JAIDM
Journal of AI and Data Mining
Vol. 1, No.2, 2013, 89-101.

Dynamic anomaly detection by using incremental approximate PCA in AODV-based MANETs

M. Alikhani*, M. Ahmadi Livani

Faculty of Electrical and Computer Engineering Tarbiat Modares University

Received 09 January 2013; accepted 28 January 2013
**Corresponding author: m.alikhany@gmail.com (M. Alikhani)*

Abstract
Mobile Ad-hoc Networks (MANETs) in contrast to other networks have more vulnerability because of having nature properties, such as dynamic topology and no infrastructure. Therefore, a considerable challenge for these networks, is a method expansion that can specify anomalies with high accuracy at network dynamic topology alternation. In this paper, two methods were proposed for dynamic anomaly detection in MANETs, namely IPAD and IAPAD. The anomaly detection procedure consists of three main phases: Training, detection and updating the two methods. In the IPAD method, to create the normal profile, we used the normal feature vectors and principal components analysis in the training phase. In detection phase, during each time window, anomaly feature vectors based on their projection distance from the first global principal component specified. In updating phase, at end of each time window, normal profile updated by using normal feature vectors in some previous time windows and increasing principal components analysis. IAPAD is similar to IPAD method with a difference that each node use approximate first global principal component to specify anomaly feature vectors. In addition, normal profile will be updated by using approximate singular descriptions in some previous time windows. The simulation results using NS2 simulator for some routing attacks show that an average detection rate and an average false alarm rate in IPAD method had 95.14% and 3.02% respectively. The IAPAD method had 94.20% and 2.84% respectively.

Keywords*: MANETs, Dynamic Anomaly Detection, Routing attacks, Incremental Principal Component Analyses.*

1. Introduction
Mobile Ad hoc Networks (MANETs) are collections of wireless and mobile nodes that there is not any fixed infrastructure, such as base stations. In recent years, the advent of wireless devices was the cause of these networks potential growth. Today, MANETs are used in military battlefield, emergency rescue and vehicular communications because of its easy and rapid development [1]. In MANETs for sake of nodes mobility, network topology changes rapidly. Due to lack of centralized management in these networks, each node accomplishes routing process.

Intrusion detection methods are divided into two main categories: Signature-based detection and anomaly detection [2]. In signature-based detection methods, known intrusion patterns compared with incoming traffic and if patterns matched, intrusion is recognized. Advantage of this method is low false alarm rate and its disadvantage is lack of new intrusion detection. In anomaly detection methods, first, a profile of network normal behavior created then any traffic deviated from created profile detected as an intrusion. Advantage of this method is new intrusions detection and its disadvantage is the high false alarm rate.

In this paper, we proposed two methods named IPCA and IAPAD, which let normal profile get updated dynamically. Proposed methods contain three phases: Training, Detection and Updating. IPAD method, in training phase, creates network normal profile by using normal feature vectors. In

detection and updating phase, a normal profile gets updated by using normal feature vectors in each time window. IAPAD method, in training phase, calculates an approximate singular description for normal feature vectors in each time window, then in detection phase, IAPAD calculates approximate covariance matrix by using approximate singular description.
In the updating phase by approximate covariance matrix, singular value parsing calculates the first approximate global principal component. Evaluations show that proposed methods have significant performance.
In section II, we imply related works in MANETs anomalies detection field. In section III, AODV protocol and in section IV, the attacks against this protocol described shortly. In section V, we have a description about how to select features. In section VI, principal components analysis is explained. In section VII, anomaly detection based on increasing principal components is explained. In section VIII, dynamic anomaly detection based on increasing approximate principal components analysis is represented. In section IV, accomplished simulation results are reported for evaluation. Finally, in section VI, we state the conclusion of this paper.

2. Related works

Huang *et al*. [3] proposed a method that uses a cross-feature analysis to capture inter-feature correlation patterns in the normal traffic. They create normal profile by using a *C* classifier and the network normal traffic. *C* classifier applied on every f_i feature and a C_i classifier will be created as sub model. Finally, these sub models will be used as normal profile. In this method, normal profile just created from training data and always is stable. Regarding to nodes dynamic behavior in MANET, fixed normal profile cannot qualify current network state well.
Huang *et al.* [4] used both specification-based and statistical-based approaches to detect attacks on AODV. First, they model AODV normal behavior by an extended finite state automaton (EFSA), according to its specifications. EFSA model is utilized for anomaly behaviors detection and they are deviated from descriptions. Statistics training algorithms with statistical properties are used for anomaly behaviors detection that is essentially statistical. In this method, normal profile is always fixed and does not alter with nodes behavior changes.
Sun *et al.* [5] proposed a method focusing on the mobility in MANETs. In this method, first, in training phase, various models of routing actions mobility has been collected and the link change rate (LCR_{recent}) average will be calculated for each mobility level. Collected routing actions utilized for normal profile creation. Then, in detection phase, each local intrusion detection system calculates link change rate for its own nodes, which are recent routing actions alternatively. Among normal profiles, a profile selected its LCR has less Euclidean distance with LCR_{recent}. In each time slot, each node calculates LCR as for its new and old neighbors. Therefore, LCR calculating does not spot the whole of inter-network nodes. However, attention must be paid for that network estate change and this is because of other network nodes with a sudden appearance and disappearance. When node's behavior in detection phase is different from training phase, using a predefined normal profile cannot describe network behavior well.
Kurosawa *et al.* [6] proposed a method with dynamic learning to detect anomalies in MANETs. This method updates training data in symmetric time slots. They used three features to model AODV protocol normal behavior that of course the protocol behavior complexity cannot be a model well with these three features. In this method, network normal profile considered as normal data average. Their method is only able to detect Blackhole attack and is not able to detect more attacks.
Nakayama *et al.* [7] proposed a method to detect dynamic anomaly that use principal components analysis for network normal profile creation. This method used normal data global covariance in sequential time slots for a created profile update. For each time slot, a weight is considered and is used as a factor in covariance calculating. This method uses weigh covariance in principal components calculating (WPCA). In this method, global covariance will be calculated inexactly.
Raj *et al.* [8] proposed a dynamic learning system to detect Blackhole attack. In this system, the node that received RREP packet compares packet's sequential number with a threshold value that updated dynamically. If the sequential number was greater than threshold value, RREP packet transmitter should be added to black list as an attacker node. This method is just able to detect Blackhole attack and is unable to detect other attacks.

