International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-8 Issue-11, September 2019

An Efficient Network Threat Detection and Classification Method using Anp-Mvps Algorithm in Wireless Sensor Networks

P. Sherubha, N. Mohanasundaram

Abstract*: *Wireless Sensor Networks (WSNs) are deployed generally in a hostile environment, where an adversary captures some nodes that are physically connected in the network. It initially reprograms the nodes and makes them replicate into a number of clones, thereby having control over them. In order to provide a distributed solution to resolve the above specified problem specified above, a framework based on Authentic Node Placement based Message Verification and Passing Strategy (ANP-MVPS) is proposed. Some of the solutions offered by existing techniques are not satisfactory due to Energy and Memory constraints. This turns to be a serious drawback for protocols used in WSN's resource constrained environment. In this work, three diverse factors are considered for investigation. They are: Firstly, modeling of Authentic Node Placement based Message Verification and Passing Strategy (ANP-MVPS) is performed to identify the distributed mechanism of clone in a network and prevent the replication of clone among them. Secondly, the parameter selection Probability of Occurrence of IP, Mean Time Intervals, Time to Live, ACK value, Time Stamp Field, SYN value, Differentiated Service Field and Sequence Number are considered before performing classification. Thirdly, an efficient Naive Bayesian classifier for security analysis based on trust value (NB-TV) is used to estimate the performance metrics like accuracy, sensitivity, specificity, F-measure, Recall etc. This method shows satisfactory results when compared to existing techniques. The simulation was carried out in MATLAB environment. The proposed method shows better trade off in contrast to prevailing techniques.

Keywords*: *Wireless Sensor networks; Clone attack; Authentic Node Placement based Message Verification and Passing Strategy; Naive Bayesian classifier for security analysis; Accuracy, Trust values.

I. INTRODUCTION

Wireless sensor networks (WSNs) are generally a collection of sensors that provides restricted resources that accumulates to attain an ultimate objective. WSNs are deployed in harsh environments to satisfy both civil and military applications. Based on the operating nature of sensors, they are unattended very often; so it is prone to various sorts of attacks. For example, an adversary possesses the ability to eavesdrop on all network communications; henceforth, an adversary can capture all nodes in the network and acquire the information stored over them- sensors are usually considered to be a non-tamper-proof [1]. So, an adversary can manipulate captured sensors and deploy them to launch diverse malicious activities in network connectivity. These sorts of attacks are known as clone attack. As clone has legitimate information, it carries out network operations as that of non-compromised node as shown in figure 1; therefore, clone nodes can launch various other sorts of attacks [2]. Some of the attacks are provided in section II. For example, clone could generate black hole, it may introduce wormhole attack with collaborating adversary, and it inject aggregate or false data in a similar way to bias the final outcome. Also, clones can also leak data. Clones may act as bottom line for various attacks and cause severe impact in the network. The threat due to clone attack is characterized by two significant factors:

- In a given network model, the neighbourhood nodes assume clones as honest nodes, as there is no global countermeasure. Thus, clones replicate themselves among their neighbours.
- In order to have an enormous number of compromised nodes, adversaries does not need to compromise all the nodes in the network; instead, only one node can be captured and compromised; therefore, a significant attack cost is sustained. Generating multiple clones of a similar node is also measured as a cheap factor.

Based on the existing techniques, it is recognized that most protocol designs are either centralized or local protocols. Both types of protocols possess certain drawbacks in their own way and attempt to cope with a clone attack [3]. In general, centralized protocols have single point of failure along with high communication cost. Meanwhile, local protocols do not identify any replicated nodes that are distributed among the region of network model [4]. In this investigation, a network authentication mechanism is designed to identify the presence of clones and prevent them from further network activity. Specifically, this approach is designed to iterate certain parameters to analyze the network events. It is designed for constant monitoring without a significant effect in network performance, and increasing the identification rate of clone attack. Network parameters such as Probability of Occurrence of IP, Mean Time Intervals, Time to Live, ACK value, Time Stamp Field, SYN value, Differentiated service field and Sequence Number are considered for monitoring the network activity. In this work, an analysis of the desirable properties of distributing mechanisms for identification of node replication due to clone attack is performed. Initially, Authentic Node Placement based Message Verification and Passing Strategy (ANP-MVPS) is designed to authentic the node that is placed in the network and verify the incoming message for any data replication.

Manuscript published on 30 September 2019.
*Correspondence Author(s)
P. Sherubha[1], Department of Computer Science and Engineering, Karpagam Academy of Higher Education, Coimbatore, India. Email: sherubha0106@gmail.com
N.Mohanasundaram[2], Department of Computer Science and Engineering, Karpagam Academy of Higher Education, Coimbatore, India. Email: itismemohan@gmail.com

If so, the message will be terminated from the specific node and denied from further processing [5]. After the completion of the authentication mechanism, parameter selection is executed in order to enhance the performance of the network. Finally, an effectual Naive Bayes classifier is implemented to recognize the accuracy of threat detection, sensitivity, specificity, F-measure and recall. The simulation of the proposed work shows better trade off, as it provides a highly efficient communication process, computation and memory, with a higher probability rate of clone detection and compared with the existing techniques like SVM, SVM-NB, NB and so on.

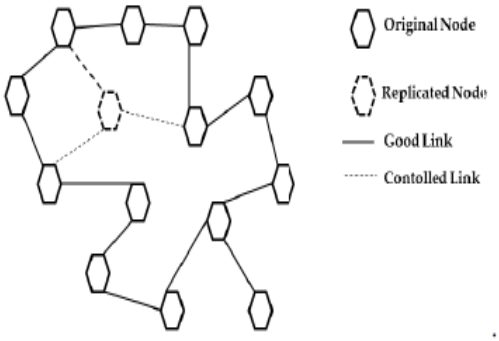

Fig 1: Clone attack scenario

The work is structured as follows: Section II describes the existing techniques that are associated with clone attack and the drawbacks encountered in it. , Section III provides a detailed explanation of the proposed method with the design of network model, threat detection, and parameter selection and finally classification based on Naive Bayes classification approach, Section IV demonstrates the numerical results, discusses parameters of the proposed outcome, comparing it with existing techniques like SVM, NB, SVM-NB and so on. Section V ends up the work with the idea for further extension of the research work. The reference section is provided at the end of the paper for designing a framework for detecting the clone attack in the network.

