Incorporating Evolutionary Computation for Securing Wireless Network against Cyber-threats

Re-Mid Semester Assignment

IQRA NATIONAL UNIVERSITY

INU

DEFINING POSSIBILITIES, CREATING OPPORTUNITIES

Student Name	**Iman**
Reg ID#	**13523**
Department	**MS (CS)**

Incorporating Evolutionary Computation for Securing Wireless Network against Cyber-threats

Abstract: Due to the rapid growth of Internet services, the demand for network protection and security against complex attacks is increasing. Today, in network security, intrusion detection systems (IDS) play an important role in detecting intrusion activities. In order to reduce the number of search dimensions and improve the classification performance of the IDS model, several hybrid evolution algorithms have been studied in the literature to solve the anomaly detection problem, but there are few disadvantages. For example, low diversity, high false negative rate and stagnation. In order to solve these limitations, in this study, we introduced a new evolutionary hybrid algorithm that combines the grasshopper hopper optimization algorithm (GOA) and simulated annealing (SA) technology, called GOSA of IDS, which extracts the most significant feature. And delete irrelevant data from the original IDS data set. In the proposed method, SA is integrated into GOA, and it is used to improve the quality of the GOA solution after each iteration. The performance of this method was evaluated on two IDS data sets (such as NSL-KDD and UNSW-NB15). It can be seen from the experimental results that the method proposed in this paper surpasses the existing advanced methods, and achieves a high detection rate of 99.86%, an accuracy of 99.89%, and a low false alarm rate of 0.009-KDD in NSL, in UNSW-NB15 Among them, the detection rate was 98.85%, the accuracy rate was 98.96%, and the false alarm rate was low, 0.084.

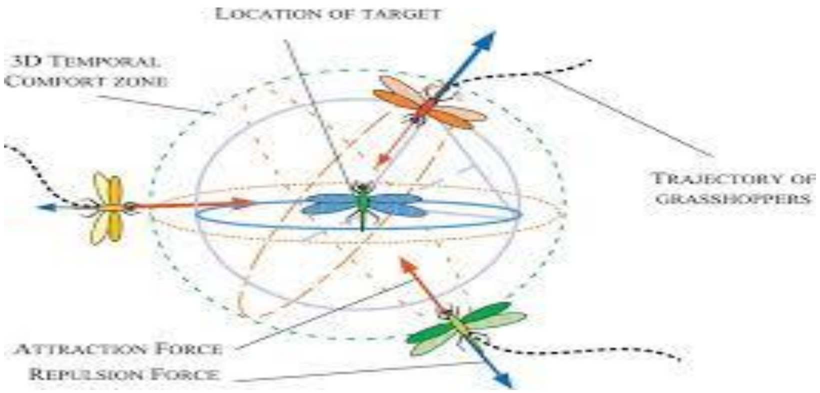
1 Introduction: In recent years, cyber security has become the subject of many reflections and has opened up a field of research. Network security contributes to the security of information systems (such as hardware, software, and related organizations), the data stored on the systems, and the services provided by these systems. Intruders can illegally access this information, and this information can also be used and abused. Sometimes, system operators intentionally cause misuse or damage. Therefore, man-made or accidental damage may be the reason for not following the safety measures. Despite the contemporary development of computer networks, security applications, and people's understanding of modern defense technologies today, modern devices against the latest cyber-attacks still cannot provide comprehensive protection. In order to solve this problem, intrusion detection systems (IDS) are now considered to be effective methods to improve detection capabilities. The intelligent IDS system can play an important role in protecting the network from cyber-attacks. Intrusion detection systems are such devices, whether they are hardware or software, they inspect network traffic, scan it for signature/heuristic scanning to identify any malicious activity (such as malware or vulnerability attacks) and generate alerts/alerts to The security team can analyze and take appropriate measures. In general, intrusion and malicious attacks on computers and their databases undermine IT security policies, namely availability, confidentiality, and integrity. By convention, the most commonly used protections for computer networks are: user authentication, firewalls, data encoding, and various old technologies. The main problem is how to increase the capacity of the intelligent intrusion detection system, which has become the central point of network security. In contrast, current IDS has detected many problems. Unlike new attacks, their detection rate is usually very low, and IDS is heavily overloaded when using audit data. The increasing complexity and number of new attacks require a new variety of solutions. Therefore, methods based on computational intelligence, such as natural-based algorithms, meta-heuristic techniques, and data mining; need to be applied to increase the productivity of IDS. In recent decades, IDS has used several machine learning algorithms, namely Naive Bayes, multi-layer perceptron, support vector machine and artificial neural network to find the type of attack. SF technology is mainly divided into filtering and packaging methods. In today's decades, meta-heuristic techniques such as teaching optimization (TLBO), colony optimization (ACO), artificial bee colony (ABC), gravity study (GSA) and tabu study (TS), differential evolution (DE) , Genetic Algorithm (GA), Particle Swarm Optimization (PSO) and Grasshopper Optimization Algorithm (GOA) have been successfully used to select the optimized subset of functions that can improve the performance of IDS. In order to balance the trade-off between these two goals, which can enrich the performance of the meta-heuristic method, several hybrid algorithms are also proposed to overcome the shortcomings of the basic method and improve the optimization ability. Therefore, several hybrid technologies called GAPSO, GA-SVM, and SVM-PSO have been introduced to detect attack types. The rest of this article is organized as follows: Section 2 introduces related work on attribute selection techniques. Section 3 introduces the attribute selection method based on the concepts of existing wrappers and SVM classification technology. Section 4 presents the proposed method plan for selecting the obvious attributes of the data set and distinguishing various attacks. Section 5 illustrates the experimental results on the data set. Section 6 discusses the conclusion.