3. AODV routing protocol

AODV protocol is a reactive routing protocol [9]. Protocol uses destination sequence number

concept in DSDV routing protocol for maintaining last routing information. Suppose, start node S attempts to communicate with destination node D. In lack of routing information aspect, S starts path discovery via a RREQ packet broadcasting to its neighboring nodes. By receiving RREQ packet having fresh routing information, each neighbor node replies S node via a RREP packet. Otherwise, a hop count field increases RREQ packet unit age and broadcasts this packet again to its neighbors. Also keeps routing information to create inverse path. N_i node to make sure about routing information freshness compares destination sequence number in RREQ packet with a D node's sequence number in its own routing table. If a D node's sequence number in routing table is lesser, this sequence number will be updated with destination sequence number at RREQ packet. If node N receives several RREP packets, select the packet that has greater destination sequence number. If destination sequence number of received RREP packets is equal, the packet will be selected which has lesser hop count. The start node starts a data packet sending as soon as first RREP packet receives. Each node for make sure about active paths validity, broadcasts a HELLO packet alternatively to its neighbors. When a node detects a link fraction, announce that to other nodes by creating a REEP packet. Figure 1 shows routing process in AODV protocol.

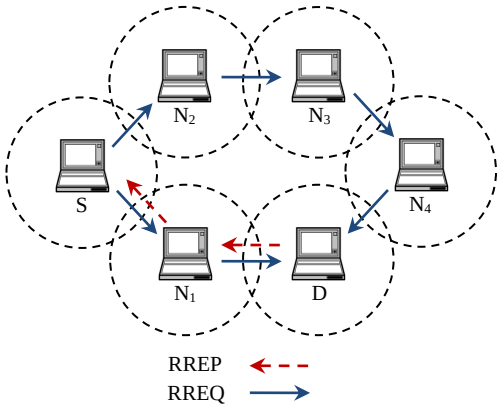


Figure 1. Routing process in AODV protocol

4.Attacks against AODV protocol

A) Classification of attacks

Attacks against AODV protocol are divided into four categories:

1) Route Disruption: A malicious node either destroys an existing route or prevents a new route from getting established.
2) Route Invasion: A malicious node adds itself into route between source and destination nodes.
3) Node Isolation: A given node is prevented from communicating with any other nodes. It differs from route disruption in the route disruption is targeting at a route with two given nodes, while node isolation is targeting at all possible routes to or from a given node.
4) Resource Consumption: The communication bandwidth in the network or storage space at individual nodes is consumed.
In the following, we give a short description of some typical routing attacks on AODV [10].

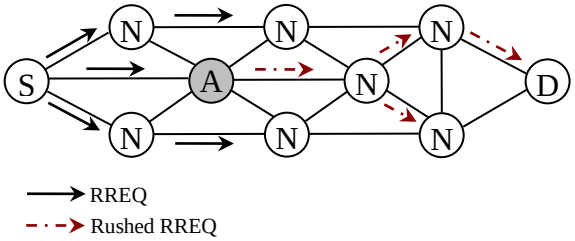


Figure 2. Rushing Attack

B)Typical Attacks

1) Rushing Attack: Each source node establishes routing process by a RREQ packet transmission. In each routing process, each intermediate node just accepts the first received RREQ packet and ignores repetitive packets. Also, each intermediate node leads received RREQ packets after a delay. Malicious node by abusing these properties, immediate after each RREQ packet receiving, sends it to the next node. By this method, probability of malicious node standing between source and destination path will increases [11]. Fig. 2 shows a rushing attack example. In this figure, N_6 and N_7 nodes receive directed RREQ packet faster than other directed packets by malicious node.
2) Neighbor Attack: In AODV protocol, each intermediate node adds its ID in the RREQ/RREP packets before forwarding it to the next node. In neighbor attack, malicious node forward RREQ or RREP packet to the next node without its ID adding. Malicious node's wrong behavior makes other nodes to save false information about its neighbors in routing tables.

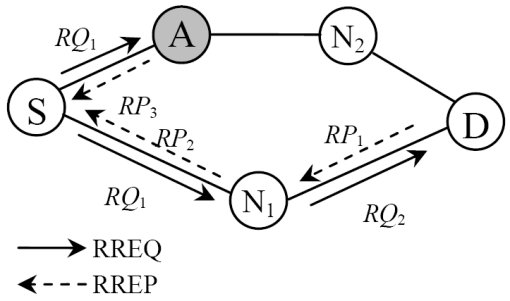


Figure 3. Blackhole Attack

3) Blackhole Attack: Malicious node with false routing information transmission claims that it has an optimized path to destination node. With this false claim, other nodes send their packets to the malicious node [12]. In AODV routing protocol, malicious node can perform this attack by sending a fake RREP packet to the source node. Figure 3 shows a Blackhole attack example. Source node S attempts to communicate with destination node D. Also, suppose, node D sequence number value in node S routing table is 20. Node N_1 by receiving RQ_1 packet forwards that to node D. malicious node A by receiving RQ_1 packet responses to node S with RP_3 packet. Node S according to destination sequence number field selects introduced path by malicious node and transmits its data to invalid node Z. Above packets details are presented in Table 1.

Table 1. The rreq/rrep packet in blackhole attack

	RQ_1	RQ_2	RP_1	RP_2	RP_3
Source IP Address	S	N_1	D	N_1	Z
Destination Sequence Number	20	20	21	21	30
Origin IP Address	S	S	S	S	S
Destination IP Address	D	D	D	D	D
Hop Count	1	2	1	2	1

4)Flooding RREQ Attack: Generally, RREQ packets will be broadcasted for new paths finding. Malicious node broadcasts because of network resources construction and alternatively many of fake RREQ packets.

5.Features definition

The appropriate feature selection for anomalies detection in routing process is the first and the most important action that must be performed. In this paper, nineteen features are used for anomaly detection in MANETs. These features are classified in four categories:

1) Traffic data related features:Each node in the network can send, receive and forwards data packet. These actions against data packets can define three features.

2) Path discovery related features:RREQ and RREP used for between source and destination nodes path finding and routing tables updating. By using these packets and various actions performed on them can define various features.

3) Path interruption related features:Some of paths disrupted cause of node mobility. Paths disrupting will be a cause of RREQ and RREP packets missing. For snatched paths reparation in AODV protocol, RERR packet is used. Proportionate these attributes can define several properties.

*4) AODV protocol specific featur*e: Difference average between destination sequence number in received RREP packet and destination sequence number in transmitted RREQ packet can be defined as a feature in each node.

The first class features are beneficial for data traffic anomaly behavior detection that can be due to a Denial of Service (DoS). The second class features are beneficial for attacks detection creating anomaly in network with routing protocol behavior change. The third class features indicant is seen routing faults rate in the network. Some of attacks alter routing faults rate through creating anomaly in the network. In Blackhole attack, malicious creates anomaly in network normal behavior by fake RREP packets that contains a great destination sequence number transmission. The fourth class features are beneficial for detect of this type of anomaly. In Table 2, name and description of each feature represented.

