Dynamic Intrusion Detection Method for Mobile Ad Hoc Network Using CPDOD Algorithm

Farhan Abdel-Fattah
Graduate Dept. of Computer Science
College of Arts and Sciences
Universiti Utara Malaysia
Sintok 06010 Kedah, Malaysia

Zulkhairi Md. Dahalin
Graduate Dept. of Computer Science
College of Arts and Sciences
Universiti Utara Malaysia
Sintok 06010 Kedah, Malaysia

Shaidah Jusoh
Graduate Dept. of Computer Science
College of Arts and Sciences
Universiti Utara Malaysia
Sintok 06010 Kedah, Malaysia

ABSTRACT
Abstract Mobile Ad hoc networks (MANETs) are susceptible to several types of attacks due to their open medium, lack of centralized monitoring and management point, dynamic topology and other features. Many of the intrusion detection techniques developed on wired networks cannot be directly applied to MANET due to special characteristics of the networks. However, all such intrusion detection techniques suffer from performance penalties and high false alarm rates. In this paper, we propose a novel intrusion detection method by combining two anomaly methods Conformal Predictor k-nearest neighbor and Distance-based Outlier Detection (CPDOD) algorithm. A series of experimental results demonstrate that the proposed method can effectively detect anomalies with low false positive rate, high detection rate and achieve higher detection accuracy.

General Terms
Security, Intrusion detection, mobile Ad hoc Network, k-Nearest Neighbors.

Keywords
MANET Intrusion detection, CPDOD, CP-KNN, Dynamic intrusion detection, Conformal Prediction.

1. INTRODUCTION
The mobile Ad hoc Network (MANET) consists of nodes which are built up from mobile devices such as mobile computers, Personal Digital Assistant (PDA) and wireless phones. The nodes communicate with each other using wireless links and forming a temporary network without the aid of an established infrastructure or a centralized administration. The absence of a centralized administration and node mobility makes all MANETs' nodes behave as both hosts and routers. In general, the cooperation of all nodes in MANET ensures reliable routing services. On the other hand, dependency and decentralized of MANET allows an adversary to exploit new type of attacks that are designed to destroy the cooperative algorithms used in ad hoc networks. Moreover due to their open medium, dynamically changing network topology and lacking central monitoring and absence of a clear line of defense, MANET is particularly vulnerable to several types of attacks like passive eavesdropping, active impersonation, and denial of services. An intruder that compromises a mobile node in MANET can destroy the communication between the nodes by broadcasting false routing information, providing incorrect link state information, and overflowing other nodes with unnecessary routing traffic information [15]. Therefore, successful implementation of MANET depends on users' confidence in its security. The security research in MANET has focused on key management, routing protocol and intrusion detection techniques [4], but past experiment have shown that encryption and authentication as intrusion prevention are not sufficient, and gravely of the security problem depend on the complexity of the system [23]. At present, completely preventing breaches of security seems unrealistic, especially in cellular Internet, MANET and wireless network. On the other hand, intrusion detection techniques used in wired networks cannot be applied directly to MANEs due to special characteristics of the networks [3, 8, 23].

Intrusion detection system (IDS) plays a very important role for detecting different types of attacks. The main function of intrusion detection is to protect the network, analyze and find out intrusions among normal audit data, and this can be considered as a classification problem. Intrusion detection system can be classified based on detection method into two basic methods misuse detection and anomaly detection methods. The misuse detection method, also known as signature-based, operate on a database of known attack signatures; the system stores patterns (or signatures) of known attacks and uses them to compare with the actual activity. Another approach to intrusion detection is called anomaly-based intrusion detection. Anomaly detection works on the assumption that "attack behavior" differs and distinct a sufficient amount from "normal user behavior". Anomaly detection algorithms have the advantage over a signature-based detection that they can detect novel attacks. Although Anomaly detections methods are able to detect new types of intrusions, most of these anomaly-based IDSs suffer from a high rate of false alarms due to a deficiency in their discrimination ability [11, 16, 21].