2 Related Work: In order to explore the problem of network intrusion detection, in recent years, researchers have used different types of methods, such as evolution and filtering methods. Function selection algorithms with learning algorithms cannot process or expand large amounts of data. In this study, we focus on the mixed packaging methods of SAT and GPA to discover attacks. In order to deal with linear and nonlinear related data attributes, the author proposes a method based on mutual information to systematically select the best classification attribute. This method uses A as a local search technique to improve the convergence speed of PSO. GOA's ease of use and capabilities have been applied in many areas, such as predicting financial constraints, partial occlusion of photovoltaic panels, wireless node location, and vibration signal analysis. Machine learning and computer networks. GOA has effectively solved many discrete, continuous, multi-objective and single-objective optimization problems for various classical meta-heuristic algorithms (such as DE and PSO) to solve global optimization problems and achieve success. This is the first time that we use the GOA algorithm for intrusion detection in a hybrid model as a search technique to select the best parameters and reduce the size of the data set. Chaos diagrams are used to effectively balance the trade between exploration and development, and also reduce the repulsion/attraction between grasshoppers. Similarly, the binary change of the optimization algorithm is expected to appear grasshopper hopper (GOA), and use them to select the most suitable. The subset of attributes related to classification is a packaging-based approach. It plays an important role in the field of machine learning and provides important attributes for better classification in large data sets. Previous studies have shown that individual search algorithms result in finding the best subset of attributes, which exacerbates over fitting of data, while heuristic search is less affected by over fitting of data. Choose attributes and face a small number of samples. To overcome this limitation, a hybrid FS method is proposed, which reduces irrelevant attributes and selects the best subset of attributes. A hybrid algorithm combining random particle swarm optimization and simulated annealing is proposed. This hybrid model combines the advantages of PSO's exploration capabilities and

SA's local research capabilities. In this method, GA is used to generate digital signatures for network parts by removing evidence from network traffic data through traffic analysis. The experimental results estimate that the method proposed in the network flow achieves 96.53% accuracy and 0.56% false alarm rate. Unlike the filtering method, the scalable method is more accurate as a method for selecting attributes using a basic radial function. In order to reduce the noise generated by the attribute difference and improve the performance of SVM, a rich kernel function is proposed by embedding the mean square error and average value of attributes in the attribute function. The core of Radial Basic Function (RBF). In order to solve the above problems, this paper integrates two evolutionary algorithms called GOSA. Consequently, the results of this prominent work are obtained as high accuracy and detection rates and low false alarm rate and processing time in comparison to related literature. In addition, to verify the performance of the proposed method, the experiment was also accompanied by the most important form of intrusion attacks found in NSL-KDD, UNSW-NB15.

3 Background details of methods: In the field of data network security, function selection is one of the important preprocessing tools to improve the performance of the classifier. In the field of data network security, function selection is one of the important preprocessing tools to improve the performance of classifiers.

3.1. Grasshopper Optimization Algorithm: Grasshopper optimization algorithm is an innovative and efficient heuristic algorithm, which is driven by the life of grasshopper. In their lifespan, there are two parts, namely nymph and adult. It can move in groups and is one of the largest groups of all living things, and grasshoppers are usually found only in the wild. Swarm behavior appears in both nymphs and adulthood, which is an inevitable aspect of locust swarms. The grasshopper nymph jumps and moves, spinning millions like a cylinder.



Generally, natural-based algorithms reasonably divide the research process into two stages: exploration and development.

4. Proposed method

4.1. Binary Grasshopper Optimization Algorithm

In binary FS problem, the agents of binary optimization algorithm can only move to closer and distant corners of this hypercube by flipping one or more bits of position vector y = {y1,y2,...,yd}, is designated as the search space of a hypercube. At the same time, the basic GOA should deal with the continuous FS problem. However, to solve the binary FS problem, this method cannot be used. Generally, the transfer function is used to generate the probability of modifying the position 0 of the i-th agent of the j-th dimension in the current iteration (t) to 1 or vice versa, as the input parameter described in Algorithm 3.

