Volume 8, No. 5, May-June 2017
ISSN No. 0976-5697
International Journal of Advanced Research in Computer Science
REVIEW ARTICLE
Available Online at www.ijarcs.info

Analysis Based Ids for Mobile Ad-Hoc Network

Pravin N. Khobragade
Department of Computer Engineering
Ramrao Adik Institute of Technology
Navi Mumbai (M. S.), India

Prof. Vimla Jethani
Department of Computer Engineering
Ramrao Adik Institute of Technology
Navi Mumbai (M. S.), India

Abstract: Data mining technique is widely used everywhere. Mobile Adhoc Network (MANET) is the wireless network in which each node moves freely. The node inside the network acts both i.e. router and host as well. I will analyze the data mining technique such as classification and clustering for separating malicious, selfish nodes from loyal node. In clustering a network is created which is further divided into partitioned network with certain node as a cluster head.
The main issue in MANET is security of the node. It is possible that the node is turned into malicious or selfish node. In order to prevent packets from the malicious node the clustering technique is use to separate the nodes having intrusive behavior from the one having normal behavior. To overcome that problem we have to use intrusion detection system (IDS) which helps to improve the security of the node.

Keywords: Analysis Based IDS, Mobile Ad-Hoc Network, NIDS, HIDS

INTRODUCTION

Data Mining

Data mining, the process of grabbing useful data from large database, is a latest technology used in industries to highlight on most important information in their data warehouses. Data mining tool analyse behaviour and feature trends, which gives business to take knowledge driven resolution. Data mining offered automated, prospective analysis which got more developed than its previous events. Data mining tool respond to business which is commonly tedious to resolve.

MANET

MANET contains mobile nodes which creates limited network without the fixed infrastructure of central administration. The direct communication between the nodes is possible within the transmission range. Nodes which are not in range can communicate via intermediate node; such scenario is also called as multihop network. In this transmission, the packet will be transmitting to the other nodes until it reaches its destination using routing protocol. For better behaviour assistance of node is required. Here cooperation refers to performing the network functions collectively by nodes for benefit of other nodes. But because of free framework and flexibility of nodes, compliance may occur which critically reduce the work of network.

MANET is weak to several types of attacks because of free framework, inadequacy of central administration, dynamic network topology, finite battery based energy of mobile nodes. These are both internal as well as external attack. Various strategies had been planned already that entirely focus on detection and prevention of outside attacks. But if suspicious node is already entered the network then almost of this schemes will be useless.

Intrusion Detection System (IDS)

Intrusion detection is a method of detecting suspicious activity to the system or computer network. Intrusion detection technique is helpful to find illegitimate intrusion into system or network. An intrusion detection system (IDS) can be software application or device application that observe system or network for suspicious activities or policy violations and generate reports to a management station. IDS comes in a various methods trying to detect malicious activities on the network. There are two type of intrusion detection systems i.e. network based (NIDS) and host based (HIDS) intrusion detection system.

Network Intrusion Detection System

Network intrusion detection system (NIDS) are located at a critical point or points in between the network to observe traffic from every device on the network. It observes traffic flow on every subnet, and compares the traffic flow with the passed subnet for known attacks. Once an attack will be found or abnormal behaviour is sensed, it will send the alert to the administrator.

An example of INDS would be installing in the subnet where firewalls are placed in order to check if somebody is trying to break into the firewall. Usually one would search all inbound and outbound traffic, causes bottleneck problem, which affects the overall speed of the network. OPNET and NetSim are generally used for simulating intrusion detection system. NID Systems are also capable of comparing signatures for similar packets to link and drop harmful detected packets which have a signature matching the records in the NIDS.

Host Intrusion Detection Systems

Host Intrusion Detection Systems (HIDS) run on individual hosts or devices on the network. The inbound and outbound packets from the device are being surveillance by host based intrusion detection. If it found any suspicious activity it will alert the user or administrator. It takes a snapshot of existing system files and matches it to the previous snapshot. If the critical system files were modified or deleted, an alert is sent to the administrator to investigate. An example of HIDS usage can be seen on mission critical machines, which are not expected to change their configurations.