II. RELATED WORKS

Rakesh Rajendran et al. [6] depict that cloud network is modelled for communication and data storage that has been utilized in diverse industries and organizations. These industrial data are exposed to attacks that specifically target cloud related resources. Henceforth, IDS is designed to execute feature selection and classification for identification of DoS attacks. Also, the anticipated model utilizes expert's advice for generating final decision based on security, and this approach attains an accuracy of about 98.5% in terms of attack detection rate. The experimental outcome acquired with the execution proves that domain expert enhances the efficiency of the anticipated approach in contrast with the prevailing classification techniques. The foremost advantage of this technique comprises diminution of false positive rate and increase in security rate.

Kai Xing et al. [7] describe a new real-time detection algorithm against clone attacks. The proposed algorithm is better in terms of higher attack detection accuracy along with resiliency cost. It can be made with the least cost of communication/storage/computation overhead. Nonetheless, real time clone detection is performed whenever the message flows in the network model, and recognizing the cloned attackers can be done in a very effectual and efficient manner. On observing the existing technique, this approach is the first model to offer real time clone attack detection in sensor networks. Shivam Dhuria et al. [8] explain the data filtering and authentication mechanism for prevention and detection of DDoS attacks in wireless sensor networks. Compared with the existing techniques, this method is extremely simple and can be deployed effectually at every network node. A slight computation of tracing data rates from neighbourhood nodes eliminates complete battery source drainage that can be due to DDoS attacks. Kiruthiga et al. [9] describe that social networks are more popular among people to communicate with their friends via internet. Users spend their time in trending social networking sites such as Myspace, Facebook and twitter to share their information. Cloning attack is one among the insidious attack encountered in Facebook. Here, the attackers generally stole the personal information and images of certain specific persons and generate a fake profile pages. As the profile page gets cloned, they start transmitting friend request using the profile cloned. For instance, if real users account is blocked and they trend to create a new one, they send a friend request to persons in contact. In the meantime, the cloned profile also send friend request to the user. In those cases, it is very complex for the person to identify the real user. In the anticipated technique, clone attack is identified during the user action time period and clicks patters of users to recognize the similarity among the real one and cloned one in Facebook. With Jaccard index and cosine similarity, the performance is analyzed among users. Clones are identified and blocked for further communication. Xiaopeng Tan et al. [10] demonstrate that intrusion detection for wireless sensor networks is a significant subject in the security domain of WSN. Owing to the imbalance class in KDD Cup 99 dataset, this investigation combines SMOTE with random forest algorithm, and anticipates an ensemble classifier for imbalanced datasets. Experimentation carried out on KDD Cup 99 dataset depicts that classification accuracy of random forest algorithm has attained 92.39%, which is better than the existing classification techniques like LibSVM, J48, Naive Bayes, AdaboostM1 and bagging methods. After merging SMOTE with KDD Cup 99 dataset, the classification accuracy of random forest has risen to 92.57%, which enhances the classification efficiency of minority classes. Random forest techniques merged with SMOTE offer an effectual solution to resolve class imbalance and enhance the classification accuracy of intrusion detection. However, this technique is easier to execute and provides a stronger generalization capability. It can be extensively utilized in the region of security of wireless sensor networks to enhance the intrusion detection of wireless sensor networks. T. Shivaiah et al. [11] depicts that SNs require tamper-resistant hardware which generates key based routing, checking and caching capabilities for clone identification. This work utilizes probabilistic method to reduce communication overhead for satisfactory detection. Whilst a DHT-based protocol offers higher security for SNs using deterministic witness and probabilistic witnesses, memory-efficient, randomly directed examination offers excellent communication and reduced storage consumption in dense SNs.

P.Thiruvannamalai Sivasankar et al. [12] anticipate a nominal three-tier security architecture for pair-wise key establishment and authentication among mobile sinks and sensor nodes. Based on polynomial pool-based key pre-distribution strategy network resilience is contrasted with polynomial pool-based key pre-distribution method. With two separate key pools, stationary access nodes execute polynomials for accumulating data while deploying mobile sink.

Vandana Mohindru et al. [13] anticipate node authentication for securing WSNs from clone attack in sensor nodes. This technique utilizes lightweight operations such as addition, subtraction of bits and XOR which raises WSNs effectually. Subsequent nodes authentication algorithm presented uses an asymmetric and symmetric cryptography technique that increases energy utilization. The anticipated algorithm produces 24-bit encrypted UID broadcasted to receiver node and UID outcome in 4.416 µJ of computational overhead, 48.75 µJ of µJ of communication overhead and 3 bytes of storage overhead is lesser than existing authentication algorithm. Also, algorithm possesses lesser success rate authentication, henceforth detection probability is higher and identification of cloned node is easier. The designed model is appropriate for energy constraint in WSNs.

A Vanathi et al. [14] depicts a security model to resolve three significant active attacks termed as MITM attack, cloning attack and Replay attack. Zero knowledge protocol idea is utilized to guarantee non-transmission of data between verifier and prover. The anticipated model utilizes a social finger print based on s-disjunct code together with ZKP to identify clone attacks and eliminate replay and MITM attack [15]. This scheme is extended to evaluate clone node detection. Various attack scenarios also analyze the performance and cryptographic strength of anticipated model.

III. PROPOSED METHODOLOGY

In this section, the proposed model is discussed in detail. The proposed model comprises more subcategories. Firstly, a network model for wireless sensor networks is designed with its deployment strategies. Secondly, a threat model for detecting the clone node is designed. Thirdly, an Authentic Node Placement based Message Verification and Passing Strategy (ANP-MVPS) is designed. The subsequent process is the parameter selection to identify the network performance. Parameter selection is carried out using CIADA dataset; with these dataset parameters such as MTI, POIP, TTL, SYN value, ACK value, Differentiated service field, Time Stamp Field and Sequence Number (SN) are considered. Finally, an effectual Naive Bayes classifier is considered for computing trust value and the associated performance metrics like accuracy, sensitivity, specificity, recall and F-measure. The outcome of the proposed design is effectual when compared to existing techniques.

