A NOVEL INTRUSION DETECTION MODEL FOR MOBILE AD-HOC NETWORKS USING CP-KNN

M. Lalli[1] and V. Palanisamy[2]

[1]Department of Computer Science,Engg & Applications, Bharathidasan University, Trichy
[2]Department of Computer Science & Engineering, Alagappa University, Karaikudi

ABSTRACT

Mobile ad-hoc network security problems are the subject of in depth analysis. A group of mobile nodes area unit connected to a set wired backbone. In MANET, the node themselves implement the network management in a very cooperative fashion. All the nodes area unit accountable to create a constellation that is dynamically, modification it and conjointly the absence of any clear network boundaries. We tend to project a completely unique intrusion detection model for mobile ad-hoc network victimization. CP-KNN (Conformal Prediction K-Nearest Neighbor) algorithmic rule is to classify the audit knowledge for anomaly detection. The non-conformity score worth is employed to cut back the classification period of time for multi level iteration. It is effectively notice anomalies with high true positive rate, low false positive rate and high confidence that the progressive of assorted anomaly detection ways. Additionally it is interfered by "noisy" knowledge (unclean data), the projected technique is strong, effective and conjointly it retains its smart detection performance and to avoid the abnormal activity.

KEYWORDS

Intrusion Detection System, Security, Intrusions, Conformal Prediction, Anomaly Detection

1. INTRODUCTION

Intrusion detection system is a device, typically another separate computer, which monitor various legitimate accesses or the system abuse their privileges [1] was used to identify malicious or suspicious events. Mobile ad-hoc network is a self configuring infrastructure less network. It is free to move independently in any direction, and will therefore change its link to its other devices frequently. The network communication was become more complex but the potential malicious outsiders who have somehow passed the screens of security controls and access controls. Prevention is although necessary, but it is not a complete computer security control; detection during an incident copes with harm that cannot be prevented in advance. But the IDS are a sensor, like a smoke detector, that raises an alarm if specific things occur.

Intrusion detection system perform various functions such as monitoring users and system activity, managing audit trails and highlighting user violation of policy or normal activity , installing and operating traps to record information about intruders but no one IDS performs all of these functions. According to this issue, many IDS techniques are used to detect various malicious activities effectively in MANETs [2].

DOI : 10.5121/ijcnc.2014.6515

The two general types of intrusion detection systems are classified into signature based and anomaly based IDS. 1) Signature based detection: Signature based model perform simple pattern-matching and report situations that match a pattern corresponding to a known attack type.

2) Anomaly based detection: Anomaly based model was build to perform the acceptable behavior and flag exceptions are used to find out the abnormal activity. The real activity is compared against a known suspicious area. 3) Hybrid of both anomaly and misuse detection: This model is a combination of signature and anomaly based to effectively find out the malicious activity. IDS model needs some of the characteristics, fault tolerance, imperviousness to subversion, scalability, adaptability, minimal overhead, configurability, Denial of service.

Due to this lack of security controls in mobile ad-hoc networks, we must picket not only against normal attacks such as denial of service, but also against selfish attacks and other malicious attacks. Intrusion prevention can be used as a first line of defense, but this system was not sufficient to prevent it directly [3]. Intrusion detection system was used as a mechanism for representing promising security failures in the system. This is simple way to classify in order to decide whether some observed traffic data is ‘‘normal’’ or ‘‘abnormal’’. The classification objective is to minimize the probability of error and to diminish the time period to get precise classification rate.

We proposed a novel intrusion detection model for mobile ad-hoc network using CP-KNN (Conformal Prediction K-Nearest Neighbor) algorithm to classify the audit data for anomaly detection. We have to calculate the non-conformity score value which is used to reduce the classification time period for multi level iteration. It is effectively detect anomalies with high true positive rate and low false positive rate of various anomaly detection methods. The proposed method is robust, effective and also it retains its good detection performance after employ the feature selection to avoid anomalous activity.

The rest of the paper is organized in the following sections. In section 2, we briefly describe about the various existing techniques used for IDS. In section 3, we describe our proposed work for IDS. The experimental result was described in section 4. In section 5, we summarize the paper.

2. RELATED WORK

Today, Intrusion detection is a mature field in mobile ad-hoc network security. Many papers focus particularly on systems based classification algorithms. The amount of work to be reported for classification-based intrusion detection in mobile ad-hoc networks is very less, but it is extensively used for wired networks.