In this paper, we propose a novel intrusion detection method based on the combination of two anomaly methods namly Conformal Predictor k-nearest neighbor and Distance-based Outlier Detection (CPDOD) algorithm. Our algorithm employs two different metrics to improve detection ability. The nonconformity metric measures whether the unknown instance is more similar to known normal instances or abnormal instances. The Outlier Factor LDOF metric identifies the similarity to normal classes and detect abnormal attacks. Our algorithm is commonly used machine learning and data mining technique. To the our best knowledge, it is the first time that Conformal Predictor K-Nearest Neighbor (CP-KNN) and Distance-based Outlier Detection (DOD) algorithms are applied to ad hoc network intrusion detection introduced by us.

This paper is organized as follows. In Section 2, provides background knowledge of the CP-KNN and DOD. Section 3 presents the intrusion detection algorithm. Section 4 illustrates the experiments and presents the results with some discussion. Finally, we summarize our work in Section 5.

2. BACKGROUND

The existing machine learning methods such as Support Vector Machines (SVM), Genetic Algorithms (GA) and Neural Networks (NN), fall under the category of inductive machine learning approach. On the other hand, there are fewer transductive algorithms, the most commonly used are the k-Nearest Neighbors (k-NN) algorithm and radial basis function networks [5, 13]. The k-NN classifier classifies an example by means of majority vote among the labels of the k nearest neighbours.

The traditional inductive machine learning approach makes two separate steps in learning process. The first step, processing the set of training examples to find general rules (decision function), and the second step using these rules to make predictions on a new example. On the contrary the transductive machine learning approach merges the two inductive approach steps into one single step by delays the process of finding general rules until a new test example is presented. Then, classify the new example points by analyzing the likelihood they belong to predefined classes and make a decision based on only a part of the training data [5, 13, 20]. The main advantage of the transductive machine learning approach is that instead of generating the decision function for the entire input space, the decision function can be generated locally and differently for each new example to be classified. This is a major advantage when the decision function to be modeled is very complex [5, 6, 18]. Neural Networks and Support Vector Machines as machine learning techniques have been concerned in making 'bare predictions', without any measure of confidence on the resulting decision [6]. These learning algorithms have been attributed to many successes in applications of pattern classification problems in recent years. Unlike traditional techniques in machine learning, transductive approach can present measures of reliability to individual points [11]. However, in MANET where the network topology dynamically changes it is difficult to draw general rules for all network activities at the same time. Therefore, transductive machine learning approach is suitable for a MANET intrusion detection system.

2.1 Conformal Predictor K-Nearest Neighbor (CP-KNN)

Gammerman et al. [5] use Transduction to present confidence measures for the decision of classifying an example point as belonging to a set of pre-defined classes. The recently introduced Conformal Predictor (CP) [6, 18] uses past experience to determine precise levels of confidence in predictions. CP introduced the computation of the confidence using Algorithmic Randomness Theory. Transaction confidence machine is a prediction technique compute a p-value for the new example *v* of any predefined class *c*. The definition of p-value is the probability of observing an example in the sample space that can be considered more extreme than a sample of data. The p-value measure how well the data (examples of a class) supports a null hypothesis that the query point belongs to a certain class. The smaller value of the p-value, the greater is the evidence against the null hypothesis.

The Conformal Prediction for k-nearest neighbor (CP-KNN) algorithm computes the similarity between new individual and other examples in the class using the K-nearest neighbor distances method. The important step when applying transductive confidence is to calculate a nonconformity score value for each example. And estimates how likely it is that a new example belongs to this class with p-values. The main idea is that the nonconformity score corresponds to the uncertainty of the point being measured with respect to all the other classified examples of a class: the higher the nonconformity score, the higher the uncertainty [10].

The CP-KNN nonconformity score is calculated using the Euclidean distances between points. Let us define $D_i^{\ y}$ as the sorted sequence of the Euclidean distances of point *i* from other points with the same classification *y*. The distance between *i* and the *jth* shortest examples in the sequence is $D_{ji}^{\ y}$ similarly let D_i^{-y} define the distances of example *i* from the other example with different classification, then D_{ji}^{-y} as the distance between *i* and the *jth* shortest examples in the sequence. α is an individual nonconformity score assign to every example. The nonconformity score for example *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, this measure 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). When there are several classes in the feature space, nonconformity score the fitness of the query example to class y with respect to all others classes in the features space. The nonconformity score of a example raises when the sum of the k nearest distances from the points of the same class becomes bigger or when the sum of the k nearest distances from the other classes becomes smaller.