6.Principal components analysis

Principal component analysis (PCA) is a well-known method for patterns analysis in data [9]. By PCA, the first principal component φ that shows data approximate distribution is calculated. Let X be an $n \times p$ data matrix, whose rows are the feature vectors and columns are the features:

$$X = \begin{bmatrix} x_1^1 & x_1^2 & \cdots & x_1^p \\ x_2^1 & x_2^2 & \cdots & x_2^p \\ \vdots & \vdots & \ddots & \vdots \\ x_n^1 & x_n^2 & \cdots & x_n^p \end{bmatrix}, \quad (1)$$

Let, $\hat{X}$ is a column-center matrix of X:

$$\hat{X} = (I - \frac{1}{n} e_n e_n^T) X \quad (2)$$

Principal components of X are obtained by singular vector decomposition (SVD) [9] of $\hat{X}$ matrix [7]:

$$\hat{X} = U\Sigma V^T \quad (3)$$

where$U$ and V are left and right singular vectors of $\hat{X}$ matrix respectively, and $\Sigma = \text{diag}(\sigma_1, \ldots, \sigma_p)$ is a diagonal matrix with singular values. In this paper, the quadruple (n, V, Σ, μ) is called as s singular description of X and represents with D_X.

Table 2. The features

Type	Feature	Description
CBR Traffic	NumSentCbrPkt	Number of sent CBR data packets
	NumRecvCbrPkt	Number of received CBR data packets
	NumFwdCbrPkt	Number of forwarded CBR data packets
Route Discovery	NumSentRReqPkt	Number of sent RREQ packets
	NumRecvSameSrcRReqPkt	Number of received RREQ packets with the same source address as the node
	NumRecvSameDstRReqPkt	Number of received RREQ packets with the same destination address as the node
	NumRecvDiffSrcDstRReqPkt	Number of received RREQ packets with the different source and destination address of the node
	NumFwdRReqPkt	Number of forwarded RREQ packets
	NumSentSameDstRRepPkt	Number of sent RREP packets with the same destination address as the node
	NumSentDiffDstRRepPkt	Number of sent RREP packets with the different destination address of the node
	NumRecvSameSrcRRepPkt	Number of received RREP packets with the same source address as the node
	NumRecvDiffSrcRRepPkt	Number of received RREP packets with the differentsource address of the node
	NumFwdRRepPkt	Number of forwarded RREP packets
Path Disrupting	NumSentRErrPkt	Number of sent RERR packets
	NumRecvRErrPkt	Number of received RERR packets
	NumFwdRErrPkt	Number of forwarded RERR packets
	NumDropRReqPkt	Number of dropped RREQ packets
	NumDropRRepPkt	Number of dropped RREP packets
Protocol Specific	AvgDiffDstSeqNum	Average difference at each time slot between destination sequence number of received RREP packet and stored sequence number in the node

In this description, V is principal components and μ is a column-center vector of X. We can use the first principal component φ for describes X. Let, C_X is covariance matrix of X:

$$C_X = \frac{1}{n-1} X^T (I - \frac{1}{n} e_n e_n^T) X \quad (4)$$

Right singular vectors of X are equal to principal components of C_X, also the kth special value of C_X is equal to the kth square of $\hat{X}$ matrix singular value:

$$C_X = \frac{1}{n-1} V \Sigma^2 V^T \quad (5)$$

where, V and $\Sigma^2 = \text{diag}(\lambda_1, \ldots, \lambda_p) = \text{diag}(\sigma_1^2, \ldots, \sigma_p^2)$ are principal components matrix and eigenvalues matrix of X respectively. According to (5) specified that the equation can gain X singular description by analyze singular values of $(n-1)C_X$ matrix.

7.Dynamic anomaly detection based on ipca

In this section, we proposed an increasing principal components analysis method named IPAD for dynamic anomaly detection in MANETs. In this method, each time window t contains several time slots. In each time slot $t_i \in t$, each node collects a x_{it} feature vector on its traffic.

$$x_{it} = [x_{it}^1, x_{it}^2, \ldots, x_{it}^p]^T, \quad (6)$$

where each x_{it}^j is a measurable feature. So, each t time window, collects each node of $X(t)$ matrix from feature vectors. In this paper, to establishing normal profile and anomaly detection, 19 mentioned features are used.

A $x_{it} \in X(t)$ feature vector is called normal if agrees with network normal traffic in t time window. Set of normal feature vectors in t time window is represented with $X_N(t)$ and set of whole normal feature vectors in m maximum time window before t is represented with $N_X(t)$.

$$N_X(t) = \bigcup_{\tau=t-m+1}^{t} X_N(\tau) \quad (7)$$

IAPAD method contains three phases: *Training*, *Detection* and *Updating* from which any of these three phases are describe on resumption.

Training Phase:

In this phase, each node collects $N_X(0)$ matrix from feature vectors by its traffic supervision at beginning, then scales each values of $N_X(0)$ features to [0,1] slot. Finally, $\varphi(0)$ first principal component calculating creates a network normal profile. Figure 4 shows the pseudocode of training phase.

Normalization of feature vectors:

The value of each feature vector can have a considerable difference with each other. So, when distance of between two feature vectors is calculated, the features with larges values conquest on features with lower values. For making sure about the whole of features, they have same affection on distance calculation, each $x_{i0}^{j} \in N_X(0),\ j=1,...,p$ feature vector values must scale with in [0,1] slot.

$$\hat{x}_{i0}^{j} = \frac{x_{i0}^{j} - \min(x^{j}(0))}{\max(x^{j}(0)) - \min(x^{j}(0))} \qquad (8)$$

That $\min(x^{j}(0))$ and $\max(x^{j}(0))$ are the smallest and greatest j feature values in $N_X(0)$ respectively.

procedure Training

input:
A set of normal feature vectors $N_X(0)$
output:
A normal profile $P(0) = (\varphi(0), \mu(0), d_{\max}(0))$
begin
Scale each feature of $N_X(0)$ to the range of [0,1]
Obtain the column-centered matrix $\hat{N}_X(0)$
Obtain the column-means vector $\mu(0)$
Compute the first principal component $\varphi(0)$
for each feature vector $x_{i0} \in N_X(0)$ **do**
Compute the projection distance $d_p(x_{i0}, \varphi(0))$
end for
$d_{\max}(0) = \max_i \{d_p(x_{i0}, \varphi(0))\}$
end procedure

Figure 4. The training phase

Establishing a Normal Profile:

For the normal profile creation, at first, each node generates $\hat{N}_X(0)$ column-centered matrix for $\hat{N}_X(0)$. Then by $\hat{N}_X(t)$ matrix singular value decomposition calculates the first $\varphi(0)$ principal component and each $x_{i0} \in N_X(0)$ feature vector's projection distance from $\varphi(0)$ is attained.

$$d_p(x_{i0}, \varphi(0)) = (\| x_{i0} - \mu(0) \|^2 - (\varphi(0)^T \cdot (x_{i0} - \mu(0)))^2)^{\frac{1}{2}} \qquad (9)$$

where $\mu(0)$ is $N_X(0)$ column-means vector.