Zhang and Lee [4] proposed the first (high-level) Intrusion Detection System (IDS) approach specific for mobile ad hoc networks. A distributed and cooperative anomaly-based IDS provides an efficient design of IDS in wireless ad hoc networks. Anomaly detection approach was based up on various routing updates on MAC layer and mobile application layer.

Huang and Lee [5] discussed about cluster based IDS which utilize a set of statistical features which was get by the routing tables and it was classified by decision tree induction algorithm C 4.5 to detect the behavior as “normal” versus “abnormal”. Abdel-Fattah et al. [6] proposed the Conformal Predictor k-nearest neighbor and the Distance based an Outlier Detection (CPDOD) algorithm which was used to detect various types of malicious activities in mobile ad-hoc networks.

Deng et al. [7] proposed two distributed intrusion detection approaches, based on hierarchical and distributed architecture respectively. The intrusion detection approach used a Support Vector Machine (SVM) classification algorithm. Using network layer, so many sets of parameters are used in distributed intrusion detection approach and that will be suggested in hierarchical distributed approach for prominent solutions.

Liu et al. [8] proposed a completely distributed anomaly detection approach. They used MAC layer data to profile the behavior of mobile nodes and then applied cross-feature analysis [9] on feature vectors constructed from the training data. The cooperative and distributed IDS utilize the various data's from MAC layer, routing values in application layers and also which was coupled with a Bayesian classifier was proposed by Bose et al [10]. Cabrera et al. [11] proposed C 4.5 training multiple classifiers, which was used to evaluate them for two types of attacks.

The true positive (TP) and false positive (FP) detection rate was demonstrated by using K-Nearest Neighbor algorithm and one-class SVM, which was proposed by authors in [12] for unsupervised anomaly detection. Especially, the performance of one class SVM algorithm was compared to the traditional supervised anomaly detection methods. But the TP and FP results are not accuracy.(98% for TP and 10% for FP [8]).

Yang Li [13] proposed TCM-KNN (Transductive Confidence Machines for K-Nearest Neighbors) machine learning algorithm which has been successfully applied to pattern recognition, fault detection and outlier detection.

The accurate value of KNN was ruined by the presence of noisy data, unrelated features and the feature scales. It is not a consistent one. More research work has been put into selecting or scaling the features to improve classification. In this paper, our proposed algorithm was used to detect anomalies with high true positive rate and low false positive rate effectively.

3. THE PROPOSED WORK

3.1 Conformal Prediction for K-Nearest Neighbor (CP-KNN)

The Conformal Prediction for k-nearest neighbor (CP-KNN) algorithm was used to calculate the resemblance between new individuals and other samples in the class using the K-nearest neighbor method. We have find out the non conformity score values for each sample and also these values are applied to find the transductive confidence. To estimate that the new samples are belongs to this particular class with p-values. The objective is, nonconformity score value is corresponds to the uncertainty of the point which was measured with respect to all other classified samples of a class with higher uncertainty, and higher nonconformity scores values. Using the Euclidean distances between points the CP-KNN nonconformity score is computed.

- The CP-KNN nonconformity score was deliberate using the Euclidean distances between points of the network parameters.
- Let, D_i^y as the sort sequence of the Euclidean distances of the point *i* from other points with the same classification *y*. The distance between *i* and the j^{th} shortest samples in the sequence is D_{ij}^y
- Similarly, D_i^{-y} is the distance between the sample *i* from the other sample with different classification, then D_{ij}^{-y} as the remoteness between *i* and the j^{th} shortest samples in the same sequence.
- α is an individual nonconformity score value which was assign to every sample. The nonconformity score value of the sample *i* with classification *y* is α_{ij} .

$$\alpha_{ij} = \frac{\sum_{i=1}^{k} D_{ij}^{y}}{\sum_{i=1}^{k} D_{ij}^{-y}} \qquad (1)$$

Therefore, the quantify value of nonconformity is the ratio of the sum of the k nearest distances from the same class (y) to the sum of the k nearest distances from all other classes (-y). There are many classes in the feature space. According to the classes, the nonconformity score for the fitness of the query sample is based to the class y with respect to all others classes in the features space. Finally, the nonconformity score of a sample was increased. When the sum of the k nearest distances from the points of the related class was higher and also the sum of the k nearest distances from the other classes was smaller.