Nonconformity score can be used in intrusion detection to measure the strangeness of an activity *i* belonging to the normal class *y* with respect to the abnormal class -*y*. The CP-KNN algorithm computes the nonconformity score of *m* training examples in class *y* and sorts their nonconformity score values in descending order $\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$. Based on Equation (1), the algorithm can also calculate the nonconformity score of the new query example *v* if it is classified as normal class *y*. Then, the p-value of the query point can be computed using Equation (2), where α_v is the nonconformity score of the new unknown example *v*.

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

As all training points are independent random samples, the strength of the evidence against *v* belonging to the class *y* is quantified by $p(\alpha_v)$, where *i* is the number of class members with nonconformity score larger than α_v. The *p*-value 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 unlikely the query point belongs to class *y*.

2.2 Distance-based Outlier Detection (DOD)

Recently, Zhang et al. [22] proposed Local Distance-based Outlier Factor (LDOF) to measure the outlier-ness of a point in the feature space. LDOF uses the relative location of a point to its neighbors to determining whether a point is an outlier with respect to all clusters. LDOF is the distance ratio representing and indicating how far the point x lies outside its neighborhood system.

Formal definition of the Local Distance-based Outlier Factor

Definition 1 (KNN distance of x_p) *Let N_p be the set of the k-nearest neighbours of object x_p (excluding x). The k-nearest neighbours distance of x_p equals the average distance from x_p to all objects in N_p. More formally, let dist(x , x') $\geq$ 0 be a distance measure between objects x and x'. The k-nearest neighbours distance of object x_p is defined as:*

$$\overline{d_{x_p}} = \frac{1}{k}\sum_{x_i \in N_p}^{k} dist(x_i, x_p).$$

Definition 2 (KNN inner distance of x_p) *Given the k-nearest neighbours set N_p of object x_p, the k-nearest neighbours inner distance of x_p is defined as the average distance among objects in N_p:*

$$\overline{D_{x_p}} = \frac{1}{k(k-1)}\sum_{x_i, x_{i'} \in N_p, i \neq i'} dist(x_i, x_{i'})$$

Definition 3 (LDOF of x_p) *The local distance-based outlier factor of x_p is defined as:*

$$LDOF_k(x_p) := \frac{\overline{d_{x_p}}}{\overline{D_{x_p}}}$$

When the Outlier Factor LDOF $\leq$1, it means that new example x_p is inside the class and surrounded by a class data. In contrast, when Outlier Factor LDOF $\geq$ 1, it means that new example x_p is outside the whole class. We use Outlier Factor LDOF to distinguish between normal and abnormal examples. It is easy to see that in any datasets, an example is outlier if Outlier Factor LDOF > 1.

Algorithm 1 CPDOD algorithm

Input: the training set T={(x_1, y_1),. . ., (x_{m-1}, y_{m-1})} and a new unlabeled example x_m, Confidence threshold τ_1,Outlier threshold τ_2

Output: The set of p-values when T is a two classes dataset normal (n) and abnormal (a)

For i = 1 to m-1 Do

calculate $D_i^{\,y}$ ***and*** D_i^{-y}

End For

Compute nonconformity scores α_n ***and*** α_a ***for all training points using Equation 1 and store***

Compute nonconformity scores α_n ***and*** α_a ***for new example x***

Compute p-values for the new example x_m, p_n ***and*** p_a ***using Equation 2***

predict the class with the largest p-value

Confidence = 1 - second highest p-value

IF Confidence $\geq \tau_1$

Return x_m ***class***

Else

Compute the average distance from x_m ***to all k nearest neighbors***

Compute the average distance among points in k nearest neighbors

Calculate the Outlier Factor LDOF for example x_m

IF (LDOF $\geq \tau_2$ ***) Then***

Classify x_m ***abnormal,***

Else

Classify x_m ***normal,***

End IF

End IF

Return x_m ***class***

3. THE CPDOD ALGORITHM

Anomaly detection systems have the capability to detect unknown attacks but at the same time they suffer from high false alarm rate when normal user profiles and system or network behavior vary widely. Anomaly detection can be combined with signature verification to detect attacks more efficiently [1, 12, 16]. Our detection method employ the combined of two detection methods Anomaly detection with signature detection to detect attacks more efficiently. At the same time, our model uses two anomaly methods CP-KNN and DOD in a conditional sequence structure that is shown in Figure 1.