4.2. Optimizing SVM parameters with GOSA

GOA is a latest nature-inspired research technique used to find approximate solutions to various real-world engineering problems. Similar to another method based on GOA population, a set of candidate solutions can be arbitrarily generated to obtain an initial group. Then, evaluate all aspiring agents by calculating the appropriateness, and regard the best research agent among the current agents as the target. Similar to other EAs, GOA stays in local optimization when applied to various optimization problems. Unlike GOA, SA is based on the Metropolis algorithm and can get rid of local optimum. Therefore, in order to combine the detection ability of GOA and the local search ability of SA, this paper uses the fusion of GOA and SA (called GOSA) to optimize the parameters σ and C of SVM.

4.3. Fitness Function

In order to maximize the classification performance of detection rate and accuracy, this paper uses the GOSA packaging method, which plays an important role in feature selection. The ultimate goal of GOSA is to select a subset of functions to obtain the best classification accuracy compared to using all available functions. The use of the adjustment function is to use the accuracy of the algorithm classification obtained due to the features selected during the evolution process.

4.4. Overall Hybrid Approach

In recent years, various evolution methods have been useful in improving the efficiency of IDS systems. In addition, many hybrid technologies should also clarify the shortcomings of each technology. In this segment, we intend to a new hybridized method utilizing GOA and SA to enrich the detection accuracy of IDS. The proposed GOSA method in this work is designed as below:

4.4.1. Data Collection and Pre-processing: Data collection is the first step of intrusion detection. This is the systematic information collection method. For the success and design of IDS, two decisive factors play an important role, namely the data source and the location of the collected data. The proposed IDS runs on the victim's adjacent router and observes incoming network traffic. Since many classifiers only use numeric values, the data conversion process is considered critical and has a significant impact on the accuracy of IDS. This article uses the nominal binary method to convert all nominal attributes to binary numeric attributes.

4.4.2. Data Normalization: In this study, we used the NSL-KDD and UNSW-NB15 datasets for experimental research. Each network record in NSL-KDD contains 41 attributes, including 6 binary, 3 nominal, and 32 digital records, while UNSW-NB15 contains 49 attributes, including floating point, integer, nominal, binary, and record. Before conducting the experiment, it is necessary to manage the discrete attributes by calculating the frequency of the discrete values and converting them into numerical attributes, and convert all attributes in a normalized form so that attributes with higher values are eliminated from the data together. Each feature of each sample is normalized by a huge value and falls within the same range as [-1, +1]. The test data also uses the same process.

4.4.3. Attack Recognition: In general, it is necessary to create a classifier to distinguish attacks to consider multiple types of problems. As for the decision limitation in this case, we can do it more directly. In the first part of the experiment in the field used in this article, if the information provided by the record category is used as regular data, the record category is consistent with the regular category, on the contrary, if an attack is envisaged, the opposite is true.

5. Experimental Results and discussion

5.1.1. NSL-KDD Dataset: This data set is an improved version of KDD Cup 99, which is a standard data set widely used to evaluate intrusion detection systems. It contains 41 attributes (six binaries, three nominal values, and 32 numbers) in each record, and includes normal activity and twenty-four types of attacks. Each connection example corresponds to one of the five labeled classes (DOS, Normal, R2L, Probe, and U2R).

DoS: An attack is considered as denial of service, such as "Smurf", "Tear", "Neptune" and "Ping of death (pod)", "Return", "Email bomb" and "Earth".

R2L: This type of attack is considered a remote attack against local attacks, such as "Httptunnel", "Xsnoop", "Phf", "Spy", "Write Wtp", "Imap", and "Multi-hop".

U2R: This type of attack is considered a user attack on the root cause, for example, "Buer over over ow", "Rootkit", "Perl" and "Load"-module".

5. Conclusion

In order to maximize the classification performance and minimize the calculation time of the intrusion detection system, in the existing literature, several evolutionary hybrid algorithms have been introduced. Different technologies have a unique understanding of solving network security problems. Therefore, integrating more than one algorithm that can reduce the gap between each other is the best way to improve the performance of IDS systems.

The proposed method not only improves the classification performance, but also reduces the calculation time. In order to improve the research ability and robustness during the evolution process, we applied the SA mechanism after the GOA stage was completed to obtain the best solution.

This situation can be identified in two cases: first, the classifier has all attributes, and then the classifier has a subset of features obtained from the proposed technique (GOSA). The proposed technique provides a high detection rate of 99.86% and an accuracy of 99.89% in NSL-KDD, as well as a low false alarm rate of 0.009 and a high detection rate of 98.85%. In UNSW-NB15, it was 98.96%, and the false alarm rate was relatively low at 0.084. This quality establishes more reclassifications through computational models.