a. Network Model

Consider a sensor node model with ‘n’ SNs, deployed randomly. Every node possesses radius ‘R’ for transmission. However, perfect communication cut off will be accomplished for real time environment; radius ‘R’ shows a circular region where nodes can be reached [17]. Due to this, outcomes attained using network model refer to lower bound detection in real time deployment. In formal computation, sensors indicate accuracy and notion of rounds as elapsed time. Every node can recognize its location (for example, GPS or protocol design). This work does not rely on real / quasi-real time clone detection. Sensor networks can accept a certain level of detection delay based on rounds. For instance, election or voting procedures are executed with intervals that last for ten rounds. But, delay is lesser than existing techniques. In SNs there are diverse mobility models for sensors: Random Way-Point mobility model [RWPMM] and two-dimensional mobility model [2DMM] [18]. In the latter model, sensors are randomly placed; hence every node position is independent of previous round. Features should posses’ average number of neighbours with small variance in all rounds [19]. This model is simple, and results drawn from it is extremely indicative of nodes’ behaviour. For example, a two-dimensional mobility model offers more reasonable outcome than the former model, when the elapsed time between the two models is identified. RWPMM is a simple mobility model utilized for describing the movement patterns of independent nodes. With this Random Way-Point mobility model, nodes are selected randomly from the network with random speed. This helps in a stable movement towards the chosen points. When a node reaches destination, it replicates two random selections and velocity for the next movement.

b. Threat model

Consider a threat model with an adversary, which concentrates on undetected replicated sensors in random environment. In general, clone attack differs from DoS attack and Byzantine failure. Clone nodes are not considered as Byzantine failure, as it does not work arbitrarily to enhance clone detection probability [20]. In addition to this, an adversary can realize replication attack that places clone nodes over the network. In Sybil attack, adversary uses one node to impersonate other nodes – nodes with numerous compromised nodes. Malicious node collusion is not considered, as this work concentrates on clone detection. However, clone sensor is revealed by multiple node connectivity at the same time and based on simple verification to collect the clues that reveal the clones’ suspicion. Assume that N_r is a replicated node with ‘c’ replicas, where ‘c’ denotes nodes that possess similar ID of N_r (comprising of original N_r itself). There are some common rounds that are encountered in the network [21]. Moreover, consider two diverse kinds of adversaries, termed as vanishing / persistent adversaries that vary in the number of rounds that it operates. The former persistent adversary attempts to clone random sensors and upholds constant control over clones continuously until termination. Indeed, vanishing adversary utilizes cloned nodes during successive round. It is introduced before ‘r’ and removed when ‘r’ expires. It is essential to consider both the vanishing and persistent adversary; the vanishing one is more complex to identify. While vanishing nodes leave their traces, the detection rate against these adversaries is lower bound for detection [22]. It is recognized that vanishing adversary works smarter as it exploits clones during rounds, for instance, decision is taken based on existing consensus or voting algorithm like leader election.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

It is very complex to identify the clone, in both the vanishing and persistent scenario. Instead, more network models attempts to enhance the chance of detecting the clone, and therefore replicated ID revocation will happen. Worst case detection is also analyzed for evaluating the performance of the anticipated method.

c. Pre-processing based Authentic Node Placement based Message Verification and Passing Strategy (ANP-MVPS)

In the proposed method, the activity of clone in the network is identified and prevented using Authentic Node Placement based Message Verification and Passing (ANP-MVPS) Strategy. To prevent clone activity, nodes in the network has to be authenticated and message passing has to be identified for clone replication. If verification fails, the message will not be forwarded to the neighbourhood nodes in the network. Authentic Node Placement based Message Verification and Passing (ANP-MVPS) Strategy works both in unicast and multicast communication models in the network. The functionality of the proposed strategy is given below:

Algorithm 1:

1. Consider S = $\{N_1, N_2, N_3, \ldots, N_n\}$, $(N_i, N_i + 1)$
2. Let BS, configured nodes
3. For i=1 to n // random placement of nodes
4. N_i ← Position (random (X), random (Y))
5. Source ID (N_i ← i);
6. End Loop
7. Let C_i be connectivity established between network nodes
8. N_i → BS (message, λ) // for node N_i
9. L^* → least distance
10. N_i → Re-establish base station;
11. Re-establish BS = ID, X, Y, λ;
12. Re-establish → information table;
13. End
14. $S_u\{L^*\}$ → S
15. S → Route transmission (S)
16. $N_i + 1$ → N_i (λ)
17. λc = (ID, X, Y, λ)
18. L^* → D
19. $for\ i = S\ to\ D$ // Route Discovery
20. Route transmission (S) N_i;
21. Re-establish → Routing table
22. End
23. $for\ i = S\ to\ D$ // Route Transmission
24. If (current N_i information == Routing table == node information table)
25. Then
26. N_i data $N_i + 1$
27. Else
28. $N_i + 1$ is blocked as it replicates clone
29. End
30. Routing table Entries = clear;
31. End

Authentic Node Placement based Message Verification and passing algorithm is utilized for route discovery and data transmission, from BS. After the verification process, the proposed model collects the time stamp, ID and placement details of nodes. The collected information will be compared with initial information during the process of registration. The outcomes of Authentic Node Placement based Message Verification and passing strategy can offer trusted nodes to guarantee secure data transmission [23]. Else, the particular node is considered as unknown node or clone, in which data transmission in the current route will be blocked and an alternative path has to be selected for further processing.

Algorithm 2:

1. Consider S = $\{N_1, N_2, N_3, \ldots, N_n\}$
2. For I= 1 to n
3. Routing table = add $(N_i\ (id), N_i\ (X), N_i\ (Y), N_i\ (msg))$ // node id, X, Y value of i^{th} node
4. End
5. For I = 1 to n
6. For J = 1 to n
7. N_i → transmit (ACK_REQ) → N_j
8. N_j → transmit (ACK_MSG) → N_i
9. $If\ \left(msg\ (N_i), msg\ N_j\ exists\ (RT)\right)\ then$
10. N_i → transmit data to N_j
11. $Else$
12. Choose subsequent neighbour
13. End for
14. End for
15. End

Authentic node placement is a time consuming process and also leads to delay. Therefore, the prevention device is suggested to eliminate the activities of clone in the network. Every node should communicate with one another by passing the authentication message. If source node identifies destination dynamically, the Message Verification and passing strategy is utilized for comparing and authenticating whether the current node is a clone or not. Generally in a network, when a node passes data to other nodes, it initially establishes request message to the node with its key [24] [25]. The key is generated by base station during network registration. Destination node has to surrender its key message, and keys should be verified by BS and ACK signal has to be produced for sharing data [26][27]. Data transmission occurs only after the processing of original signal form base station [28]. The algorithm for Message Verification and passing is given in Algorithm 2.

d. Parameter selection

After examining datasets completely, choose certain essential parameters for classification. Parameters such as MTI, POIP, TTL, SYN value, Time Stamp Field, ACK value, Sequence Number (SN) and Differentiated service field are considered for classification process. MTI is Mean Time Intervals of IP in window; 'T' is arrival time of IP. POIP is probability of time intervals of IP. Values of these entries are considered in accordance to IP packets in CIADA dataset. Significant attributes are considered for performing Naive Bayes classification.

