INTRUSION DETECTING USING COMPLEX NETWORK IN WSN

Mehreen mirza,Neha Bathla
Student,Assistant Professor
Computer Science Department
yiet

Abstract

In today's world the wireless sensor network has great significant in application like defense surveillance, patient health monitoring, traffic control etc. As WSN utilize radio frequencies so there is threat of interference in network. These threats also include distributed denial of service in which the messages that are sent over the network may be attacked by unauthorized user. It would harm the confidentiality of the network user and the services of network. There are various algorithm that are utilized to detect clone attack in WSN but these schemes only stress on prevention of attack after it is occurred. This would leads to the loss of data and more consumption of limited network resources. So in this work we introduce a new algorithm that is based on DCA along with random walk detection. It would detect earlier the clone attack in WSN and prevent the data loss. The parameters like throughput, energy consumption etc are utilized to analyze the performance of this technique.

Keywords: *DCA, WSN, clone detection*

1. Introduction

Wireless sensor networking stays a standout amongst the most requesting and rising exploration territories of our chance. A Wireless Sensor Network (WSN) is a gathering of self-ruling nodes, which transmits information in wireless channel with little transmission capacity utilization and recurrence. [1]

The various applications such as military application, data collection and monitoring utilize the sensor network because it gives minimal effort solution. Every hub can discover their neighbor nodes in network and this give assistance in courses arrangement in the gathering.[2] These kinds of assaults lessen the ability of WSN, with the goal that they can't work for a drawn out stretch of time. It has often consequences for utilization assets in the network and expands the energy utilization, delay, and decreases the throughput. [3]

The un-ability of authorized user to access network resources that may be website or whole system is known as clone attack. A Distributed clone attack is a synchronized assault which is done on the accessibility of services of some specific network with the assistance of traded off processing frameworks in a roundabout way, so tracking the cloned packets turns out to be more troublesome [3] The principle point of this paper is to shield the Wireless Sensor Network from flooding, a kind of clone assault. Flooding can deplete all network assets, for example, data transfer capacity, energy and processing power and so on and plan another location plot named early identification of clone assault utilizing distributed method. This plan recognizes the attacker based on the quantity of transmissions relating to the quantity of neighbors of a hub and

these transmissions are contrasted and the limit esteem registered and PDR of different nodes in the network. [4]

Clone attack (additionally called hub replication attack) is a serious attack in WSNs. In this attack, a foe catches just a couple of hubs, duplicates them and afterward conveys subjective number of imitations all through the system. The catch of hubs is conceivable in light of the fact that sensor hubs are typically unprotected by physical protecting because of cost contemplations, and are frequently left unattended after deployment. On the off chance that we don't distinguish these reproductions, the system will be helpless against a vast class of inner attacks.[5] For instance, the foe presently can catch the movement passing the reproductions (which may contain the previously mentioned areas of troopers), infuse false information into the system (which might be false summons), slander different hubs and even disavow true blue hubs. Hitherto, most conventions for identifying hub replication have depended on a put stock in base station to give worldwide location. Additionally a portion of the current verification strategies [4, 5] can't identify such attacks, since every one of the reproductions hold real keys. The current methodologies fall into following two classes:

A. Brought together Detection The clearest recognition conspires requires every hub to send a rundown of its neighbors and their guaranteed areas to the base station. The base station would then be able to analyze each neighbor rundown to search for imitated hubs. On the off chance that it finds at least one copy, it can repudiate the imitated hubs by flooding the system with a confirmed renouncement message. [6]

B. Nearby Detection: To abstain from depending on a focal base station, we could rather depend on a hub's neighbors to perform replication identification. Utilizing a voting system, the neighbors can achieve an agreement on the authenticity of a given hub. Sadly, while accomplishing recognition in a disseminated design, this technique neglects to distinguish circulated hub replication in disjoint neighborhoods inside the system. For whatever length of time that the duplicated hubs are no less than two bounces from each other, a simply neighborhood approach will fail. [7]