How data mining used in IDS

Intrusion detection is used to detect the nodes which are malicious nodes, selfish nodes and loyal nodes. Data is clustered using data mining technique and grouped as per category i.e. malicious, selfish and loyal nodes. Once the

© 2015-19, IJARCS All Rights Reserved

data is clustered together then that particular nodes will be discarded to prevent the cluster head.

LITERATURE SURVEY

Zohaib et al. [1] discussed the clustering problem in MANET. Over the course of the last decade, there have been several improvements in the performance of Integer Linear Programming (ILP) and Boolean Satisability (SAT) solvers. These improvements have encouraged the application of SAT and ILP techniques in modeling complex engineering problems. One such problem is the Clustering Problem in Mobile Ad-Hoc Networks (MANETs). The Clustering Problem in MANETs consists of selecting the most suitable nodes of a given MANET topology as cluster heads, and ensuring that regular nodes are connected to cluster heads such that the lifetime of the network is maximized.

This paper proposes the development of an improved ILP formulation of the Clustering Problem. Additionally, various enhancements are implemented in the form of extensions to the improved formulation, including the establishment of intra cluster communication, multi hop connections and the enforcement of coverage constraints. The improved formulation and enhancements are implemented in a tool designed to visually create network topologies and cluster them using state of the art Generic ILP and SAT solvers. Through this tool, feasibility of using the proposed formulation and enhancements in a real life practical environment is assessed. It is observed that the Generic ILP solvers, CPLEX, and SCIP, are able to handle large network topologies, while the 01 SAT based ILP solver, BSOLO, is effective at handling the smaller scale networks.

It is also observed that while these enhanced formulations enable the generation of complex network solutions, and are suitable for small scale networks, the time taken to generate the corresponding solution does not meet the strict requirements of a practical environment. This paper puts forward an improved ILP formulation to solve the clustering problem in MANETs.

The proposed model presented the use of a Star Ring backbone. Additionally, the proposed formulation included the ability to enforce coverage constraints to ensure that only connections that are within the physical limitations of the node are established. The enhancements include the ability for nodes within the same cluster to communicate without going through the designated cluster head, and the ability to establish multi hop links. Using the proposed ILP formulations and enhancements together with a custom designed tool, it was possible to test the performance and analysis the feasibility of Generic ILP.

Zhao et al. [2] presented loose virtual clustering protocol (LVC). Power heterogeneity is common in mobile ad hoc networks (MANETs).

With high power nodes, MANETs can improve network scalability, connectivity, and broadcasting robustness. However, the throughput of power heterogeneous MANETs can be severely impacted by high power nodes. To address this issue, a loose virtual clustering based (LVC) routing protocol used for power heterogeneous (LRPH) MANETs. To explore the advantages of high-power nodes, they have developed LVC algorithm to construct a hierarchical network and to eliminate unidirectional links. To reduce the interference raised by high power nodes, routing algorithms are got developed to avoid packet forwarding via high power nodes. Via the combination of analytical modeling, simulations, and real world experiments, we demonstrate the effectiveness of LRPH on improving the performance of power heterogeneous MANETs. Because of development of a LVC based routing protocol named LRPH for power heterogeneous MANETs.

LRPH is considered to be a double edged sword because of its high power nodes. We designed an LVC algorithm to eliminate unidirectional links and to benet from high power nodes in transmission range, processing capability, reliability, and bandwidth. They have developed routing schemes to optimize packet forwarding by avoiding data packet forwarding through high power nodes. Hence, the channel space utilization and network throughput can be largely improved. Through a combination of analytical modeling and an extensive set of simulations, the effectiveness of LRPH over power heterogeneous MANETs. A proof of concept system on Microsoft WinCE has been also implemented, and real world experiments have been conducted and validated our theoretical and simulation endings well.

Jiang [3] discussed a prediction-based link availability estimation to quantify the link reliability. A Mobile Ad hoc Network (MANET) is a collection of wireless mobile terminals that are able to dynamically form a temporary network without any aid from xed infrastructure or centralized administration. One critical issue for routing in MANETs is how to select reliable paths that can last as long as possible since terminal mobility may cause radio links to be broken frequently. To solve this problem, a criterion that can judge path reliability is needed. The reliability of a path depends on the number of links and the reliability of each link constituting the path.