In intrusion detection, the nonconformity score was used to measure the peculiarity of an activity i belongs to the normal class y with respect to the anomalous class -y. CP-KNN algorithm used to computes the nonconformity score of m training samples in class y and arrange their nonconformity score values in sliding order. Based on the Equation (1), the algorithm can also be calculating the nonconformity score of the latest query sample v if it is classified as normal class y. Then, the p-value of the query point was compute using Equation (2), where the nonconformity score of the new unknown sample was defined as v.

$$P(\alpha_v) = \frac{\#\{i=(1,\ldots..,m): \alpha_i \geq \alpha_v\}}{m+1} \qquad (2)$$

The entire training points are independent random samples. The strength of the indication against v was belongs to class y, where i is the number of class members with larger nonconformity score values. If the value is larger than the p-value then it shows how likely the query point is to be classified as y, by referring to the distribution of all points in the same class. The smaller the p-value the more improbable query point was belongs to class y.

3.2 Conformal Prediction for K-Nearest Neighbor (CP-KNN) Algorithm

In this paper we proposed CP-KNN, which is used to classify the audit data for anomaly detection. Conformal prediction is to determine the precise levels of confidence in new predictions. The error probability is defined as ϵ. The error probability and the method both are combine together and make a point prediction of a label y. It produces a set of labels that typically containing the point prediction is also contains y with probability$1-\epsilon$. Conformal prediction is a method for producing point predictions such as nearest neighbor method. In this method K depends upon the data; a larger values of *k* will reduce the effect of a noise on the classification, but it will make the boundaries between classes with less distinct. The class which was predicted to be the nearby class of the training sample (*k* = 1) is a nearest neighbor. The CP-KNN classification process is shown in Figure 1.

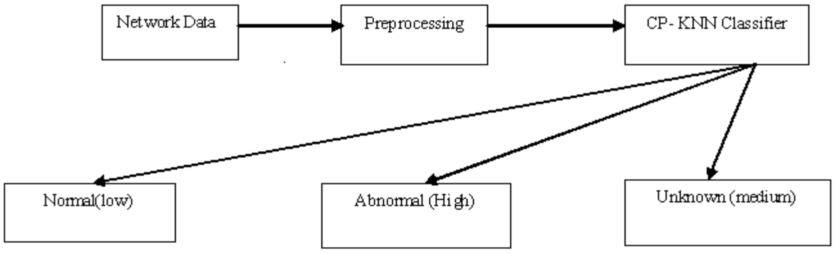


Figure 1. CP-KNN classification

The classification process consists of three stages such as network data collection, preprocessing and CP-KNN classification. The data collection stage provides the data for various network activities and also the number of features needs to be selected to represent the ad-hoc network activity which can be used to detect various attacks. The initial subset features selection is part of the preprocessing, which is implemented by feature selection algorithm. Different feature selection corresponds to different type of attacks. Depends upon the CP-KNN classification rate we have to detect ‘‘normal’’ versus ‘‘abnormal’’ behaviors. This will give higher performance when we removed the irrelevant features. The overview of CP-KNN design architecture is shown in Figure 2.