We introduce both measures (CP-KNN nonconformity score and Outlier Factor LDOF) to MANET intrusion detection and use them together as follows. The CP-KNN algorithm computes the nonconformity score of the query point with respect to all classes and gets a sequence of p-values. Since the smaller the p-value, the more evidence against the null hypothesis that the query point belongs to the class, the algorithm predicts that the point belongs to the class with the largest p-value. The confidence of this prediction, which is how unlikely the prediction is wrong, is equal to the complement of the second largest p-value. The nonconformity score of the CP-KNN algorithm estimates the fitness of a new example to one class with respect to all other classes. The Outlier Factor LDOF measures absolute deviation from the class of interest. Thus, any activity that is significantly different from the normal data in nonconformity score measure is regarded as an intrusive. Outlier Factor LDOF is used to measure the deviation of an activity from the normal data.

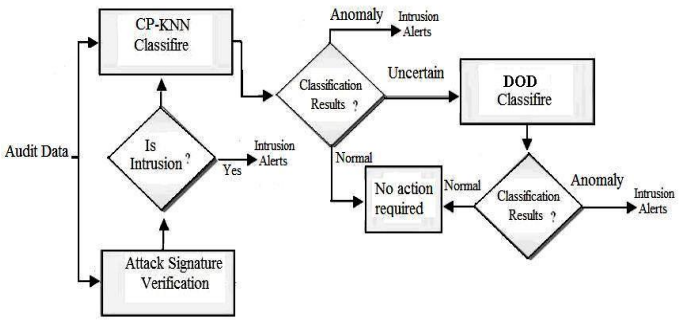

Figure 1: Structure of CPDOD algorithm

This proposed classifier also grades the confidence of the decisions based on the results of the two classifiers and their individual confidence metrics, and offers insights into the nature of the ad hoc network behaviors being tested. Our proposed detection framework measures an attack or vulnerability with a metric known as Risk Index (RI). RI is the metric used by the detection framework to discover if the mobile node is under attack or not. Risk Index is a digits number, which takes values between number 1 and number 10. In this research work, the network is classified into three states include normal state, uncertain state and vulnerable state. The network is in the normal state when there is no attack. This is indicated by the RI range from 1 to 3 and its confidence measure for normal classification range from 90% to 100% ; the network is in the vulnerable state when there is an intrusion and is indicated by the RI range from 8 to 10 with it's confidence measure for anomaly classification range from 90% to 100%. The intermediate state of the network between normal and vulnerable state is referred to as the uncertain state which is confidence measure less than 90% for one of both classes normal or abnormal. At the first, our detection algorithm use the CP-KNN as a main classifier to analyses the collected data. At this step if the data is sufficient to find the class of the activity (depending on the two labeled classes) the This proposed classifier also grades the confidence of the decisions based on the results of the two classifiers and their individual confidence metrics, and offers insights into the nature of the ad hoc network behaviors being tested. Our proposed detection framework measures an attack or vulnerability with a metric known as Risk Index (RI). RI is the metric used by the detection framework to discover if the mobile node is under attack or not. Risk Index is a digits number, which takes values between number 1 and number 10. In this research work, the network is classified into three states include normal state, uncertain state and vulnerable state. The network is in the normal state when there is no attack. This is indicated by the RI range from 1 to 3 and its confidence measure for normal classification range from 90% to 100% ; the network is in the vulnerable state when there is an intrusion and is indicated by the RI range from 8 to 10 with it's confidence measure for anomaly classification range from 90% to 100%. The intermediate state of the network between normal and vulnerable state is referred to as the uncertain state which is confidence measure less than 90% for one of both classes normal or abnormal. At the first, our detection algorithm use the CP-KNN as a main classifier to analyses the collected data. At this step if the data is sufficient to find the class of the activity (depending on the two labeled classes) the system initiates an intrusion alarm if the classification result is anomaly. On the other hand if the result is normal the system does not do any actions. But if the CP-KNN find the class with low confidence measure less than 90%, the system go to the next step by using DOD algorithm, which is use normal data only to make the decision. System initiates an intrusion alarm if the classification result is anomaly. On the other hand if the result is normal the system does not do any actions. But if the CP-KNN find the class with low confidence measure less than 90%, the system go to the next step by using DOD algorithm, which is use normal data only to make the decision.