Many routing metrics in terms of number of links have been proposed, such as the shortest path routing. However, how to measure link availability or reliability in order to nd more reliable paths has not been addressed adequately in the literature. (By a link being available, mean that the radio quality of the link satisfies the minimum requirement for successful communication. Link availability is used to measure probability or degree that a link is available. The terms availability and reliability are used interchangeable in this paper.)The quantity makes use of some instantly available information and also considers the dynamic nature of link status in order to properly react the link reliability. Then, this quantity has been further used to develop routing metrics for path selection in terms of path reliability to improve routing performances.

Srivastava et al. [4] introduced the Mobile Ad-Hoc Networks (MANET) consist of peer to peer infrastructure less communicating nodes that are highly dynamic. As a result, routing data becomes more challenging. Ultimately routing protocols for such networks face the challenges of random topology change, nature of the link (symmetric or asymmetric) and power requirement during data transmission. Under such circumstances both, proactive as well as reactive routing are usually inefficient. The zone routing protocol (ZRP) that adds the qualities of the proactive (IARP) and reactive (IERP) protocols. In ZRP, an updated topological map of zone centered on each node, is maintained. Immediate routes are available inside each zone.

© 2015-19, IJARCS All Rights Reserved

In order to communicate outside a zone, a route discovery mechanism is employed. The local routing information of the zones helps in this route discovery procedure.

In MANET security is always an issue. It is possible that a node can turn malicious and hamper the normal flow of packets in the MANET. In order to overcome such issue we have used a clustering technique to separate the nodes having intrusive behavior from normal behavior. This technique as effective k-means clustering which has been motivated from k-means. They have proposed an Intrusion Detection System on each node of the MANET which is using ZRP for packet flow. Then they used effective k-means to separate the malicious nodes from the network. Thus, the Ad-Hoc network will be free from any malicious activity and normal flow of packets will be possible.

Lung-Chung Li and Ru-Sheng Liu [5] discussed the address key management in cluster-based mobile ad hoc networks (MANETs). Ensuring secure communication in an ad hoc network is extremely challenging because of the dynamic nature of the network and the lack of centralized management. For this reason, key management is particularly difficult to implement in such networks. They have presented a fully distributed ID-based multiple secrets key management scheme (IMKM). This scheme is implemented via a combination of ID-based multiple secrets and threshold cryptography. It eliminates the need for corticated based authenticated public- key distribution and provides an efficient mechanism for key update and key revocation schemes, which leads to more suitable, economic, adaptable, scalable, and autonomous key management for mobile ad hoc networks. They have proposed a secure, efficient, and scalable distributed ID-based multiple secrets key management scheme (IMKM) for cluster-based mobile ad hoc networks. In order to address the highly dynamic topologies and varying link qualities of ad hoc networks, the master secret key is generated and distributed by all cluster heads. As a result, not only are central instances avoided, which constitute single points of attack and failure, but this also leads to more autonomous and exible key update methods.

Nicklas Beijar [6] suggested Zone Routing Protocol (ZRP). Karmore et al. [7] discussed Detecting Intrusion on AODV based Mobile Ad Hoc Networks by k- means Clustering method of Data Mining. Jabas et al. [8] suggested MANET Mining: Mining Temporal Association Rules. Nadeem et al. [9] discussed on a Survey of MANET Intrusion Detection & Prevention Approaches for Network Layer Attacks. Nadim and Howarth [10] used ABID(Anomaly based Intrusion Detection) to detect intrusion in the network, this requires traffic traces that contain only normal activities to build a training profile. However, in contrast with fixed networks, data resources such as that reflect normal activities or events are not currently available for MANETs. Therefore they assumed that the initial behavior of the network formed on-the- fly is free from anomalies. To illustrate the implementation of GIDP (Generalized Intrusion Detection and Prevention) they assumed a clustered MANET organization. They selected the most capable nodes in terms of their processing abilities as cluster heads (CHs) and the others nodes becomes cluster nodes (CNs).