A clear answer for protect against clone attacks is to give the base station a chance to gather the area data (e.g. area, neighbor list, and so forth.) from every sensor and screen the system centralized. This approach experiences high correspondence overhead by asking for excess data from the system. Further, a "shrewd" clone may report the area of the first hub, influencing the base station to flop in distinguishing the imitation. In [8], propose for one-jump networks that the base station (BS) can store the one of a kind flag trademark for every gadget, and in this way gadget cloning can be distinguished as needs be. Nonetheless, in a multi-bounce sensor organize; it is unreasonable for BS to track the flag attributes of sensors multi-jumps away. In restricted voting/trouble making identification plans [8], hubs inside an area concur/vote on the authenticity of a given hub in view of their nearby perceptions. By the by, these plans are not fit for identifying clones with typical conduct, and may fizzle when various clones in closeness intrigue. Moreover, limited voting/trouble making identification plots intrinsically do not have the capacity to identify dispersed clones that may show up at wherever in the system.

2.Literature Review

Clone attack detection methodology is proposed by [9]. The framework employed by Kontaxis et. Al can be used by the users to determine whether they are under clone attack or not. The components employed in this framework involves

a. Information distiller

This component is used in order to extract the information from legitimate social networking site. Information that could be used to identify the user is extracted by this component and maintained within the buffer.

b. Profile Hunter

Profile hunter used to locate the profile of the users. In case multiple records corresponding to single user is fetched then clone attack is detected.

c. Profile verifier

This component verifies the records filtered by profile hunter. The filtered information is compared against the profile of the user to find the nearest matches. In case matches do occur, profile clone attack is detected.User footprint analysis is proposed by [10]. User may have multiple accounts over the various services over the internet. All the services over the internet uses digital mechanisms.

Topological feature extraction mechanism is proposed by[11] for clone attack detection. In clone attack detection, earliest techniques assume that distinguished keywords are used by malicious users. But this may not be the case all the time. in order to tackle the situations, features like images, topological features etc. must be analysed. Topological analysis allow the user to construct the profile on the basis of heterogeneous features hence producing accurate result associated with the clone attack.

The clone attack detection techniques as proposed by [12] can be considered for such attack resolution. According to Dave et. Al., attack can either be on the access restricted information and anonymous data attacks. To tackle the situations attributes similarity based privacy preservation solutions are proposed. Several techniques corresponding to attribute similarity are used in order to determine the clone attacks.

Social networking is one of the most widely used internet activity as proposed by [9]. it is prone to profile clone attacks and its preservation is compulsory. Kontaxis et al proposed mechasnism for detection of profile clone attacks by the use of archituctural design and prototype system for detecting similarity of attributes in case profile of the user is copied. Experiment result shows better result of clone attack detection hence proving worth of the study.

Clone attack is a problem over the online social media. Detecting and preserving the state of the online social media is a need of the hour. Online social media plays a role of complex network. To detect the profile cloning attacks from such a network technique has been proposed by [13]**.** Entire social media is diided into tow parts. First part considered and dra the social network as a graph. In the second part, graph is divided into subparts based on the similarity of profile. The modular approach considered ultimatley led to the formation of smaller networks consisting of only those nodes having similar characterstics or properties thus facilitate detection of clone attacks. Online social media is a huge network of users. As the uers of the online social media grows, so does the chances of clone attack. To detect the clone attack a new approach for clone attack detection is proposed by [14]. Clone attacks causes the similar profiles from one or more users. In order to determine the similarity, strength of suers profiles matching is determined. The strength determines profile clone attack by the said mechanism. degree of modularity achieved through this technqiue is not perfect and required certain degree of modifications.

2. Dendritic Cell Algorithm

The Dendritic Cell Algorithm is roused by the Danger Theory of the mammalian invulnerable framework, and particularly the part and capacity of dendritic cells. The Danger Theory was proposed by Matzinger and recommends that the parts of the gained resistant framework are to react to signs of threat, as opposed to segregating self from non-self. The hypothesis proposes

that antigen showing cells, (for example, partner T-cells) initiate an alert flag giving the essentially co-incitement of antigen-particular cells to react. Dendritic cells are a kind of cell from the natural invulnerable framework that reacts to some particular types of threat signals.
There are three principle types of dendritic cells: 'immature' that collect parts of the antigen and the signs, 'semi-mature' that are immature cells that internally choose that the neighborhood signals represent sheltered and present the antigen to T-cells resulting in tolerance, and 'mature' cells that internally choose that the nearby flags represent peril and present the antigen to T-cells resulting in a reactive reaction.