On resume, the maximum of projection distance of all x_{i0} feature vectors from $\varphi(0)$ is calculated and uses that for anomaly detection:

$$d_{\max}(0) = \max_i \{d_p(x_{i0}, \varphi(0))\} \qquad (10)$$

Finally, uses $(\varphi(0), \mu(0), d_{\max}(0))$ triplet for $P(0)$ normal profile creation.

Detection Phase:

In this phase, each node during each t time window collects $X(t)$ matrix from feature vectors by its traffic supervision. Then scale features values of each $x_{it} \in X(t)$ feature vector by using minimum and maximum features values in $N_X(t-1)$ and then compares scaled feature vectors with $P(t-1) = (\varphi(t-1), \mu(t-1), d_{\max}(t-1))$ normal profile to detect anomaly traffic.

procedure Detection

input:
A normal profile $P(t-1) = (\varphi(t-1), \mu(t-1), d_{\max}(t-1))$
A set of feature vectors $X(t)$
output:
A set of normal feature vectors $X_N(t)$
begin
$X_N(t) = \varnothing$
Scale each feature of $X(t)$ using min and max of $N_X(t-1)$
for each feature vector $x_{it} \in X(t)$ **do**
Compute the projection distance $d_p(x_{it}, \varphi(t-1))$
if $d_p(x_{it}, \varphi(t-1)) \le d_{\max}(t-1)$ **then**
$X_N(t) = X_N(t) \bigcup \{x_{it}\}$
end if
end for
end procedure

Figure 5. The detection phase

Anomaly Detection:

For anomaly detection, each node calculates projection distance of each $x_{it} \in X(t)$ feature vector from $\varphi(t-1)$ that $\varphi(t-1)$ is the first global principal component until $t-1$ time window. If calculated projection distance were greater than $d_{\max}(t-1)$, x_{it} would be detected as an anomaly feature vector:

$$\begin{cases} d_p(x_{it}, \varphi(t-1)) > d_{\max}(t-1) & : \text{Anomaly} \\ d_p(x_{it}, \varphi(t-1)) \le d_{\max}(t-1) & : \text{Nomal} \end{cases} \qquad (11)$$

Figure 5, shows the pseudo code of detection phase.

Normal Profile Updating Phase:

In this phase, each node at each t time window ending, if normal network state is detected, updates normal profile in this time window by using normal feature vectors. Thus, first, add collected normal feature vectors in t time window to $N_X(t-1)$:

$$N_X(t) = N_X(t-1) \bigcup X_N(t) \tag{12}$$

That $X_N(t)$ is set of collected normal feature vectors in t time window. Nodes mobility in MANETs is cause of topology similar to network behavior alternation. Each set of feature vectors shows the network state and its connection time. By considering to the rapid behavior changing of network, this feature vectors set cannot show the network state in further times well. So, weight to each set of feature vectors can be useful for dynamic anomaly detection. Assume, $x_{i\tau} \in N_X(t)$ normal feature vector in τ time window be collected, an oblivion relation calculate this feature vector weight in current t time window:

$$w_{i\tau}(t) = \begin{cases} w_0 e^{-\alpha r_m(\tau,t)\Delta T(t-\tau)} & (t-m) \le \tau \le t \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

That $\alpha \in [0,1]$ and w_0 parameters is determined by a user. ΔT is time window length and $r_m(\tau,t)$ is network topology changing rate between τ and t time windows. Network topology changing rate is determined by using neighbor nodes number:

$$r_m(\tau,t) = \frac{|N(\tau) - N(t)| + |N(t) - N(\tau)|}{n}, \tag{14}$$

That n is the number of whole nodes in the network. $N(\tau)$ and $N(t)$ are neighbor nodes index in τ and t time windows, respectively. Each node just uses a set of collected normal feature vectors in maximum m previous time window. $w_{i\tau}(t)$ weights be bounded by (15) relation:

$$\sum_{\tau=t-m+1}^{t} w_{i\tau}(t) = 1 \tag{15}$$

If weight of one normal feature vector is lesser from a ε threshold value, the feature vector will be deleted form $N_X(t)$ set.

For normal profile updating, at first, each node by using relation (2), generates $\hat{N}_X(t)$ column-centered matrix for $N_X(t)$, then by $\hat{N}_X(t)$ matrix singular value analysis, calculates the first $\varphi(t)$ global principal component and finally the maximum projection distance of whole $x_{i\tau} \in N_X(t)$ feature vectors from $\varphi(t)$ is attained:

$$d_{\max}(t) = \max_{i,\tau} \left\{ d_p(x_{i\tau}, \varphi(t)) \right\} \tag{16}$$

The $(\varphi(t), \mu(t), d_{\max}(t))$ triplet shows $P(t)$ updated normal profile. In Figure 6, represent normal profile updating in each node.

procedure Updating

input:
A set of normal feature vectors $N_X(t)$
output:
A normal profile $P(t) = (\varphi(t), \mu(t), d_{\max}(t))$
begin
for each feature vector $x_{i\tau} \in N_X(t)$ **do**
Update the weight $w_{i\tau}(t)$
if $w_{i\tau}(t) < \varepsilon$ **then**
$N_X(t) = N_X(t) \setminus \{x_{i\tau}\}$
endif
end for
Obtain the column-centered matrix $\hat{N}_X(t)$
Obtain the column-means vector $\mu(t)$
Find the global first principal component $\varphi(t)$
for each feature vector $x_{i\tau} \in N_X(t)$ **do**
Compute the projection distance $d_p(x_{i\tau}, \varphi(t))$
end for
$d_{\max}(t) = \max_i \left\{ d_p(x_{i\tau}, \varphi(t)) \right\}$
end procedure

Figure 6. The updating phase

8.Increasing proximate components analysis based dynamic anomaly detection:

In IPAD method, first global principal component calculating is accomplished strictly. In this method, at each time window ending for normal profile updating, a set of normal feature vectors in previous time windows is used. This problem is cause of calculating complexity and memory usage in crescent in each node. For this problem solution, an increasing proximate principal components analysis based method is proposed which named IAPAD that decreases calculating complexity and memory usage in each node. In this method, each node in current time window t, calculates $\tilde{D}_N(t) = (n(t), \tilde{V}(t), \tilde{\Sigma}(t), \mu(t))$ proximate singular description for $X_N(t)$ normal feature vectors. The time window in this description, $n(t)$ is number of normal feature vectors, $\tilde{V}(t)$ is the

matrix which contains k is the most important principal component, $\tilde{\Sigma}(t)$ is the matrix contains k the greatest special value and μ is $X_N(t)$ column-means vector. The k is the minimum value that relation (17) confirmed that:

$$\frac{\sum_{j=1}^{k}\lambda_j^2}{\sum_{j=1}^{p}\lambda_j^2} \geq T_\alpha \qquad (17)$$

where p and λ_j^2 are the number of features and $X_N(t)$ matrix special value, respectively. T_α Named as proximate quality threshold bound and thus by using k the most important principal component can describe T_α percent of data dispersion. The set of proximate singular descriptions in maximum m time window before t is presented with $\tilde{N}_D(t)$:

$$\tilde{N}_D(t) = \bigcup_{\tau=t-m+1}^{t} \tilde{D}_N(\tau) \qquad (18)$$

Each node instead of $N_X(t)$ normal feature vectors set maintaining, and keeps $\tilde{N}_D(t)$ singular description set.