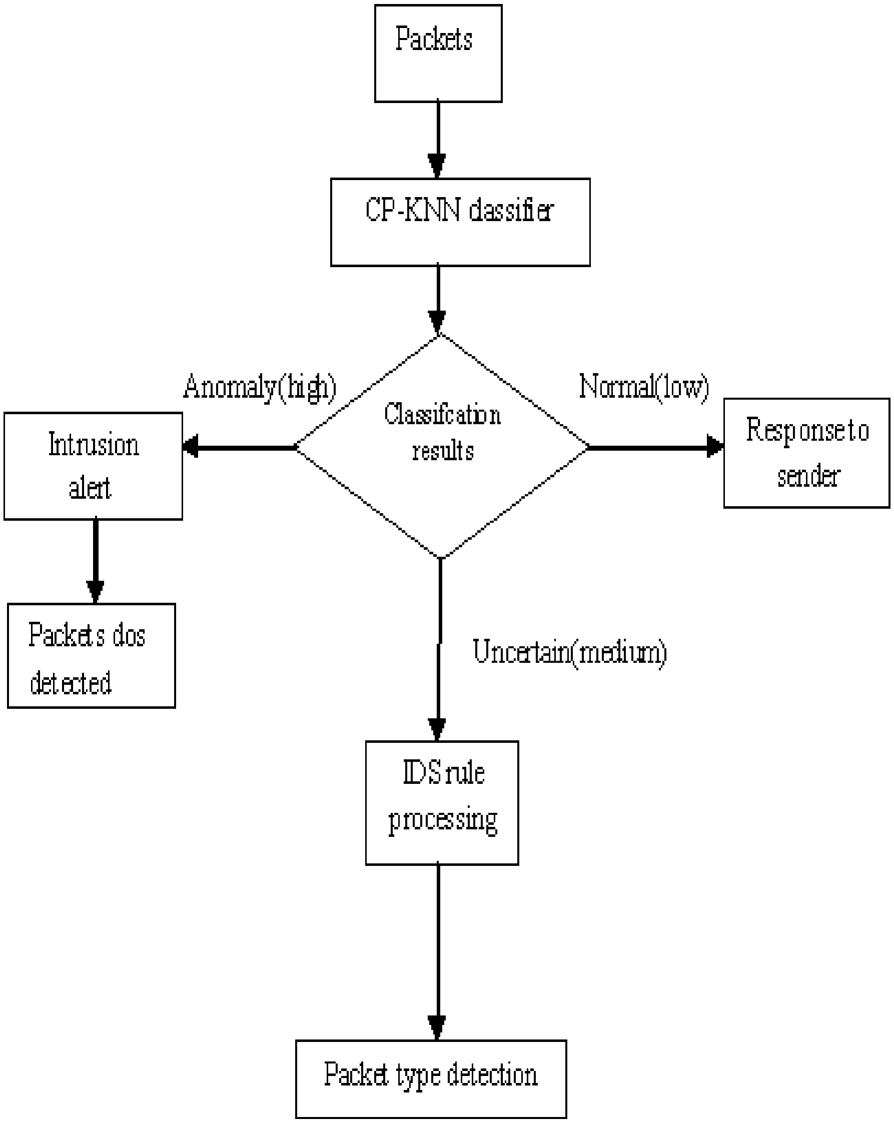
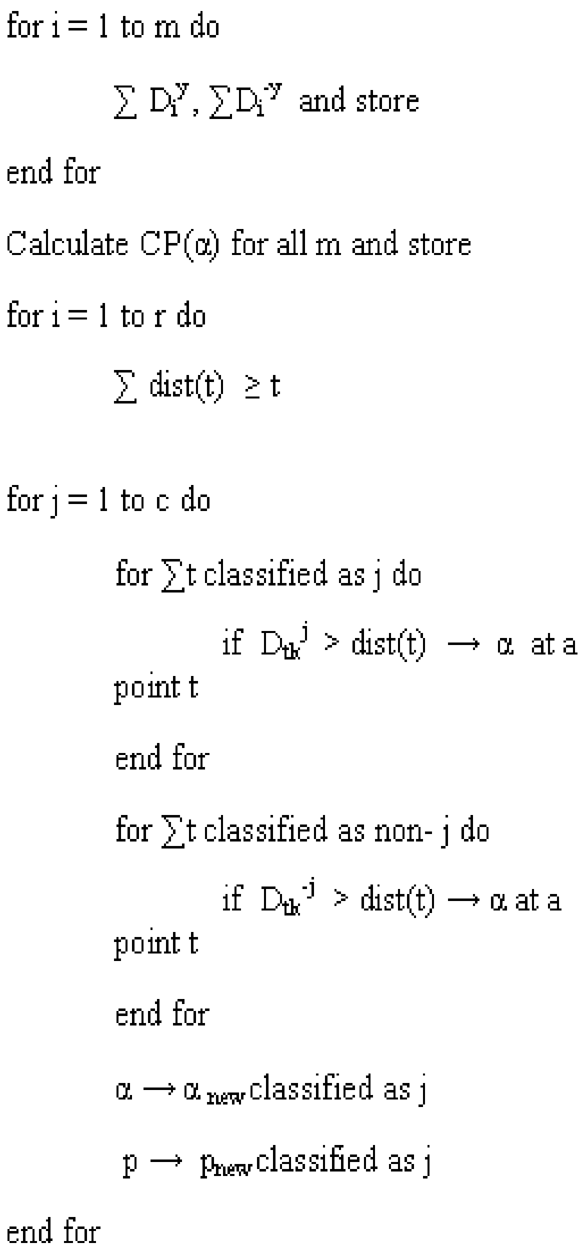


Figure 2. Design Architecture of CP-KNN

The CP-KNN algorithm is shown in Figure 3.



```
for i = 1 to m do
        $\sum D_i^y, \sum D_i^{-y}$ and store
end for
Calculate CP($\alpha$) for all m and store
for i = 1 to r do
        $\sum$ dist(t) $\geq$ t

for j = 1 to c do
        for $\sum$t classified as j do
                if $D_{tk}^j$ > dist(t) $\rightarrow$ $\alpha$ at a
        point t
        end for
        for $\sum$t classified as non- j do
                if $D_{tk}^{-j}$ > dist(t) $\rightarrow$ $\alpha$ at a
        point t
        end for
        $\alpha \rightarrow \alpha_{new}$ classified as j
        p $\rightarrow$ $p_{new}$ classified as j
end for
```