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

e. Naive Bayes (NB) classifier

In this section, NB classifier is utilized for identifying the severity of clone attacks in the sensor environment. The assumptions of Naive Bayes classifier provides enhanced outcome, as Authentic Node Placement based Message Verification and Passing is performed in the pre-processing stage of investigation. The pre-processing strategy enhances the performance of SN.

With respect to Bayes theorem, classification algorithm of NBs is evaluated with strong assumption. NBs follows simple hypothesis in which some features are unrelated to each other and some features are independent of hypothesis space. Therefore, when sample 'X' owns features X_i and X_j correspondingly, $(X_i \neq X_j)$, sample probability 'X' is classified to Y is $P(X|Y) = P(Y|X_i, X_j) = P(Y|X_i) * P(Y|X_j)$

Probability of detection model for Naive Bayes classifier for trust value computation is described below:

1) Naive Bayes classifier depicts that, sample Y is a random Boolean variable; X is n-D Boolean vector and specified as $X = < X_1, X_2, \ldots, X_n >$, where X_i is an i[th] random Boolean attribute. In accordance to, Bayesian theorem as in Equation 1:

$$P(Y|X) = \frac{P(X|Y) * P(B)}{\sum_{i=1}^{n} P(X_i|Y) * P(Y)} \quad (1)$$

2) Consider a target function $f: X \rightarrow Y$, most probable value of Y can be deduced with $f(x): V_{MVP} = argmax\ P(Y|X_1, X_2, \ldots, X_n)$

3) Naive Bayes classifier for trust value can be signified as $posterior = probability * \frac{prior}{marginal\ probability}$

$$V_{MVP} = argmax \frac{P(Y) * P(X_1, X_2, \ldots, X_n|Y)}{P(X_1, X_2, \ldots, X_n)} \quad (2)$$

4) As denominator $P(X_1, X_2, \ldots, X_n)$ specifies the feature values X_i, which have been provided (value is provided as 1 or constant). It does not depend on Y as in Equation 3:

$$V_{MVP} = argmax\ P(Y) * P(X_1, X_2, \ldots, X_n|Y) \quad (3)$$

5) Conditional independence specifies every feature X_i is independent of subsequent features X_j. Eventually, the equation obtained is given in Eq. (4):

$$V_{NBTV} = argmax\ P(Y) * \prod_{i=1}^{n} P(X_i|Y) \quad (4)$$

While applying Naive Bayes classifier to clone attack detection, Naive Bayes can construct Trust valued inputs. It is evaluated that the classification output is effectively a supervised learning process, by virtue of larger amount of datasets.

f. Trust Value computation of NB classifier

In this example, certain parameters are considered for performing the target classification. Noticeably, problem encountered here is the utilization of training data merged with chosen parameters to identify activities as "Clone Encountered", or "Generic node". When new objects such as repeated IP address, packet from similar nodes are identified, the decision can be generated based on formulating the priority probability. Initially,

Priority probability for 'Clone Encountered' P (Clone) = 5/7 = 0.714

Priority probability of 'generic node' P (Non-clone) = 3/7 = 0.4285

Following this, the conditional probability for every feature has to be evaluated. For attribute outlook, consider,

P (POIP | yes) = 2/4 = 0.5
P (POIP | No) = 2/3 = 0.6667

Likewise, for other two attributes SYN value and SN value, conditional probabilities can be provided as follow:

P (SYN | Yes) = ¼ = 0.25;
P (SYN | No) = 3/3 =1
P (SN | Yes) = ¼ = 0.25;
P (SN | No) = 2/3 = 0.6667

Therefore, with Naive Bayes classifier, posterior probability of attack identification *P (Yes | New) = P (Yes) * P (POIP | Yes) * P (SYN | Yes) * P (SN |Yes) = 0.714 * 0.5 * 0.25 * 0.25 = 0.02231*. Similarly, probability of generic node in network connectivity is *P (No | New) = P (No) * P (POIP | No) * P (SYN | No) * P (SN |No) = 0.4285 * 0.6667 * 1 * 0.0667 = 0.1905*. After the evaluation, comparing the probability of occurrence is evaluated with *P (No | New) = P (Yes | New)*, which specifies the occurrence of clone nodes in the network or simply the generic node connected to the network.

When every node in the network is deployed, nodes have the ability to gather entire trust information from 1H neighbours and employ trust assistant. With respect to the collected trust information, node encounters minimum and maximum trust value as T_{min} and T_{max} with the neighbourhood nodes. Then, it determines the trust value among T_{min} and T_{max} as trust threshold T_{limit} to employ trust assistants while examining trust value of other nodes. Henceforth, sum of trust value helps in selecting the trust assistants as in Eq. (5) & (6):

$$N_i \in Trust\ value \quad (5)$$
$$T_{N_i} \geq T_{limit} \quad (6)$$

Where, $T_{min} \leq T_{limit} \leq T_{max}$. When node evaluates the neighbourhood nodes' trust value, it will query the trust assistance T_{A_i} about the neighborhood node Y.

$$N_i\ \overrightarrow{Query\ (Y)}\ T_{A_1} \quad (7)$$

The assistant T_{A_i} will offer Y's trust value in neighbourhood to evaluate nodes N_i.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

$$T_{A_i} \xrightarrow[T(Y)]{} N_i \quad (8)$$

After determining trust value T (Y) from the trust assistantT_{A_i}, NBs algorithm is used to identify accuracy of trust information with respect to trust training data. During parameter selection, three features are considered.