4. Experiments and Evaluation

4.1 Simulation Environment

Simulators are the most common tools used for testing MANET intrusion detection systems [9, 15]. Simulators help researchers to study the performance and the reliability of their proposed IDS without using real mobile nodes. In order to evaluate our techniqe we simulated MANET by using Global Mobile information systems Simulation library (GloMoSim) [7]. It builds a scalable simulation environment for wireless and wired network systems. Parsec, is a C-based simulation language based on parallel discrete-event simulation, is used to design GIoMoSim. We have taken Ad hoc On Demand Distance Vector (AODV) [17], one of the popular MANET routing algorithms [14], as a network routing protocol. Specifically, in the simulation, nodes having the same transmission range of 200

meters with the channel capacity of 2000 bps. We use the Distributed Coordination Function (DCF) of IEEE 802.11 for wireless LANs as the MAC layer protocol. In this simulation, 30 mobile nodes were set to move in the area of 1000 meters x 1000 meters. To simulate the nodes mobility, a Random Waypoint model (RW) is used. All nodes were set to move independently with the same average speed.

Table 1: Description of the datasets

Data set	Attack	Examples	sampling period
BHAT10	Black Hole	3000	10 seconds
BHAT30	Black Hole	1000	30 seconds
BHAT60	Black Hole	500	one minute
RCAT10	Resource Consumption	3000	10 seconds
RCAT30	Resource Consumption	1000	30 seconds
RCAT60	Resource Consumption	500	one minute
DRAT10	Dropping Routing Traffic	3000	10 seconds
DRAT30	Dropping Routing Traffic	1000	30 seconds
DRAT60	Dropping Routing Traffic	500	one minute
BRDAT	Three attacks	3000	10 seconds

4.2 Attacks

Wireless ad hoc network routing protocols are designed based on the concept that all the nodes must participate in the routing process. These protocols assume a trusted and cooperative network environment. Many researchers [2, 3, 23] discussed various type of attacks can be performed easily against the ad hoc network routing protocols. We choose to implement three common attacks to evaluate the performance of our Dynamic Intrusion Detection algorithm.

• Black Hole Attack: In MANET that use reactive protocols such as AODV, protocols create and maintain routes between nodes by assigning special increasing sequence numbers to find a path to a destination node. Because the existing of a route is determined from the destination's sequence number, an attacker or a compromised node can transmit and inject fake routing information to the network. Thus act as to have the fresh enough route information to the destination node. Sequence Number works as a time stamp and let nodes to determine how new and fresh their information on the other node is. However when a node transmits any type of routing control packets like Route_Request packet, Route_Reply packet, Route_Error packet, it increases its own sequence number. Higher sequence number gives evidence of more new and correct information. And mobile node that has highest sequence number, its information is considered and route to the destination is established based on this trusted information. In this scenario, The attacker node can make the black hole attack by inserting itself into the active route, and send Route_Reply packet with highest destination sequence number (advertises itself as having a suitable, direct path to the destination node), even if it does not have any route. The malicious node then just drops all the receiving packets without any forwarding and creates a black hole in the ad hoc network, as the attack name implies.

• Resource Consumption Attack: In this type of attack the malicious or compromised node attempt to consume both the network and node resources by broadcasting and sending frequent excessive routing control packets. This routing traffic can be Route_Request packet or Route Reply packet. The destinations node addresses that are used in the Resource Consumption Attack do not exist in the ad hoc network. In order for these packets not to be removed by the protocol implementation rule, the attacker change the destination node address each time. The goal of this malicious attack is to overflow the network with fake routing control packets to consume all the available network bandwidth with unrelated traffic and to consume power from the mobile nodes. At the same time, effectively paralyze the ad hoc network.

• Dropping Routing Traffic Attack: To conserve the battery life and nodes resources, a mobile node may decide not to participate and cooperate in the routing process. Mobile node acts selfishly by dropping all routing packets that is not send to it. At the same time it processes its routing packets. The node may conserve its energy and resources by acting selfishly but it may also cause many networks problems such as segmentation. If any mobile node is only connected with other nodes through malicious node then it becomes isolated and unreachable.

