Classification for Malicious Packet Detection in Computer Networks," *IEEE Access*, vol. PP, p. 1, 2024, doi: 10.1109/ACCESS.2024.3362246.

[21] W. A. Al Al-Bayt and J. University, "The Cuckoo Feature Filtration Method for Intrusion Detection (Cuckoo-ID)," 2020. [Online]. Available: www.ijacsa.thesai.org

[22] A. Alzaqebah, I. Aljarah, O. Al-Kadi, and R. Damaševičius, "A Modified Grey Wolf Optimization Algorithm for an Intrusion Detection System," *Mathematics*, vol. 10, no. 6, pp. 1–16, 2022, doi: 10.3390/math10060999.

[23] T. H. Pham and B. Raahemi, "Bio-Inspired Feature Selection Algorithms With Their Applications : A Systematic Literature Review," *IEEE Access*, vol. 11, no. April, pp. 43733–43758, 2023, doi: 10.1109/ACCESS.2023.3272556.

[24] R. C. Jeyavim Sherin and K. Parkavi, "Investigations on Bio-Inspired Algorithm for Network Intrusion Detection – A Review," *Int. J. Comput. Networks Appl.*, vol. 9, no. 4, pp. 399–423, 2022, doi: 10.22247/ijcna/2022/214503.

[25] N. F. Johari, A. M. Zain, N. H. Mustaffa, and A. Udin, "Firefly algorithm for optimization problem," in *Applied Mechanics and Materials*, 2013, vol. 421, pp. 512–517. doi: 10.4028/www.scientific.net/AMM.421.512.

[26] O. Almomani, "A Hybrid Model Using Bio-Inspired Metaheuristic Algorithms for Network Intrusion Detection System," *Comput. Mater. Contin.*, vol. 68, no. 1, pp. 409–429, Mar. 2021, doi: 10.32604/cmc.2021.016113.

[27] M. Sazzadul Hoque, "An Implementation of Intrusion Detection System Using Genetic Algorithm," *Int. J. Netw. Secur. Its Appl.*, vol. 4, no. 2, pp. 109–120, Mar. 2012, doi: 10.5121/ijnsa.2012.4208.

[28] A. S. Joshi, O. Kulkarni, G. M. Kakandikar, and V. M. Nandedkar, "Cuckoo Search Optimization- A Review," in *Materials Today: Proceedings*, 2017, vol. 4, no. 8, pp. 7262–7269. doi: 10.1016/j.matpr.2017.07.055.

[29] P. Dash, L. C. Saikia, and N. Sinha, "Comparison of performances of several Cuckoo search algorithm based 2DOF controllers in AGC of multi-area thermal system," *Int. J. Electr. Power Energy Syst.*, vol. 55, pp. 429–436, 2014, doi: 10.1016/j.ijepes.2013.09.034.

[30] Institute of Electrical and Electronics Engineers. Ukraine Section. I & M/CI Joint Societies Chapter and Institute of Electrical and Electronics Engineers, *IDAACS'2017 : proceedings of the 2017 IEEE 9th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS) : September 21-23, 2017, Bucharest, Romania*.

[31] J. Matos, C. M. Rebello, E. A. Costa, L. P. Queiroz, M. J. B. Regufe, and B. R. Idelfonso, "Bio-inspired Algorithms in the Optimisation of Wireless Sen- sor Networks : State of the Art Review," *Futur. Internet*, pp. 1–27, 2022.

[32] P. Melin, *Studies in Computational Intelligence 915 Recent Advances of Hybrid Intelligent Systems Based on Soft Computing*. 2020.

[33] G. Li, Y. Jin, M. W. Akram, X. Chen, and J. Ji, "Application of bio-inspired algorithms in maximum power point tracking for PV systems under partial shading conditions – A review," *Renewable and Sustainable Energy Reviews*, vol. 81. Elsevier Ltd, pp. 840–873, Jan. 01, 2018. doi: 10.1016/j.rser.2017.08.034.

[34] H. Suhaimi, S. I. Suliman, I. Musirin, A. F. Harun, and R. Mohamad, "Network intrusion detection system by using genetic algorithm," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 16, no. 3, pp. 1593–1599, 2019, doi: 10.11591/ijeecs.v16.i3.pp1593-1599.

[35] I. Dutt, S. Borah, and I. Maitra, "Intrusion Detection System using Artificial Immune System," *Int. J. Comput. Appl.*, vol. 144, no. 12, pp. 19–22, Jun. 2016, doi: 10.5120/ijca2016910496.

[36] H. Lin, "DeepShield: A Hybrid Deep Learning Approach for Effective Network Intrusion Detection," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 7, pp. 1094–1104, 2023, doi: 10.14569/IJACSA.2023.01407117.