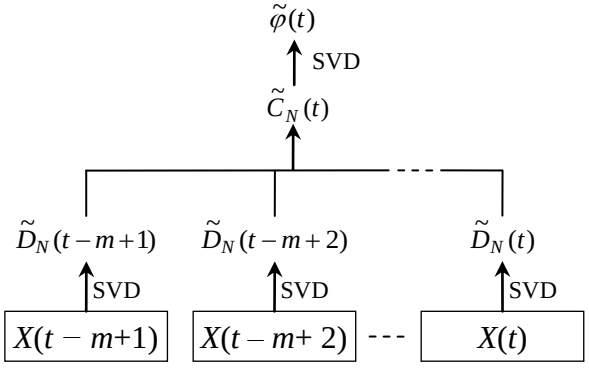


Figure 7. IAPAD Description

IAPAD method contains three phases: *Training*, *Detection* and *Updating* noted above are similar to IPAD method with a difference that instead of $\varphi(t)$ first global principal component, used from $\tilde{\varphi}(t)$ first proximate global principal component. For calculation of $\tilde{\varphi}(t)$, first, $\tilde{C}_N(t)$ proximate covariance matrix for normal feature vectors in maximum m previous time window calculated [13]:

$$\tilde{C}_N(t) = \frac{1}{n-1}\sum_{\tau=t-m+1}^{t}(\tilde{V}(\tau)\tilde{\Sigma}^2(\tau)\tilde{V}^T(\tau) + n(\tau)(\mu(\tau)-\mu)(\mu(\tau)-\mu)^T) \qquad (19)$$

where μ is global column-means vector.

$$\mu = \frac{1}{n}\sum_{\tau=t-m+1}^{t} n(\tau)\mu(\tau) \qquad (20)$$

Value of n is equal to normal feature vectors total in maximum m time window:

$$n = \sum_{\tau=t-m+1}^{t} n(\tau) \qquad (21)$$

Then, by $\tilde{C}_N(t)$ matrix singular value analysis, the first $\tilde{\varphi}(t)$ proximate global principal component is calculated. Figure 7 shows a description for this method.

9.Time complexity analysis:

In this section, the first $\varphi(t)$ global principal component time complexity calculating in IPAD method and the first $\tilde{\varphi}(t)$ proximate global principal component in IAPAD will be compared with each other.

In IPAD method, in each t time window, each node, first, generates $\hat{N}_X(t)$ column-centered matrix for $N_X(t)$. This matrix generation is done in time of $O(mnp)$, that n is normal feature vectors number average in each time window, then by $\hat{N}_X(t)$ matrix singular value analysis, calculates the first $\varphi(t)$ global principal component in time of $O(mnp^2)$. Therefore, $\varphi(t)$ calculating has $O(mnp^2)$ time complexity.

In IAPAD method, in each t time window, each node, at first, calculates $\tilde{D}_N(t)$ proximate singular description for normal feature vectors in this time window. This singular description calculating is accomplished in time of $O(np^2)$. Then calculates $\tilde{C}_N(t)$ proximate covariance matrix by using singular description in maximum m previous time window. This covariance matrix calculating is accomplished in time of $O(mp^2)$. Finally, by $\tilde{C}_N(t)$ matrix singular value analysis, calculates the first $\tilde{\varphi}(t)$ proximate global principal component in time of $O(p^3)$. Therefore, $\tilde{\varphi}(t)$ calculating has $O(np^2)$ time complexity. Mention is require that $p < n$ and $m < n$.

10.Experiment results:

In this section, at first, impact of routing attacks on MANETs performance will be studied, and then accomplished experiment results are described for proposed IPAD and IAPAD

performance evaluation.

Simulation Environment:

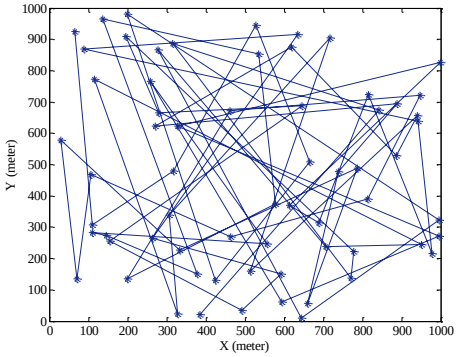
We conducted MANET simulations using the NS2 simulator [14]. In this simulation, CBR traffic model with 512-byte data packet length generated through *cbrgen.tcl* program and RWP [15] mobility model in a region dimensioned 1000m×1000m and 5sec pause time, generated by the *setdest* program. The number of whole network nodes, includes 30 nodes. Table 3 shows a detail of simulation parameters represented.

Table 3. Simulation Parameters

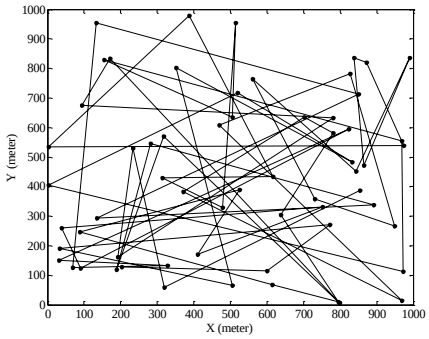
Parameter	*Value*
Simulation Time	10000(s)
Mobility Model	RWP
Pause Time	5(s)
Maximum Mobility	35(m/s)
Maximum Connections	30
Maximum bandwidth	2(Mbps)
Number of Malicious Nodes	1
Simulation Area	1000(m) × 1000(m)
Transmission Rate	250(m)
Traffic Model	CBR
Routing Protocol	AODV

In RWP mobility model, each node for a specific time length (pause time) locates in a simulation region and after this time ending, a random destination selection with a steady speed moves from [0, *maxspeed*] slot to the destination. The node after reaching to a new location, positions there within pause time and then begins mobility process again.

Figure 8 shows malicious node and a node of the network mobility model. In this figure, pause time 5 seconds and speed bound [0, 35m/s] has been selected. Regarding to Figure 8(a) specified that malicious node attends steady in different location of simulation environment. Therefore, any anomaly behavior from malicious affects entire network. This local distribution is also seen for other network nodes (Figure 8(b)).