4.3 Experimental Data Set

In this work, 10 source nodes and 10 destination nodes are selected randomly to generate Constant Bit Rate (CBR) traffic as the background traffic. The transmission rat is 2 packets per second with the packet size 512 byte. In our simulation, in order to give the nodes enough time to finish the network initialization process, we collect the traffic data after a warm up time of 400 seconds. In our experiments, the data is sampled in three sampling periods, 10 seconds, 30 seconds, and 60 seconds. Details on the data sets are shown in table 1. Selecting the correct set of features is an important step when formulating the classification tasks. We mainly consider the features have been commonly used in the MANET intrusion detection research [3, 8]. The feature will select as a sensitive feature based on the confidence measure during CPDOD algorithm training phase, when some feature give high confidence for normal prediction or anomaly prediction. Therefore, the features will depend on the region and type of attack.

• Routing packet propagation features: The routing packets (Route_Request , Route_Reply, Route_Error and Hello messages) can take four directions, sent packet by a source node, forwarded packet by intermediate node, receive packet by the destination and dropped packet by the node who does not has active route for the packet. In total, there are sixteen features for Routing packet propagation.

• Route table changes: Routing table is an electronic database or file. This table stores all ad hoc network activities, such as, add route, delete route, found route and stale route. Changes in the routing table can capture the basic view of ad hoc network topology update, and the relationship between the mobile nodes.

• Data packet transmission: We consider data packets in two layers, network layer and transport layer. At the network layer the data packets can be in the four directions (sent, forward, received or dropped). While at the transport layer the data

packets can be sent by source node or received by destination node.

4.4 Performance of Dynamic Intrusion Detection Model

One of the most important problems facing MANET intrusion detection is the high false alarms rate or False Positive Rate (FPR), generating during intrusion detection process [15, 19]. The researcher in MANET intrusion detection focuses on ether to minimize false alarms rate or to maximize detection rate (the rate of attacks detected successfully). High detection rate and low false positive rate are required for any good intrusion detection system. In order to determine the relationship between false alarms rate and detection rate we used Receiver Operating Characteristic (ROC) curve as a performance evaluation metric to evaluate our intrusion detection algorithm. In intrusion detection, the ROC curve is usually used to measure the performance of the detection model [15, 19]

- True positive (TP): examples predicted positive which are correctly predicted
- True negative (TN): examples predicted negative which are correctly predicted
- False positive (FP): negative examples that are incorrectly predicted positive
- False negative (FN): positive examples that are incorrectly predicted negative
- Accuracy (ACC): The percentage of correction predictions to the total number of predictions

Table 2: Performance comparison

PREDICTION MODEL	TPR	FPR	ACC
Using one classifier CP-KNN	0.973	0.0228	0.9800
Using CP-KN and DOD	0.9933	0.0060	0.9867

Table 2 shows the results of the detection model using one classifier CP-KNN algorithm, and the detection model using both classifiers CP-KNN and DOD using the same metrics. It shows that the detection model that uses both CP-KNN and DOD measures achieves a higher detection rate than using single classifier. The false positive rate is also decreased. Moreover, the prediction accuracy of the combined prediction model is higher than the model using a single classifier.

Table 3: Experimental results on CPDOD

DATA SET	TPR	FPR	ACC
BHAT10	0.99432	0.001958	0.99734
BHAT30	0.97005	0.011584	0.97973
BHAT60	0.98022	0.00669	0.98874
RCAT10	0.99605	0.001917	0.99767
RCAT30	0.96444	0.01167	0.97874
RCAT60	0.99162	0.008642	0.99139
DRAT10	0.99228	0.001931	0.99689
DRAT30	0.97473	0.015633	0.9804
DRAT60	0.97202	0.014939	0.97973
BRDAT	0.98851	0.011494	0.98007

Table 4: Experimental results on C4.5

DATA SET	TPR	FPR	ACC
BHAT10	0.99	0.008	0.997
BHAT30	0.968	0.058	0.958
BHAT60	0.972	0.033	0.98
RCAT10	0.99	0.008	0.99
RCAT30	0.96	0.019	0.97
RCAT60	0.984	0.033	0.985
DRAT10	0.99	0.008	0.99
DRAT30	0.97	0.025	0.98
DRAT60	0.989	0.06	0.97
BRDAT	0.981	0.015	0.98

Table 5: Experimental results on K-NN

DATA SET	TPR	FPR	ACC
BHAT10	0.944	0.201	0.9
BHAT30	0.833	0.167	0.73
BHAT60	0.891	0.287	0.84
RCAT10	0.944	0.056	0.9
RCAT30	0.957	0.043	0.917
RCAT60	0.914	0.201	0.915
DRAT10	0.979	0.23	0.926
DRAT30	0.87	0.15	0.84
DRAT60	0.943	0.057	0.91
BRDAT	0.957	0.081	0.915

Table 3, Table 4 and Table 5 show the detail running results of three machine learning detection model on various attack datasets.