Let us assume that estimation of training is 0.5. this determines learning rate. after characterizing learning rate activation function is required to be computed. Normalization is connected to introduce uniformity in calculations.

$$Nodes_i = Dataset_{Traffic_{tupples_i}}$$

$$whereidefinenumberoftupleswithinthetrafficdataset$$

- Hidden layer definition

Processing layer contains the neurons which are defined by normalization function

$$Normalization_i = \frac{Total_{Nodes}}{K}$$

Where, K is the parameter while value vary within the range of 2 to 4.

- Activation function definition

Activation function indicates the activation of weight function which is gradient descent in this case.

$$update_i = lerning_{rate} * gradient\_of_{parametes}$$

$$learningrateisconstantforeachiteration$$

- Backpropagation for weight adjustment

This is used to check prescribed tolerance. If it is not achieved then weight is adjusted by the random factor between [1-5].
If (Perscribed_tolerance)
Then list error rate
Else
Adjust weights $w_{i_j} = W_{i_j} + rand(1,5)$

4.Proposed Methodology

In numerous social networking destinations, network topological structure and properties esteems are the entire data. Hubs represent to clients and edges represent to the relationship among them. In every hub, there are a few characteristics, for example, name, sexual orientation, training, interests, area and social exercises. Clearly network topological structure and trait data can be utilized to recognize some shrouded designs in groups. In this examination, DENDRITIC CELL ALGORITHM(DCA) clustering calculation is connected to distinguish groups in social network diagrams.
This approach is used to obtain optimized number of nodes and weights present within MLP. It takes number of tuples within the dataset as input and calls the MLP described in the second section.

4.1 Proposed Model

The iterations are repeated until optimal weights along with the number of nodes in the input layer are achieved. The pseudo code for the same is described in

- Initialization

Clusters data from the KNN is served as initial population.

- Fitness function

Fitness function involves minimization of errors obtained after MLP
If (Minimum_Error)
Convergence with prediction and error rate
Else
Move to the next phase
End of if

- Based on fitness function, selection of population
- Performing crossover

2 or 4 point crossover is performed and nomination is obtained

$$Normalization_i = \frac{Normalization_i}{K}$$

Normalization selects the parents for next phase

- Performing mutation

Mutation mutates the parents and new chromosomes are generated.
Model for the same is given as under