Figure 3. Conformal Prediction K-Nearest Neighbor Algorithm

4. EXPERIMENTAL RESULTS

These works based to find out the different type of attacks that are correspond to different types of feature selection. Distributed and dynamic nature of the local features is composed in an ad-hoc environment. The features should be within the node itself or its communication activities by overhead transmissions to and from the nodes. The proposed CP-KNN is reducing the classification time period for multi level iteration. The dissimilar features are removed to get the higher performance. It is interfered by "noisy" data (unclean data), the proposed method is robust, effective and also it retains its good detection performance after employ the feature selection to avoid anomalous activity. The Classifier Rate(P) is calculated as,

Classifier Rate (P) =(Number of data classified / Total Number of Data)* Iteration Level

The Iteration level is defined as the classifier classified the data with number of various levels.

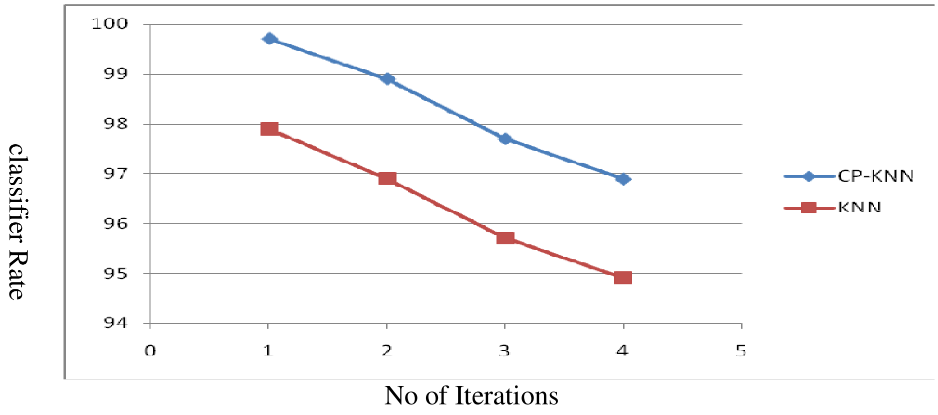


Figure 4. Classifier Rate for CP-KNN vs KNN

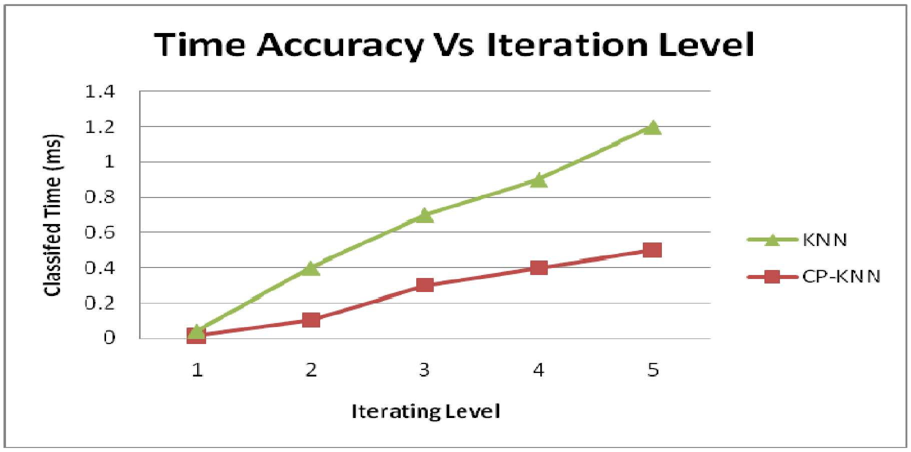


Figure 5. Time Accuracy Graph for CP-KNN and KNN

In Figure 5. the graph is based on the classifier that makes the time consumption for its classification at multi level iteration. In that experiment, the accuracy for K-Nearest Neighbor is 90% and Conformal Prediction K-Nearest Neighbor accuracy is 94%.