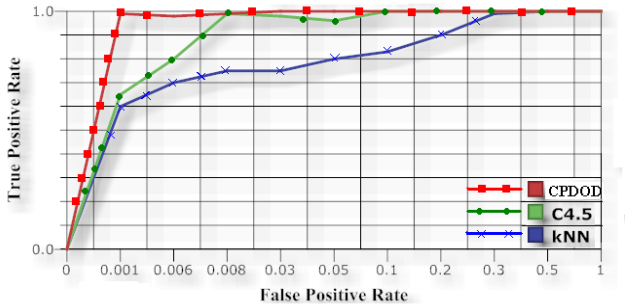

Figure 2: RCO curves showing the performance of our method and other two algorithms over Black Hole Attack dataset

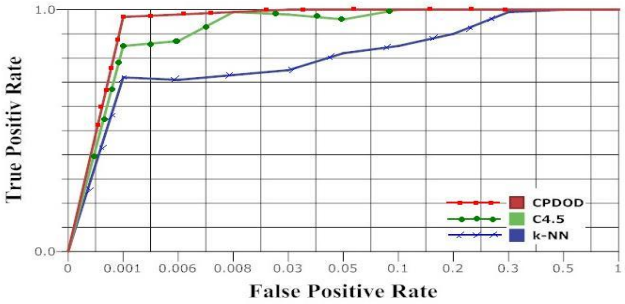

Figure 3: RCO curves showing the performance of our method and other two algorithms over Resource Consumption Attack dataset

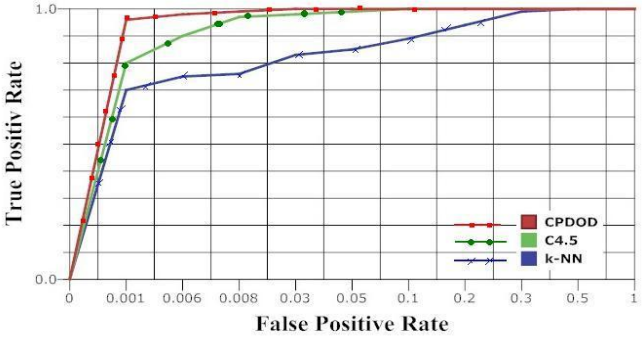

Figure 4: RCO curves showing the performance of our method and other two algorithms over Dropping Routing Traffic Attack dataset

Figure 2 shows the ROC curves of the three algorithms on the black hole attack dataset. It has been seen that all the anomaly detection algorithm can detect black hole attack, but our detection method achieves a higher detection performance. Figure 3 and Figure 4 shows the performance metric obtained by each algorithm over resource consumption attack dataset and dropping routing traffic attack dataset.

5. Conclusions and future work

In this paper, we propose a novel intrusion detection method using the combined two anomaly methods Conformal Predictor K-Nearest Neighbor (CP-KNN) and Distance-based Outlier Detection (DOD). The proposed algorithm employs a combined model that uses two different measures (nonconformity metric measures and Outlier Factor LDOF metric) to improve its detection ability. nonconformity metric measures whether the unknown instance is more similar to known normal instances or abnormal instances. The Outlier Factor LDOF metric identifies the similarity to normal classes and can detect abnormal attacks. We implemented our detection algorithm and tested the detection approach over three common attacks dataset (resource consumption attack, dropping routing traffic Attack and black hole attack) to evaluate the performance of our Dynamic Intrusion Detection method. A series of experimental results demonstrate that the proposed method can effectively detect anomalies with low false positive rate, high detection rate and achieve higher detection accuracy.

6. REFERENCES

[1] Daniel Barbará, Carlotta Domeniconi, and James P. Rogers. Detecting outliers using transduction and statistical testing. In KDD '06: Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining, pages 55_64, New York, NY, USA, 2006. ACM.