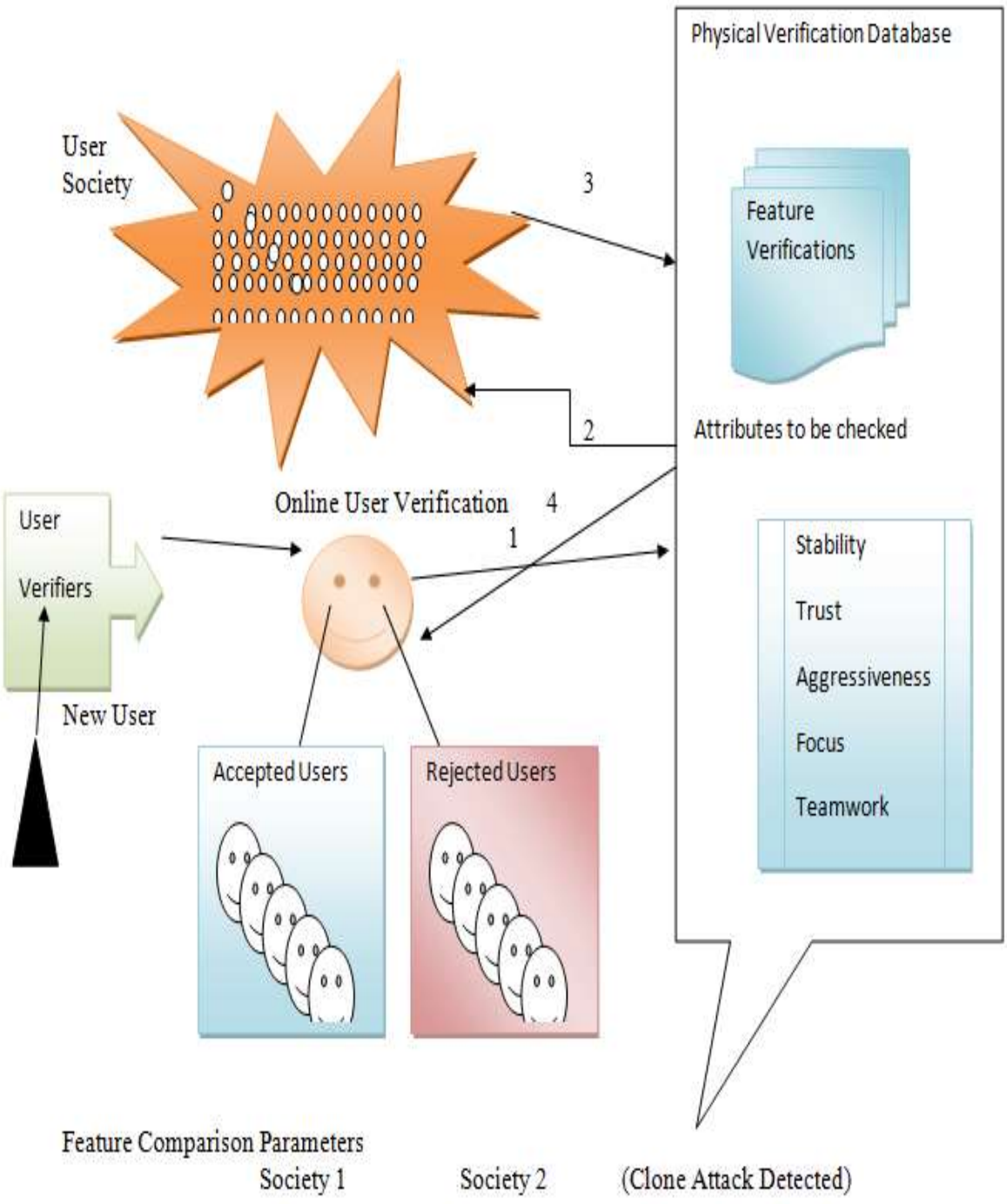

Figure 1. Proposed Model of Proposed Dendritic Cell Algorithm(DCA)

The result of this methodology is optimal predictions although time consumption could be high. This algorithm primarily used to obtain the least possible error in the prediction process.

Pseudo Code

The accompanying demonstrates a pseudo code of the calculation where it acknowledges a property expanded diagram and restore a clustered chart as yield.

Algorithm DCA
Input: Attack Dataset
Output: Classification accuracy

---

1. Input number of nodes G, a
2. $A < -adj(G)$
3. K= a X E[G]
4. Compute the attribute matrix, C
5. S_{ij}= 1 if (I,j) € TopKpair(C), 0 otherwise
6. W= A+S
7. Cluster ←Apply Random walk for clustering
8. Return clusters and classification accuracy

---

3. Metrics for Clone Attack Detection Method

Metric defines the mechanism to estimates the classification of attacks performed on network.**Modularity** defines the network nodes divisibility. More the modularity, more complex network there will be and more chances of attacks.
In order to calculate the modularity we have used the following formula:

$$\boldsymbol{Q = \frac{1}{2m}\sum ij[A_{ij} - \frac{k_i k_i}{2m}]\delta(C_i, C_j)}$$

Where A is adjacency matrix, $k_i$is the level of vertex of i , m is the all number of boundary in the Structure. The component of $A_{ij}$of adjacency matrix is 1 if vertices i and j are associated.

Classification accuracy is used to evaluate the accuracy of result. Classification accuracy of proposed system with DCA gives better performance as compared to naïve bayes and support vector machine. Classification accuracy is obtained as under

$$Accuracy = 1 - MSE$$

MSE is mean square error which must be low to prove worth of the study.