(a)



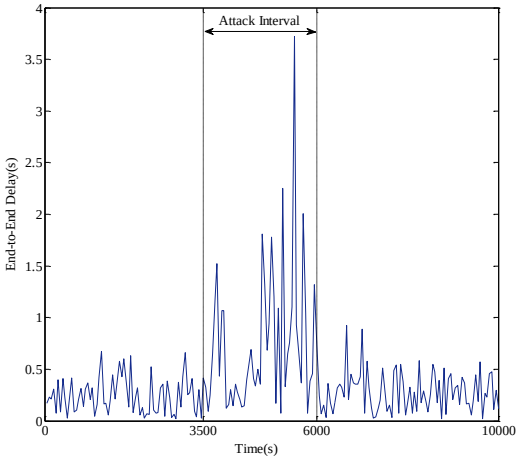
(b)

Figure 8. Mobility model in RWP: a) Malicious Node b) A Node of The Network

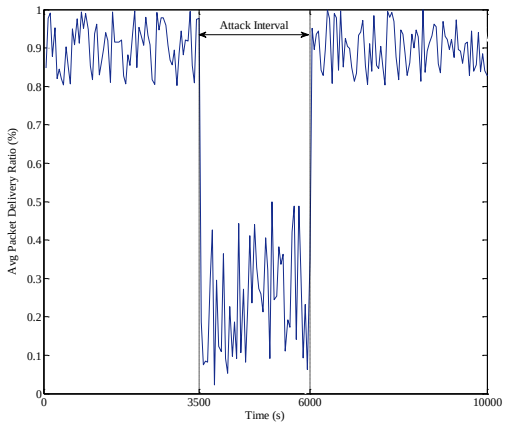
Impact of routing attacks on Network Performance

In this section, we will study about impact of routing attacks on MANET performance by using NS2 simulator. There are many parameters such as End-To-End Delay and Packet Delivery Ratio for MANET performance measurement [16]. End-To-End Delay refers to the time taken for a packet to be transmitted across a network from source to destination. The packet delivery ratio of a receiver is defined as the ratio of the number of data packets, which actually received over the number of data packets transmitted by the senders. Routing attacks are cause of network performance decrement by anomaly creation in the network.

Figure 9 shows, impact of blackhole attack on End-To-End Delay parameters and Packet Delivery Ratio represented respectively. As seen on this figure, in the above attack occurrence time, network performance decreased noticeably.



(a)



(b)

Figure 9. Impact of Blackhole Attack: a) On the average End-To-End Delay Parameter b) On the average Packet Delivery Ratio

Performance Evaluation

To establish the normal profile, a set of feature vectors are collected by each node of network normal traffic. This set of feature vectors collecting time length are considered 1000sec and time slot length for any feature vector collecting considered 5sec. One of the nodes selected is a malicious node. This node, accomplished rushing, neighbor, blackhole and flooding RREQ attacks is distinctly in 3500-6000sec-time interval. An experiment, used for normal feature vectors sets and proximates singular descriptions in maximum m=5 previous time window, and the length of time window is also selected ΔT=200s.

For performance evaluation of anomaly detection methods, two measures used detection rate (DR) and false alarm rate (FAR). Detection rate is a percent of anomaly feature vectors that have been detected successfully. False alarm rate is a percent of normal feature vectors that have been detected as anomaly feature vectors inaccuracy.

Figure 10 shows detection rate and false alarm rate averages in IPAD and IAPAD have been compared within different values of time window length ΔT=500, 400, 300, 350, 200, 150, 100s represented detection rate and false alarm rate in this figure. This calculated as blackhole, rushing, neighbor and flooding RREQ attacks detection rate and false alarm rate averages.

Regarding to the above figure specified that by time window length decrement or by the other hand, by normal profile rapid updating, detection rate increases. In addition, IAPAD method has similar performance with IPAD method and it has lesser time complexity and lesser usage memory. Table 4 shows detection rate and false alarm rate averages in IAPAD method has been compared by the m parameter different values. Regarding to this table specified that by m value decrement, for normal profile updating used from lesser singular description. Therefore, the updated normal profile cannot model current time network normal traffic well with regarding to the low number of singular description, false alarm rate increased. By m value increasing, using from old singular descriptions to normal profile is updated. Therefore, the updated normal profile cannot model current time network normal traffic well with regarding to the high number of old singular descriptions, and detection rate are decreased.

In Figure 11, for one of network nodes, feature vectors projection distance from the first proximate global principal component during blackhole attack represented. In 3500-6000sec-time distance, projection distance of many collected feature vectors during each time window from the first proximate global principal component calculated to its prior time window is greater than a threshold bound. So, this feature vectors detected as anomaly and above time distance is considered as attacks time distance.

Figure 12 shows detection rate and false alarm rate averages in IAPAD method compared with each other through different time window length values. Regarding to this figure, specified that by time window length decrement to 200 seconds, detection rate increases noticeably. For time windows with under 200sec, detection rate against false alarm rate is so fiddling. So, in accomplished experiment, time window length selected as ΔT=200.

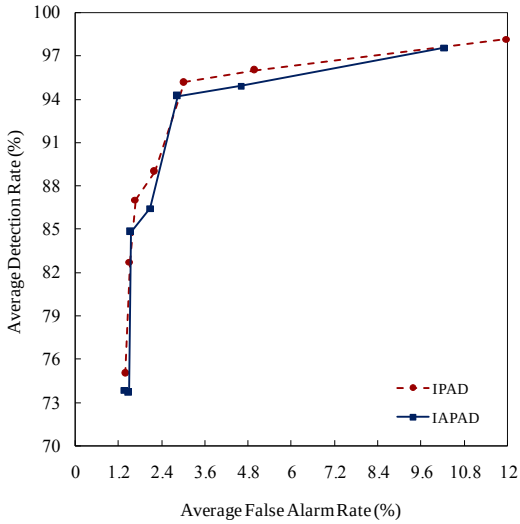


Figure 10. Detection rate and false alarm rate averages in IPAD and IAPAD

Table 4. Detection rate and false alarm rate averages in IAPAD

m	Rushing		Neighbor		Blackhole		Flooding	
	DR	***FAR***	***DR***	***FAR***	***DR***	***FAR***	***DR***	***FAR***
1	100	36.07	100	29.02	100	30	100	31.21
3	95.83	8.76	97.5	6.06	95.42	5.99	95.17	7.81
5	95.83	3.27	96.53	2.63	92.22	2.29	92.23	3.16
10	87.5	1.38	93.75	2.02	85	1.13	89.95	1.80