[2] Varun Chandola, Arindam Banerjee, and Vipin Kumar. Anomaly detection: A survey. ACM Comput. Surv., 41(3):1_58, 2009.

[3] Hongmei Deng, Roger Xu, Jason Li, Frank Zhang, Renato Levy, andWenke Lee. Agent-based cooperative anomaly detection for wireless ad hoc networks. In ICPADS '06: Proceedings of the 12th International Conference on Parallel and Distributed Systems, pages 613_620, Washington, DC, USA, 2006.

[4] Yingfang Fu, Jingsha He, and Guorui Li. A distributed intrusion detection scheme for mobile ad hoc networks. Computer Software and Applications Conference, Annual International, 2:75_80, 2007.

[5] Alex Gammerman and Volodya Vovk. Prediction algorithms and con_dence measures based on algorithmic randomness theory. Theor. Comput. Sci., 287(1):209_217, 2002.

[6] Alexander Gammerman and Vladimir Vovk. Hedging predictions in machine learning. Comput. J., 50(2):151_163, 2007.

[7] GloMoSim. Glomosim website, June 2007.

[8] Yi-an Huang, Wei Fan, Wenke Lee, and Philip S. Yu. Cross-feature analysis for detecting ad-hoc routing anomalies. In ICDCS '03: Proceedings of the 23rd International Conference on Distributed Computing Systems, page 478, Washington, DC, USA, 2003. IEEE Computer Society.

[9] A. Karygiannis, E. Antonakakis, and A. Apostolopoulos. Host-based network monitoring tools for manets. In PE-WASUN '06: Proceedings of the 3rd ACM international workshop on Performance evaluation of wireless ad hoc, sensor and ubiquitous networks, pages 153_157, New York, NY, USA, 2006. ACM.

[10] Yang Li, Binxing Fang, Li Guo, and You Chen. Network anomaly detection based on tcm-knn algorithm. In ASIACCS '07: Proceedings of the 2nd ACM symposium on Information, computer and communications security, pages 13_19, New York, NY, USA, 2007. ACM.

[11] Yang Li and Li Guo. An active learning based tcm-knn algorithm for supervised network intrusion detection. Computers & Security, 26(7-8):459_467, 2007.

[12] Yihua Liao and V. Rao Vemuri. Use of k-nearest neighbor classifier for intrusion detection, 2002.

[13] Tom M. Mitchell. Machine Learning. McGraw-Hill, New York, 1997.

[14] C. Siva Ram Murthy and B.S. Manoj. Ad Hoc Wireless Networks: Architectures and Protocols. Prentice Hall PTR, Upper Saddle River, NJ, USA, 2004

[15] Hadi Otrok, Joey Paquet, Mourad Debbabi, and Prabir Bhattacharya. Testing intrusion detection systems in manet: A comprehensive study. Communication Networks and Services Research, Annual Conference on, 0:364_371, 2007.

[16] Animesh Patcha and Jung-Min Park. Network anomaly detection with incomplete audit data. Comput. Netw., 51(13):3935_3955, 2007.

[17] Charles Perkins and Elizabeth Royer. Ad-hoc on-demand distance vector routing. In Proceedings of the 2nd IEEE Workshop on Mobile Computing Systems and Applications, pages 90_100, 1997.

[18] Glenn Shafer and Vladimir Vovk. A tutorial on conformal prediction. J. Mach. Learn. Res., 9:371_421, 2008.

[19] W J Ulivla. Evaluation of intrusion detection system. J. J. Res. Natl. Inst. Stand. Technol., 108(6):453_473, 2003.

[20] Liwei vivian Kuang. Dnids: A dependable network intrusion detection system using the csi-knn algorithm, 2007

[21] Fu Xiao and Xie Li. Using outlier detection to reduce false positives in intrusion detection. In NPC '08: Proceedings of the 2008 IFIP International Conference on Network and Parallel Computing, pages 26_33, Washington, DC, USA, 2008. IEEE Computer Society.

[22] Ke Zhang, Marcus Hutter, and Huidong Jin. A new local distancebased outlier detection approach for scattered real-world data. CoRR, abs/0903.3257, 2009.

[23] Yongguang Zhang, Wenke Lee, and Yi-An Huang. Intrusion detection techniques for mobile wireless networks. Wirel. Netw., 9(5):545_556, 2003.