4. **Result and Discussion**The proposed algorithm will reduce the deception over the social media. The physical check mechanism is used in this case. Dataset derived from UCI website is used to detect the problems within the network. DCA algorithm produce low modularity thereby reducing the error rate within the network.

Simulation Parameters

Simulation parameters includes the network size defined in the form of datset. Dataset is derived from the UCI website. Dataset is represented in the form of a graph. Simulation table for the proposed system is given as under

Parameter	Quantity	Depth and unit
Network Size	Minimum 2000 nodes	989 adjacent Nodes
Number of Nodes	10000	UCI website derivation
Entropy	Single entropy for every attack	<=1
Classification Accuracy	Single classification accuracy for every simulation	>=0 and <=1
MSE	Single MSE value for every simulation	>=0 and <=1
Links or edges	Minimum 2000	989 edges between nodes

Table 1: Simulation parametric Table

The proposed work is focused on the security aspect of social media. Media Platform websites like facebook and twitter grant the users to create their accounts easily without strong verification. The proposed work suggests the strong verification process so that deception from the online social media can be reduced.

Classification Accuracy

Classification accuracy of proposed system with DCA gives better performance as compared to naïve bayes and support vector machine. Classification accuracy is obtained as under

$$Accuracy = 1 - MSE$$

Number of nodes	ExistingAlgorithm(Naïve Bayes%)	SVM Classification Accuracy(%)	Proposed Algorithm(DCA%)
50	90.33	91.265	95.22
100	91.44	92.255	95
150	91.99	92.322	96.44
200	92.66	93.255	97.45

Table 1 Classification Accuracy in tabular form of Naïve Bayes, SVM and Proposed DCA