Variance and Mean of previous trust value of neighborhood node is VTV and MTV correspondingly. SD is specified as SDTV. With trust value N_i, evaluate variance of trust value to mean of previous trust value, for instance, VTV ≥ SDTV or VTV < SDTV. The conditional probability of the features is provided as P [(VTV ≥ SDTV) | true] and P [(VTV< SDTV) | false]

Hence, with Naive Bayes classifier, the N_i determines the posterior probability of newly received trust value P (True | New) = P (True) * P (VTV ≥ SDTV) | True] * P [(VTV ≥ SDTV) | true]. In addition, P (False | New) = P (False) * P (VTV ≥ SDTV) | False] * P [(VTV ≥ SDTV) |False]. Next, N_i compare P (True | New) and P (False | New). If P (True | New) > P (False | New), it determines that new trusted value is reliable.

After identifying the trust value reliability from trust assistant, T_{A_i} evaluating node requires diverse techniques to calculate final trust value. Then, trust threshold limit is provided for detecting the malicious neighbourhood nodes in network. Trust value of node is considered as a secure member.

$$Y \neq N_i \quad (9)$$

$$T_{F(Y)} < T_{not} \quad (10)$$

With obtained result, N_i takes decision, for instance if "Y' fulfils trust value requirement ofN_i, then 'Y' will be included in communication process of N_i else it will be eliminated. As an evaluation node does not take decision on its own information, it may use its trust assistants' information. Evaluation node determines reliable and objective decisions of nodes.

IV. SIMULATION SETUP AND DISCUSSION

Here, experimental results of our ten machine learning techniques with five class classification methodologies using CIADA intrusion detection dataset are provided in order to detect network intrusions and compare them with the existing approaches to evaluate the efficacy of our network intrusion detection model. This work use MATLAB for our experimentation in Pentium-4 Machine with 2.86GHz CPU, and 1GB RAM. The framework of the proposed approach is shown in Figure 1.

All the experiments are conducted using CIADA dataset that has 60438 training instances, 22544 instances for testing with 42 attributes and 38 attack types for five class classifications to build an efficient network intrusion detection system.

The definitive objective is to estimate the classification performance of attacks in sensor environment. It is essential to plot normal activities and incoming events for clone detection; some performance metrics are estimated for results acquired from CIADA datasets.

1) True Positive: Nodes that are appropriately classified to network node.

2) False Positive: Nodes are incorrectly classified as appropriate node as it is a clone.

3) Sensitivity = It measures clones' proportion that is correctly identified in the network. It determines probability of correctly recognizing clone from subset.

$$Sensitivity = \frac{Number\ of\ true\ positive\ assessments}{Number\ of\ all\ positive\ assessments} \quad (11)$$

4) Specificity = It measures original nodes' proportion that are correctly recognized. It determines probability of appropriately recognizing original class from negative class subset.

$$Specificity = \frac{Number\ of\ true\ negative\ assessments}{Number\ of\ all\ negative\ assessments} \quad (12)$$

5) Precision: It is distinct as fraction of elements classified appropriately as positive to sum of true positive and false negative elements.

$$Precision = \frac{True\ positive}{True\ positive + False\ negative} \quad (13)$$

6) Accuracy: It is ratio of appropriately recognized instances to total instances.

$$Accuracy = \frac{Number\ of\ correct\ assessment}{Number\ of\ all\ assessments} \quad (14)$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (15)$$

7) F-measure = F-measure is utilized as a measure of test performance for positive class. It is harmonic mean of recall and precision.

$$F-measure = 2 * \frac{Precision * sensitivity}{Precision + sensitivity} \quad (16)$$

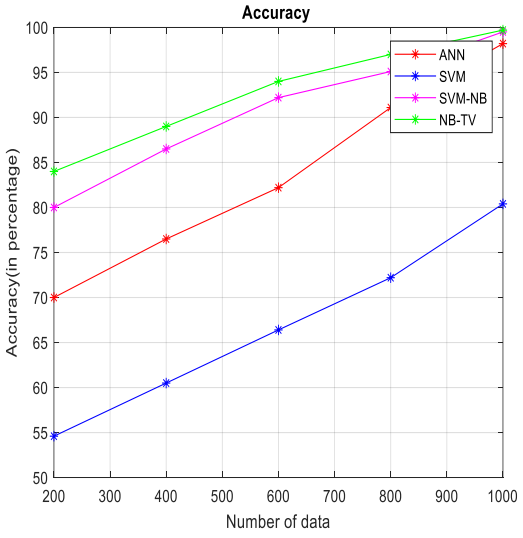

Fig 2: Graphical representation of Number of data Vs Accuracy rate

Figure 2 shows the graphical representation of accuracy computation of the proposed NB+TV with the existing techniques such as ANN, Naive Bayes and SVM.

The accuracy attained due to the proposed approach is better than the existing techniques. SVM-NB attains 80% accuracy while ANN is 70%, Naive Bayes is 74%, SVM is 54.6 for 200 data respectively. With respect to 1000 data, the accuracy attained is 99.5% for the proposed method.

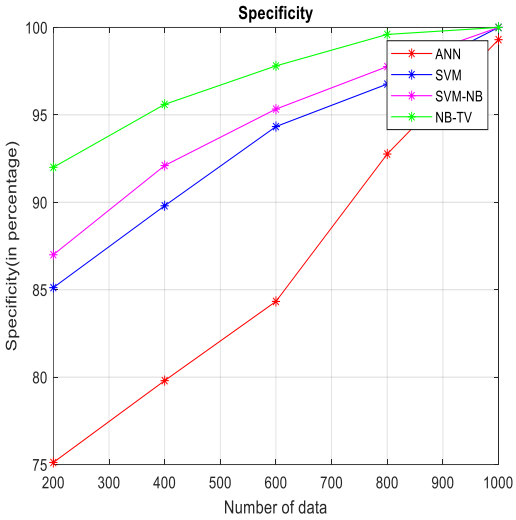

Fig 3: Graphical representation of Number of data Vs Specificity

Figure 3 shows the graphical representation of specificity computation of the proposed NB+TV with the existing techniques such as ANN, Naive Bayes and SVM. Specificity attained due to the proposed approach is better than the existing techniques. SVM-NB attains 87% specificity while ANN is 75.12%, Naive Bayes is 72.8%, and SVM is 85.12 for 200 data respectively. With respect to 1000 data, the accuracy attained is 100% for the proposed method.