5. CONCLUSION

Due to this vulnerability of ad hoc networks, the intrusion anticipation measures such as encryption and authentication. They are used to reduce the various types of intrusions, however they cannot be totally eliminate them. For this reason, researchers necessitate intrusion detection; it will be act as a frontline of security in mobile ad- hoc networks. In this paper we proposed a novel intrusion detection model for mobile ad-hoc network using CP-KNN (Conformal Prediction K-Nearest Neighbor) algorithm to classify the audit data for anomaly detection. The non-conformity score value is used to reduce the classification time period for multi level iteration. The proposed work was effectively detecting various anomalies with high true positive rate, low false positive rate and high confidence rate. In addition it is interfered by "noisy" data (unclean data), the proposed method is robust, effective and also it retains its good detection performance to avoid anomalous activity.

REFERENCES

[1] Abarna Sri.R, Lalli. M , "NIDS in Manet Using KMCA" International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 9, 2013

[2] Farhan Abdel-Fattah, Zulkhairi Md. Dahalin, Shaidah Jusoh, "Distributed and Cooperative Hierarchical Intrusion Detection on MANETs", International Journal of Computer Applications, Volume 12-No.5, December 2010

[3] Aikaterini Mitrokotsa, Christos Dimitrakakis, "Intrusion Detection in MANET using Classification algorithms: The effects of cost and model selection", Elsevier- Ad Hoc Networks 2012.

[4] Y. Zhang, W. Lee, Y. Huang, Intrusion detection techniques for mobile wireless networks, Wireless Networks 9 (5) (2003) 545–556.

[5] Y. Huang, W. Lee, A cooperative Intrusion detection system for ad hoc networks, in: Proceedings of the 1st ACM Workshop on Security of Ad Hoc and Sensor Networks (SASN'03), Fairfax, VA, USA, 2003, pp. 135–147

[6] F. Abdel-Fattah, Z. Md. Dahalin, S. Jusoh, Dynamic intrusion detection method for mobile ad hoc networks using CPDOD algorithm, International Journal of Computer Applications, vol. 2, Published by the Foundation of Computer Science, 2010. pp. 22–29.

[7] H. Deng, Q. Zeng, D.P. Agrawal, SVM-based intrusion detection system for wireless ad hoc networks, in: Proceedings of the 58th IEEE Vehicular Technology Conference (VTC'03), vol. 3, Orlando, FL, USA, 6–9 October 2003, pp. 2147–2151.

[8] Y. Liu, Y. Li, H. Man, Mac layer anomaly detection in ad hoc networks, in: Proceedings of the 6th Annual IEEE SMC Information Assurance Workshop (IAW '05), West Point, NY, USA, 15–17 June 2005, pp. 402–409.

[9] Y. Huang, W. Fan, W. Lee, P. Yu, Cross-feature analysis for detecting ad-hoc routing anomalies, in: Proceedings of the 23rd International Conference on Distributed Computing Systems, Rhode Island, USA, 2003, p. 478.

[10] S. Bose, S. Bharathimurugan, A. Kannan, Multi-layer intergraded anomaly intrusion detection for mobile ad hoc networks, in: Proceedings of the IEEE International Conference on Signal Processing, Communications and Networking (ICSCN 2007), February 2007, pp. 360–365.

[11] J.B.D. Cabrera, C. Gutiérrez, R.K. Mehra, Ensemble methods for anomaly detection and distributed intrusion detection in mobile adhoc networks, In Information Fusion 9 (2008) 96–119.

[12] D. Djenouri, O. Mahmoudi, M. Bouamama, D.L. Llewellyn-Jones, M. Merabti, On securing manet routing protocol against control packet dropping, in: Proceedings of the IEEE International Conference on Pervasive Services (ICPS '07), Istanbul, Turkey, 15-20 July 2007, pp. 100–108.

[13] Yang Li, Binxing Fang, Li Guo,You Chen, "Network Anomaly Detection Based on TCM-KNN Algorithm", ASIACCS'07, March 20–22, 2007, Singapore.

Authors

Author 1 : Lalli. M is working as a Assistant Professor in the Department of Computer Science Engineering and Applications, Bharathidasan University, Trichy, TamilNadu, India. She has 12 Years of experience in teaching. Her area of interest is Manets. Other areas of on interest include Information Security and Computer Networks and she is pursuing Ph.D in mobile Ad-hoc network securityunder the guidance of Dr. V. Palanisamy.

Author 2. Palanisamy.V is working as a Head of the Department of Computer Science & Engg Alagappa university. Karaikudi, Tamil Nadu, India. He has 20 Years of Teaching and 15 Years of experience in research. His area of Specialization is Wireless networks and network security. Other areas of an interest is include Algorithms.