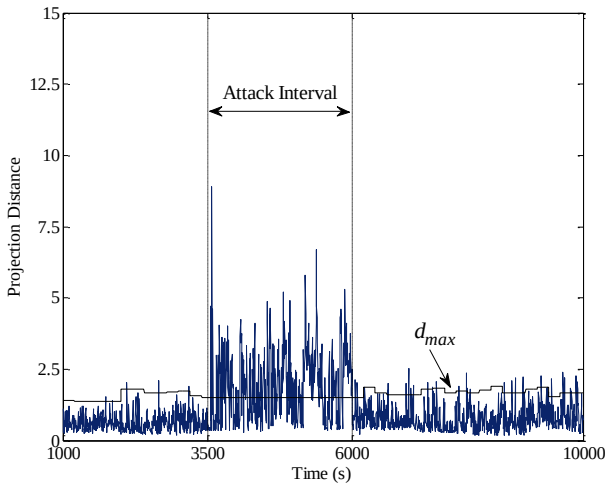


Figure 11. Projection distance from the first proximate global principal component during blackhole attack

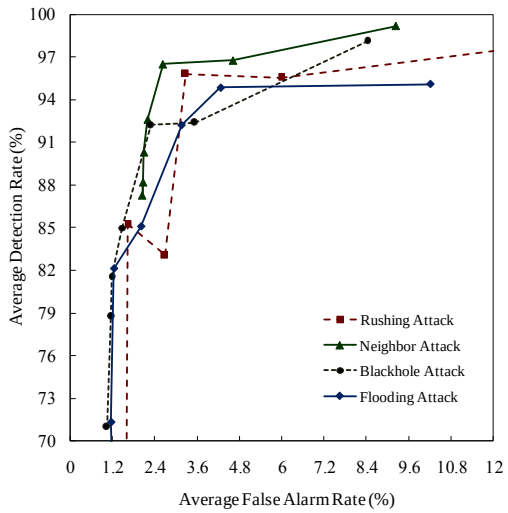


Figure 12. Detection rate and false alarm rate averages in IAPAD method compared with each other by different time window length values

Cumulative Percent Variance (CPV) [17] is a standard that represents the described variance percent by the most important principal components. In fact, complex variance percent determines importance level of each principal component to complex variance percent calculating for each i principal component used the relation (22):

$$CPV_i = \frac{\lambda_i}{\sum_{j=1}^{p} \lambda_j} \qquad (22)$$

where λ_i is special value corresponding with *i*th principal component.

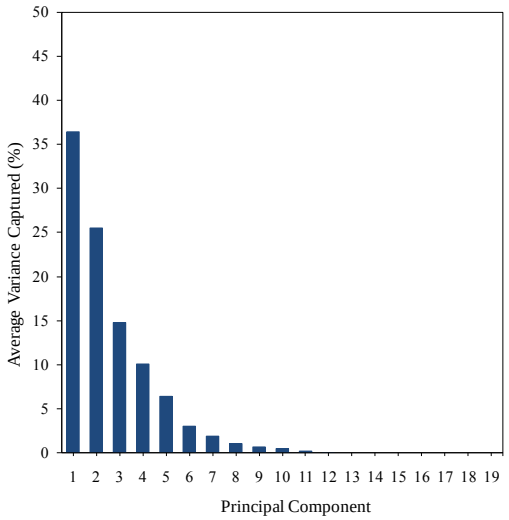


Figure 13. CPV average in detection phase represented for each principal component in IAPAD method

Figure 13 shows, cumulative percent variance average in detection phase represented for each principal component in IAPAD method. Regarding to this figure specified that the first principal component describes only 36.46 percent of total variance. Therefore, for better data dispersal modeling, it is necessary the second principal component considered with 25.45 percent of total variance in proximate singular description calculation time. In this face, with *k*=2 principal component can describe 69.91 percent of data dispersal. In accomplished experiment, threshold bound of approximation quality considered equal to $T_\alpha = 50\%$.

Table 5 shows impact of updating in performance on the IAPAD method represented for various type of attacks. Regarding to this table, the detection rate and false alarm rate averages in IAPAD method in the face of normal profile updating are 94.20 and 2.84 percent, respectively and in the face of nonupdating are 59.72 and 1.70 percent in respectively.

Table 5. Impact of updating in performance on the IAPAD method

	With Updating		Without Updating	
	DR	***FAR***	***DR***	***FAR***
Rushing	95.83	3.27	38.60	0.64
Neighbor	96.53	2.63	74.60	2.88
Blackhole	92.22	2.29	58.4	1.32
Flooding	92.23	3.16	67.28	1.98
Average	94.20	2.84	59.72	1.70

Figure 14 shows detection rate and false alarm rate in IPAD, IAPD and WPCA [7] methods compared with each other. Regarding to this figure specified that detection rate average in IPAD and IAPAD methods is 4/40 and 3/46 percent better than WPCA method, when false alarm rate average in WPCA method is 0/53 and 0/35 percent better than IPAD and IAPAD methods.

Table 6 shows detection rate and false alarm rate averages in IPAD and IAPAD and WPCA methods compared with each other by the breakdown of each rushing, neighbor, blackhole and flooding RREQ attacks.

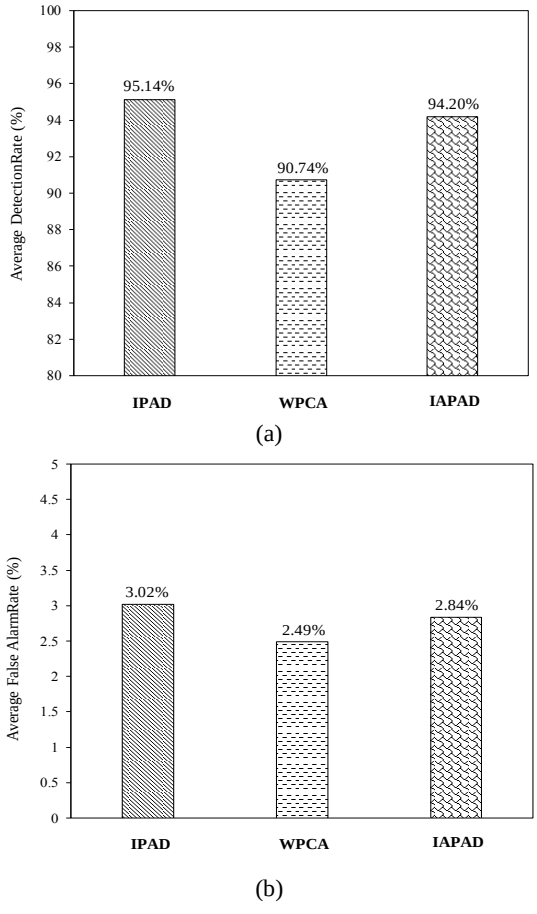


Figure 14. Comparison of the performance of IPAD, IAPD and WPCA: a) average detection rate b) average false alarm rate.