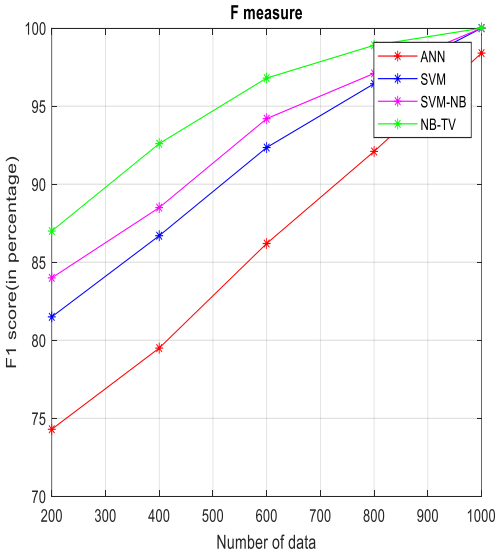

Fig 4: Graphical representation of Number of data Vs F1 score

Figure 4 shows the graphical representation of F1 score computation of the proposed NB+TV with the existing techniques such as ANN, Naive Bayes and SVM. F1 score attained due to the proposed approach is better than the existing techniques. SVM-NB attains 84% F1 score while ANN is 74.3%, Naive Bayes is 72%, and SVM is 81.5 for 200 data respectively. With respect to 1000 data, the accuracy attained is 100% for the proposed method.

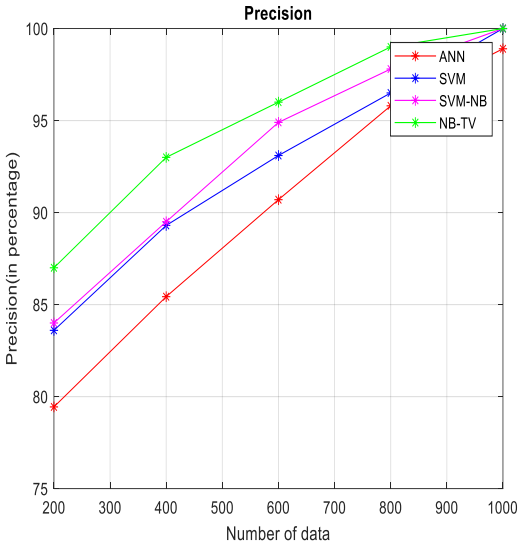

Fig 5: Graphical representation of Number of data Vs Precision

Figure 5 shows the graphical representation of Precision computation of the proposed NB+TV with the existing techniques such as ANN, Naive Bayes and SVM. Precision attained due to the proposed approach is better than the existing techniques. SVM-NB attains 84% Precision while ANN is 79.44%, Naive Bayes is 81%, and SVM is 83.6 for 200 data respectively. With respect to 1000 data, the accuracy attained is 100% for the proposed method.

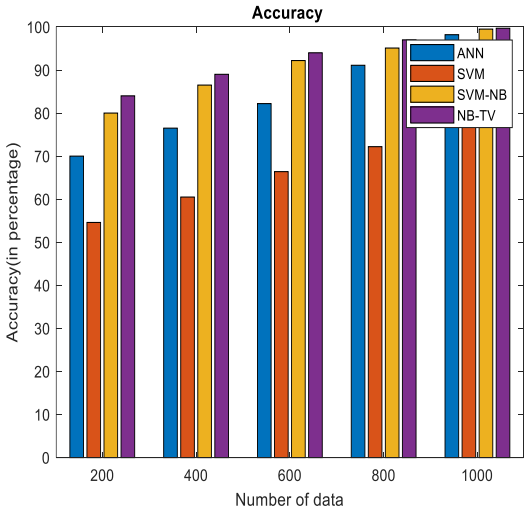

Fig 6: Comparison chart of data Vs Accuracy rate

Figure 6 depicts graphical representation of Accuracy of NB+TV with the existing techniques like ANN, Naive Bayes and SVM. Accuracy attained due to the proposed approach is better than the existing techniques. With respective to 1000 data, the accuracy attained is 100% for the proposed method.

Table-I: Accuracy comparison of NB+TV with existing techniques

Number of data	200	400	600	800	1000
ANN	70	76.5	82.2	91.1	98.2

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

SVM	54.6	60.5	66.4	72.2	80.4
SVM-NB	80	86.5	92.2	95.1	99.5
NB-TV	84	89	94	97	99.7

Table I shows the Accuracy comparison of NB+TV with the existing techniques like ANN, SVM, NB+TV. The accuracy of the existing techniques with 1000 data is 98.2%, 80.4%, 99.5% correspondingly. In addition, the proposed NB+TV offer 99.7% accuracy respectively. The proposed approach shows better trade off than the prevailing techniques.

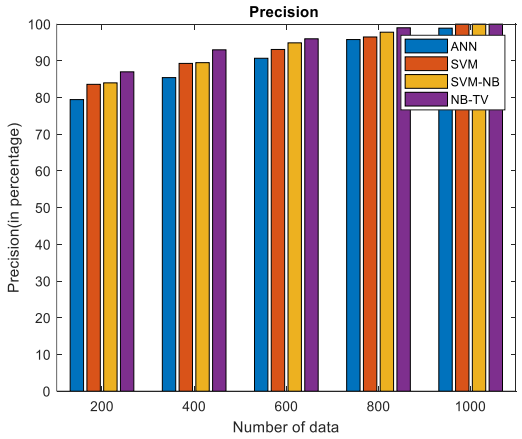

Fig7: Comparison chart of data Vs Precision

Figure 7 shows the graphical representation of Precision of the proposed NB+TV with the existing techniques such as ANN, Naive Bayes and SVM. Precision attained due to the proposed approach is better than the existing techniques. With respect to 1000 data, precision attained is 100% for the proposed method.

Table-II: Precision comparison of NB+TV with existing techniques

Number of data	200	400	600	800	1000
ANN	79.44	85.43	90.7	95.8	98.9
SVM	83.6	89.3	93.1	96.5	100
SVM-NB	84	89.5	94.9	97.8	100
NB-TV	87	93	96	99	100

Table II depicts the Precision comparison of NB+TV with the existing techniques like ANN, SVM, SVM+NB. The Precision of the existing techniques with 1000 data is 98.2%, 100%, 100% correspondingly. In addition, the proposed SVM+TV offer 100% accuracy. The proposed approach shows better trade off than the prevailing techniques.