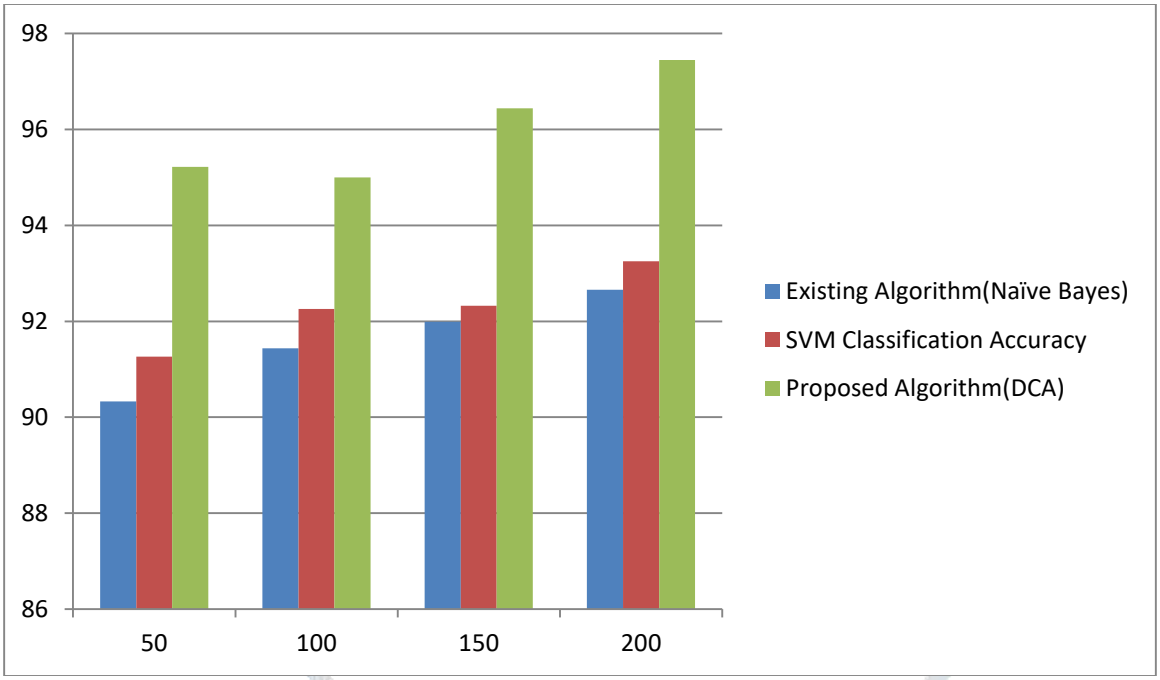

Figure 2: Plots of Classification Accuracy

MSE-Mean Square Error

Mean squared deviation" redirects here. It is not to be confused with Mean squared displacement.Instatistics, the **mean squared error** (**MSE**) or **mean squared deviation** (**MSD**) of an estimator (of a procedure for estimating an unobserved quantity) measures the average of the squares of the errors or deviations—that is, the difference between the estimator and what is estimated. Attack detection is the prime objective of proposed system.The classification accuracy indicates that attack is securely detected.

Number of nodes	Existing Algorithm(Naïve Bayes MSE %)	SVM MSE (%)	Proposed Algorithm(DCA MSE %)
50	9.67	8.235	4.78
100	8.56	7.222	5
150	8.01	6.256	3.56
200	7.34	6.125	2.55

Table 2: Plot of Mean square error

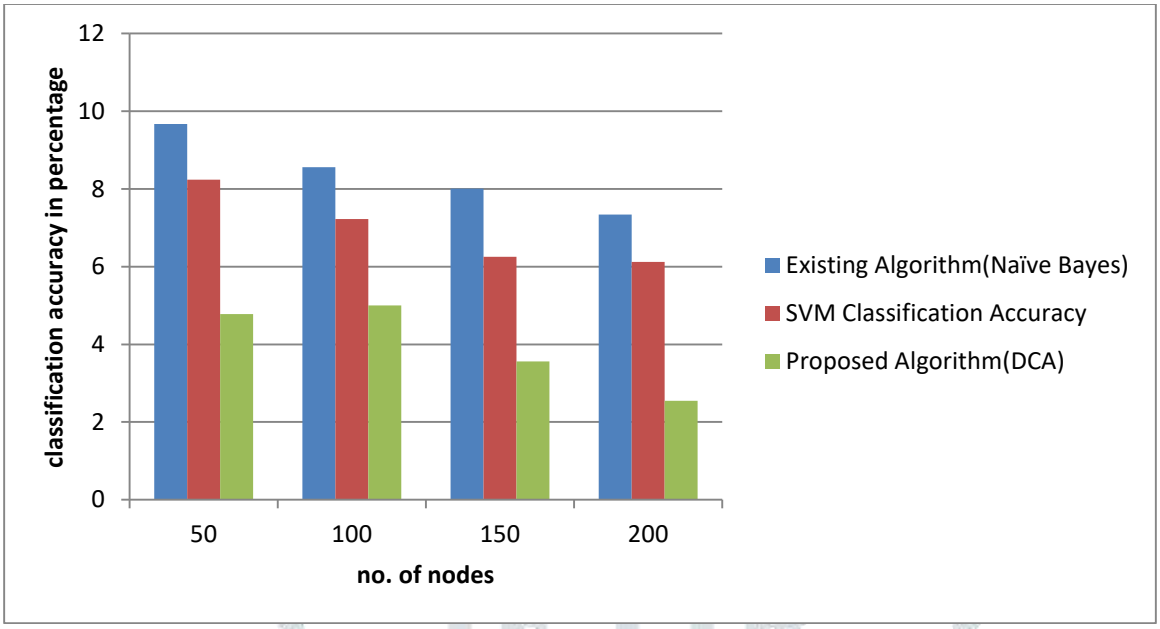

Figure 3: Mean square error of existing and proposed system

5. Performance Comparison

The performance comparison of existing system is done with DCA. The performance of proposed DCA method has been compared with the existing methods in terms of modularity,mean square error and classification accuracy. The result shown in Tabledemonstrates that the modularity and classification accuracy are better than existing method. The proposed method has gained the highest mean values among all of the competitors on each data set. On the other hand, the production of the proposed technique considering the value of modularity outshout the other technique in most of the times and equals to their results in other times. The reason for these better performances of the proposed method is that it considers modularity and optimizes it along with classification accuracy as a multi-objective task.