Table 6. Comparison of the performance of IPAD and IAPAD and WPCA methods

	Rushing		Neighbor		Blackhole		Flooding	
	DR	***FAR***	***DR***	***FAR***	***DR***	***FAR***	***DR***	***FAR***
IPAD	**98.33**	4.6	**96.53**	2.84	91.88	2.35	**93.83**	**2.27**
IAPAD	95.83	3.27	96.53	2.63	**92.22**	**2.29**	92.23	3.16
WPCA [7]	90.83	**2.18**	95.34	**2.22**	86.52	2.39	90.25	3.18

11.Conclusion

Regarding to dynamic topology in MANETs, the cause of alternation in network behavior, using from a predefined normal profile cannot describe network behavior well. Therefore, it is necessary to update normal profile coincident with network nodes and topology behavior alternations. In this paper, two increasing principal components analysis based on methods named IPAD and IAPD proposed for dynamic anomaly detection in MANETs. Proposed methods contain 3 phases: Education, detection and normal profile updating. In IPAD method, in education phase, by using normal feature vectors and principal component analysis, network traffic usual profile will be created. In detection phase, during each time window, a set of feature vectors to be collected and anomaly feature vectors based on their projection distance detected from the first global principal component. In the updating phase, in each time window ending, usual profile will be updated by using normal feature vectors in this time window and previous time windows. Updating is accomplished by using increasingly principal components analysis and an oblivion relation. IAPAD method is similar to IPAD method with this difference that any node in any time window calculates a proximate singular description of normal feature vectors in the time window. In addition, instead of the first global principal component used from the first proximate global principal component for anomaly feature vectors detection. Usual profile updated by using proximate singular description in current and previous time windows. For MANETs implementation and also rushing, neighbor, blackhole and flooding RREQ attacks used from NS2 simulator. Routing attacks are the cause of network performance decrement through creating anomaly in the network. By using the End-To-End Delay and Packet Delivery Rate parameters, impact of above attacks network performance is studied. The performance evaluation of proposed IPAD and IPAD methods used two standards,

which are detection and false alarm rate. Regarding to the accomplished experiment results, IAPAD method has a similar performance with IPAD method when it has less time complexity and usage memory.

Time windows length in the usual profile updating time can be affective on the detection rate and false alarm rate increment of decrement. The IAPAD method performance is evaluated by various values of time window length. Regarding to experiment results in this method, $\Delta T=200s$ time window length establishes a better balance between detection rate and false alarm rate. In IAPAD method for proximate singular description, calculation is used for k the most important principal component. Using standard of Cumulative Percent Variance (CPV) importance level of each principal component in detection phase is calculated. Regarding to experiment results in normal face, $k=2$ principal components can describe 61.91 percent of data dispersal. Various experiment accomplishments, performance of IPAD and IAPAD methods are compared with WPCA method for rushing, neighbor, blackhole and flooding RREQ attacks detection. Experiment results show that detection rate average in IPAD and IAPAD methods respectively 4.40 and 3.46 percent better than WPCA method. False alarm rate average in WPCA method is only 0.53 and 0.35 percent better than IPAD and IAPD methods.

Acknowledgment

This work was supported by Iran Telecommunication Research Center (ITRC) under contract number 88-12-128.

References

[1] Chlamtac, I., Conti, M. and. Liu, J. J.-N. (2003). Mobile ad hoc networking: imperatives and challenges. Ad Hoc Networks, 1(1), 13–64.

[2] Debar, H., Dacier, M. and Wespi, A. (2000). A revised taxonomy for intrusion detection systems," Annals of Telecommunications. 55(7–8), 361–78.

[3] Huang, Y. A., Fan, W., Lee, W. and Yu, P. S. (2003) Cross-feature analysis for detecting ad-hoc routing anomalies, in Proceedings of the 23rd International Conference on Distributed Computing Systems. 478–487, Washington DC, USA.

[4] Huang, Y. A. and Lee, W. (2004) Attack analysis and detection for ad hoc routing protocols, in Proceedings of the 7th International Symposium on Recent Advances in Intrusion Detection (RAID'04), 125–145, Riviera, French.

[5] Sun, B., Wu, K. and Pooch, U. (2004). Towards adaptive intrusion detection in mobile ad hoc networks," in Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM'04), 6, 3551–3555, Dallas, TX, USA.

[6] Kurosawa, S., Nakayama, H., Kato, N., Jamalipour, A. and Nemoto, Y. (2007). Detecting blackhole attack on AODV-based mobile ad hoc networks by dynamic learning method," International Journal of Network Security, 5(3), 338–346.

[7] Nakayama, H., Kurosawa, S., Jamalipour, A., Nemoto, Y. and Kato, N. (2009). A dynamic anomaly detection scheme for AODV-based mobile ad hoc networks," IEEE Transactions on Vehicular Technology, 58(5), 2471–2481.

[8] Raj, P. N. and Swadas, P. B. (2009). DPRAODV: A dynamic learning system against blackhole attack in AODV-based MANET," International Journal of Computer Science Issues, 2(1), 54–59.

[9] Perkins, C. E., Royer, E. M. B. and Das, S. R. (2003). Ad hoc on-demand distance vector (AODV) routing. RFC 3561, July 2003.

[10] Ning, P. and Sun, K. (2003). How to misuse AODV: A case study of insider attacks against mobile ad-hoc routing protocols," in Proceedings of the 4th Annual IEEE Information Assurance Workshop, pp. 60–67, West Point, NY, USA, June 2003.

[11] Chun Hu, Perrig, A. and johnson, D. B. (2003). Rushing attacks and defense in wireless ad hoc network routing protocols," in Proceeding of the 2nd ACM workshop on Wireless security, pp. 30–40, San Diego, CA, USA, September 2003.

[12] Hongsong, Y. C., Zhenzhou, J. and Mingzeng, H. (2006). A novel security agent scheme for AODV routing protocol based on thread state transition. Asia Journal of Information Technology. 5(1), 54–60.

[13] Qu, Y., Ostrouchov, G., Samatova, N. and Geist, A. (2002). Principal component analysis for dimension reduction in massive distributed data sets, in Proceedings of the SIAM International Conference on Data Mining, Washington, DC, USA, April 2002.

[14] Fall, K. and Varadhan, K. (2002). The NS manual, The VINT Project, UC Berkeley, LBL, USC/ISI, and Xerox PARC, April 2002.

[15] Camp, T., Boleng, J. and Davies, V. (2002). A survey of mobility models for ad hoc network research," Wireless Communications and Mobile Computing. 2(5), 483–502.

[16] Nguyen, H. L. and Nguyen, U. T. (2008). A study of different types of attacks on multicast in mobile ad hoc networks," Ad Hoc Networks, 6(1), 32–46.

[17] Jolliffe, I. T. (2002). Principal Component Analysis, Second Edition, New York: Springer-Verlag.