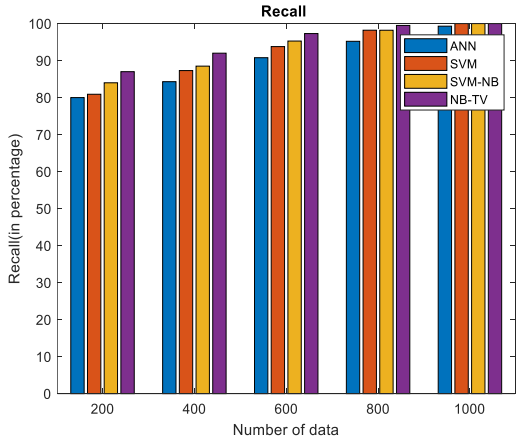

Fig 8: Comparison chart of data Vs Recall

Figure 8 shows the graphical representation of Recall of the proposed NB+TV with the existing techniques such as ANN, Naive Bayes and SVM. Recall due to the proposed approach is better than the existing techniques. With respect to 1000 data, Recall is 100% for the proposed method.

Table-III: Recall comparison of NB+TV with existing techniques

Number of data	200	400	600	800	1000
ANN	80	84.3	90.78	95.21	99.3
SVM	80.9	87.30	93.78	98.21	100
SVM-NB	84	88.5	95.28	98.2	100
NB-TV	87	92	97.3	99.5	100

Table III depicts the Recall comparison of NB+TV with the existing techniques like ANN, SVM, SVM+NB. The Recall of the existing techniques with 1000 data is 99.3%, 100%, 100% correspondingly. In addition, the proposed SVM+TV offer 100% Recall. The proposed approach shows better trade off than the prevailing techniques.

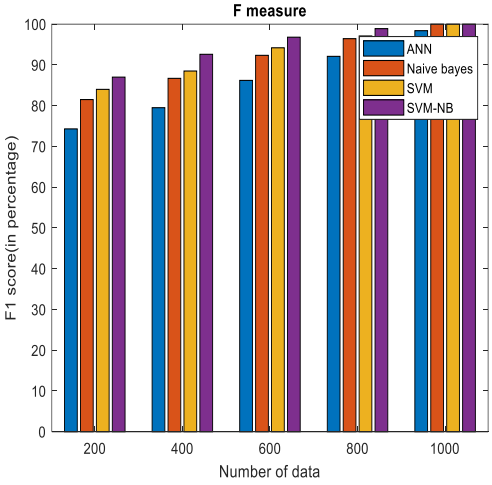

Fig 9: Comparison chart of data Vs F1 score

Figure 9 shows the graphical representation of F1 score of the proposed NB+TV with the existing techniques such as ANN, Naive Bayes and SVM. F1 score attained due to the proposed approach is better than the existing techniques. With respective to 1000 data, F1 score attained is 100% for the proposed method.

Table-IV: F-measure comparison of NB+TV with existing techniques

Number of data	200	400	600	800	1000
ANN	74.3	79.5	86.2	92.1	98.4
SVM	81.5	86.7	92.35	96.44	100
SVM-NB	84	88.5	94.2	97.1	100
NB-TV	87	92.6	96.8	98.9	100

Table IV depicts F-measure comparison of proposed NB+TV with the existing techniques like ANN, SVM, SVM+NB. F-measure of the existing techniques with 1000 data is 98.4%, 100%, 100% correspondingly. In addition, the proposed SVM+TV offer 100% F-measure. The proposed approach shows better trade off than the prevailing techniques.

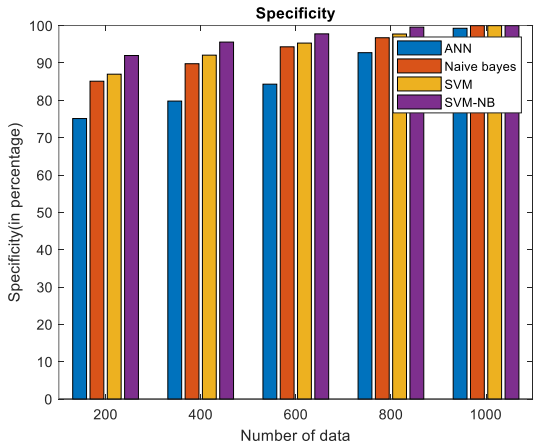

Fig 10: Comparison chart of data Vs specificity

Figure 9 shows the graphical representation of specificity of NB+TV with the existing techniques such as ANN, Naive Bayes and SVM. Specificity attained due to the proposed approach is better than the existing techniques. With respect to 1000 data, specificity attained is 100% for the proposed method.

Table-V: Specificity comparison of NB+TV with existing techniques

Number of data	200	400	600	800	1000
ANN	75.12	79.8	84.33	92.76	99.3
SVM	85.12	89.8	94.33	96.76	100
SVM-NB	87	92.1	95.33	97.76	100
NB-TV	92	95.6	97.8	99.6	100

Table IV depicts Specificity comparison of NB + TV with the existing techniques like ANN, SVM, SVM+NB. Specificity of the existing techniques with 1000 data is 99.3%, 100%, 100% correspondingly. In addition, the proposed SVM+TV offer 100% Specificity. The proposed approach shows better trade off than the prevailing techniques.

Table-VI: Execution time of SVM-NB with existing techniques

Algorithms	Execution time (in Sec)
ANN	20.644
SVM	29.07
SVM-NB	29.18
NB-TV	27.35

Table VI shows the execution time of NB+TV with the existing techniques. The execution time of ANN, SVM, SVM-NB techniques are 20.644s, 29.07s, and 29.18s correspondingly. The execution time of NB+TV is 27.35s.

V. CONCLUSION

In past decades, identifying the presence of clone attack with 100% accuracy has been a challenging task. With selected parameters, Naive Bayes classifier with trust value gives 99.7% accuracy. The execution time of anticipated work is about 27.35 Secs. Henceforth, with the acquired results it is identified that anticipated work outperforms prevailing techniques based on accuracy, sensitivity and specificity. The proposed NB-TV identifies the probability of occurrence of the clone nodes in the network based on parameters like sequence number, SYN value and probability of occurrence of IP address.

The trust-value-based application plays a considerable role in providing a high level security in wireless sensor networks. Here, a trust value is generated to monitor the activities of nodes in the network. The feasibility of utilizing Naive Bayes classifier is analysed and used to recognize clones' presence in network. Misbehaving nodes in the network are eliminated and no further processing is performed in route and therefore successive node will be selected for transmission. The performance metric such as Accuracy, Sensitivity, Specificity, F-measure and recall is computed in this investigation with respective to CIADA dataset. The accuracy attained is 99.7%. The proposed method outperforms the existing techniques. The future direction will be blacklisting the nodes that cause traffic and congestion in the network using effectual machine learning algorithms. The abnormal IP will be detected and fed to server, which can block such IPs while they try to make further processing.