Dataset Attack Consideration	Modularity			Classification Accuracy			Mean Square Error		
	Naïve bayes	SVM	DCA	Naïve bayes	SVM	DCA	Naïve bayes	SVM	DCA
DDOS	10.235	15.021	17.256	90.33	91.265	95.22	9.67	8.235	4.78
U2R	9.2652	14.025	16.256	91.44	92.255	95	8.56	7.222	5
R2L	9.02356	13.2562	15.2652	91.99	92.322	96.44	8.01	6.256	3.56
Normal	8.0256	12.236	14.236	92.66	93.255	97.45	7.34	6.125	2.55
PRB	8.0001	12.0023	14.0025	90.33	91.265	95.22	7.24	6.025	2.45

Table 3:Resultant Modularity and Classification Accuracy Comparison of Proposed and Existing Methods

6. Conclusion

This paper is showing the refreshed self-retouching, Randomized, Efficient, and Distributed DCA-random walk estimation for the distinguishing proof of center point replication strikes when diverge from the Line-Selected Multicast and Randomized , Efficient, and Distributed traditions. The principal duty of this paper is the new suggestion of DCA-Random walk that is skilled for perceiving center replication strike while standing out from the .That DCA is more grounded in its area limits than Naïve Bias. We assume that the new method makes the capable and strong results in future still our examination is going on this territory.

References

[1] Reyaz Ahmad sheikh, "Detection of Clone Attack in Wsn\n," *IOSR J. Comput. Eng.*, vol. 16, no. 5, pp. 48–52, 2014.

[2] A. Kaur, "DDOS Attack Detection on Wireless Sensor Network using DSR Algorithm with Cryptography," vol. 175, no. 3, pp. 16–23, 2017.

[3] V. Subramanian, "Proximity-based attacks in wireless sensor networks," *Int. J. Sci. Eng. Res.*, vol. 3, no. May 2013, pp. 2–5, 2013.

[4] V. A. Khandekar, P. Singh, and G. Shrivastava, "Simulation Approach To Detect Clone Attack in," no. May, pp. 7–13, 2013.

[5] M. H. Ansari and V. Tabatabavakily, "Classification and A analysis of clone attack detection procedures in mobile wireless sensor networks," vol. 2, no. 11, pp. 1–7, 2012.

[6] W. Znaidi, M. Minier, and S. Ubéda, "Hierarchical node replication attacks detection in wireless sensor networks," *Int. J. Distrib. Sens. Networks*, vol. 2013, 2013.

[7] M. Conti, R. Di Pietro, L. V. Mancini, and A. Mei, "Distributed Detection of Clone Attacks in Wireless Sensor Networks," vol. 8, no. 5, pp. 685–698, 2011.

[8] N. Shruthi and C. K. Vinay, "Network Layer Attacks : Analysis & Solutions , A Survey," vol. 18, no. 2, pp. 67–80, 2016.

[9] G. Kontaxis, I. Polakis, S. Ioannidis, and E. P. Markatos, "Detecting Social Network Profile Cloning," *IEEE*, 2013.

[10] A. Malhotra, "Studying User Footprints in Different Online Social Networks."

[11] S. Y. Bhat and M. Abulaish, "Communities A gainst Deception in Online Social Networks 1 The Platform 2 The Mischef," *Ieee*, vol. 2014, no. 2, pp. 8–16, 2014.

[12] D. Dave, N. Mishra, and S. Sharma, "Detection Techniques of Clone Attack on Online Social Networks : Survey and Analysis," *Elsevier*, pp. 179–186.

[13] M. Kharaji and F. Rizi, "An IAC Approach for Detecting Profile Cloning in Online Social Networks," vol. 6, no. 1, pp. 75–90, 2014.

[14] F. S. Rizi, M. R. Khayyambashi, and M. Y. Kharaji, "A New Approach for Finding Cloned Profiles in Online Social Networks," *ACEEE Int. J. Netw. Secur.*, vol. 6, no. April, pp. 25–37, 2014.