REFERENCES

1. B. Zhu, V.G.K. Addada, S. Setia, S. Jajodia, and S. Roy, "Efficient Distributed Detection of Node Replication Attacks in Sensor Networks," Proc. Ann. Computer Security Applications Conf. (ACSAC '07), pp. 257-266, 2007.
2. Ficco, M., Rak, M.: Stealthy denial of service strategy in cloud computing. IEEE Trans. Cloud Comput. 3(1), 80–94 (2015)
3. Arul Xavier, V.M., Annadurai, S.: Chaotic social spider algorithm for load balance aware task scheduling in cloud computing. Clust. Comput. (2018).
4. Patel, A., Taghavi, M., Bakhtiyari, K., Ju´nior, J.C.: An intrusion detection and prevention system in cloud computing: a systematic review''. J. Netw. Comput. Appl. 36(1), 25–41 (2013)
5. Kim, H.-Y.: An energy-efficient load balancing scheme to extend lifetime in wireless sensor networks. J. Clust. Comput. 19, 279–283 (2016)
6. Rakesh Rajendran, S. V. N. Santhosh Kumar, "Detection of DoS attacks in cloud networks using intelligent rule based classification system", Cluster Computing https://doi.org/10.1007/s10586-018-2181-4
7. Kai Xing, "Real-time Detection of Clone Attacks in Wireless Sensor Networks", International Conference on Distributed Computing Systems, IEEE 2008.
8. Shivam Dhuria and Monika Sachdeva, "Detection and Prevention of DDoS Attacks in Wireless Sensor Networks", Springer Nature Singapore Pte Ltd. 2018
9. Kiruthiga. S, "Detecting Cloning Attack in Social Networks Using Classification and Clustering Techniques", International Conference on Recent Trends in Information Technology, IEEE 2014.
10. Xiaopeng Tan, Shaojing Su, "Wireless Sensor Networks Intrusion Detection Based on SMOTE and the Random Forest Algorithm", Sensors 2019, 19, 203; doi:10.3390/s19010203
11. T. Shivaiah, "On the Node Clone Detection in Wireless Sensor Networks", IJCSMC, Vol. 3, Issue. 9, September 2014, pg.176 – 189
12. P.Thiruvannamalai Sivasankar, "Active key management scheme to avoid clone attack in wireless sensor network", ICCCNT 2013 July 4-6, 2013, Tiruchengode, India
13. Vandana Mohindru, "Node authentication algorithm for securing static wireless sensor networks from node clone attack", Int. J. Information and Computer Security, Vol. 10, Nos. 2/3, 2018
14. 1A Vanathi, "Cloning Attack Authenticator in Wireless Sensor Networks", IJCST Vol. 3, Iss ue 1, Spl. 5, Jan. - March 2012
15. Reyaz Ahmad sheikh, "Detection of Clone Attack in WSN", IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 16, Issue 5, Ver. V (Sep – Oct. 2014), PP 48-52
16. Aalsalem, M.Y., Khan, W.Z., Saad, N.M., Hossain, M.S., Atiquzzaman, M. and Khan, M.K. (2016)'A new random walk for replica detection in WSNs', PloS one, Vol. 11, No. 7, p.e0158072.
17. Atmel Corporation Product Document (2009) [online] http://www.atmel.com/dyn/resources/ prod_documents/doc2467.pdf (accessed 25 January 2017).

18. Baronti, P., Pillai, P., Chook, V.W., Chessa, S., Gotta, A. and Hu, Y.F. (2007) 'Wireless sensor networks: a survey on the state of the art and the 802.15. 4 and ZigBee standards', Computer Communications, Vol. 30, No. 7, pp.1655–1695.
19. Cho, K., Jo, M., Kwon, T., Chen, H.H. and Lee, D.H. (2013) 'Classification and experimental analysis for clone detection approaches in wireless sensor networks', IEEE Systems Journal, Vol. 7, No. 1, pp.26–35.
20. Conti, M., Di Pietro, R., Mancini, L.V. and Mei, A. (2007) 'A randomized, efficient, and distributed protocol for the detection of node replication attacks in wireless sensor networks', in Proceedings of the 8th ACM International Symposium on Mobile Ad hoc Networking and Computing, ACM, September, pp.80–89.
21. De Meulenaer, G., Gosset, F., Standaert, F.X. and Pereira, O. (2008) 'On the energy cost of communication and cryptography in wireless sensor networks', in IEEE International Conference on Wireless and Mobile Computing Networking and Communications, WIMOB'08, IEEE, October, pp.580–585.
22. Akyildiz, I.F., Su, W., Sankarasubramaniam, Y. and Cayirci, E. (2002) 'Wireless sensor networks: a survey', Computer Networks, Vol. 38, No. 4, pp.393–422.
23. Amor, N.B.; Benferhat, S.; Elouedi, Z. Naive Bayes vs decision trees in intrusion detection systems.In Proceedings of the ACM Symposium on Applied Computing, Nicosia, Cyprus, 14–17 March 2004;pp. 420–424.
24. P. Sherubha, "A Detailed survey on security attacks in Wireless sensor networks", International Journal of Soft Computing 11 (3), 221-226
25. P. Sherubha, "Energy efficient and bandwidth aggregation techniques in Wireless heterogeneous devices: A Survey", International journal of pure and applied mathematics, 2018.
26. P. Sherubha, "Clone Attack Detection using Random Forest and Multi Objective Cuckoo Search Classification", International Conference on Communication and Signal Processing, April 4-6, 2019, India
27. P. Sherubha, "An Efficient Intrusion Detection and Authentication Mechanism for Detecting Clone Attack in Wireless Sensor Networks", Jour of Adv Research in Dynamical & Control Systems, Vol. 11, No. 5, 2019.
28. P. Sherubha, "An Adaptive FSCSO-RKPEM-based Feature Selection and Classification Techniques for Threat Identification in WSNs", Volume 53, ISSUE 2 (MAY - AUG), 2019 Caribbean Journal of Science, ISSN: 0008-6452.