GIDP is a hybrid IDP approach that uses a combination of anomaly-based and knowledge based ID. A cluster head first gathers data in the form of two matrices which are network characteristic matrix (NCM) and a derived matrix (DM). The NCM contains data related to the network routing protocol. The DM consists of parameters which reflects the network performance and can be derived from NCM parameters. Then the cluster head employs two phases: training and testing. When the network is established, the CH continuously gathers NCM and DM information and applies the GIDP training module for N time intervals (TI), resulting in initial training profiles (ITPs) of NCM and DM. The ITPs reflects the normal behavior of the nodes in the network and the expected network performance. In the testing phase the CH applies the testing module after each TI. The testing phase consists of several tasks. Firstly it detects intrusion in the network. If there is no intrusion then it updates the ITPs in order to adapt the variation in the network behavior as time progresses. If there is intrusion, in the second task the CH identifies the attack or attacks using existing information in the knowledge base. In the case of known attacks the CH identifies intruding nodes using intruder identification rules specific to the known attack. To optimize the probability of identifying intruders correctly with a low level of false positives, it maintains a test sliding window (TSW), in which d detections of a node are required in p time intervals (TI). If this detection threshold is passed then the CH will blacklist the node and isolate the node by informing all CNs. If attack identification detects an attack that does not match the rules for known attack then CH applies the attack inferences. Attack inference stores the rule trace of current TI as Detected Rule Trace and looks for its match in a TSW. If the match is found in a TSW then CH confirms the new attack by constructing adding a rule for the new attack in the set of rules stored in knowledge base.

KEY CONCEPTS

Data Mining Technique

Clustering

Clustering is an automated process to group related records together. Related records are grouped together on the basis of having similar values for attributes. It is a main task of exploratory data mining, and a common technique for statistical data analysis, used in many fields, including machine learning, pattern recognition, image analysis, information retrieval, and bio-informatics.

Example in a library, there is a wide range of books in various topics available. The challenge is how to keep those books in a way that readers can take several books in a particular topic without hassle. By using clustering technique, we can keep books that have some kinds of similarities in one cluster or one shelf and label it with a meaningful name. If readers want to grab books in that topic, they would only have to go to that shelf instead of looking for entire library.

In this technique, data points are clustered together based on their similarity factors and is often nearness according to some defined distance. Clustering is an effective way to find hidden patterns in data that humans might miss. It is useful for ID as it can cluster malicious and nonmalicious activity separately.

Classification

Classification is a classic data mining technique based on machine learning. Basically classification is used to classify each item in a set of data into one of predefined set of

© 2015-19, IJARCS All Rights Reserved

classes or groups. Classification method makes use of mathematical techniques such as decision trees, linear programming, neural network and statistics. In classification, we develop the software that can learn how to classify the data items into groups.

Example we can apply classification in application that given all records of employees who left the company, predict who will probably leave the company in a future period. In this case, we divide the records of employees into two groups that named leave and stay. And then we can ask our data mining software to classify the employees into separate groups.

It is a data mining technique used to map data instances into one of the various predefined categories. It can be used to detect individual attacks but it has high rate of false alarm. Various algorithms like decision tree induction, Bayesian networks, k-nearest neighbor classifier, case-based reasoning, genetic algorithm and fuzzy logic techniques are used for classification techniques. The classification algorithm has been then applied to audit data collected which then learns to classify new audit data as normal or abnormal data.

Intrusion Detection Mechanism

Intrusion Detection System (IDS) is a system which is used to prevent attacks in MANET. Attacks in MANET are classified as active and passive attacks. A passive attack do not disturb the actual functioning of the network whereas the active attack affects the actual functioning of the network as in this unauthorized person tries to extract data being exchanged. The key objective of attacker is to analyze the traffic and location from which the data has been sent. The main purpose of IDS is to provide security to MANET. There are a variety of intrusion detection systems which are used to identify different types of attacks using different ID techniques. One of the intrusion detection technique is anomaly based intrusion detection (ABID) which is also known as behavior based intrusion detection system. Second one is knowledge based intrusion detection system. It is also called as knowledge based intrusion detection system (KBID) and the third one is specification based intrusion detection system (SBID).

***Anomaly Based Intrusion Detection system* (ABID)**

This system is based on the observation of the deviation from the normal traffic. This system basically consists of two parts namely testing and training. Training phase is to capture the knowledge from the existing normal traffic and testing phase is further used to test the present scenario against the anomalies. This method is workable for selected number of attacks and those attacks should occur due to abnormal functioning of the network. It involves training from the operation of non malicious nodes. Good quality training is a key to success of the anomaly based intrusion system.

***Knowledge Based Intrusion Detection system* (KBID)**

It is based on knowledge base which is nothing but a repository where we store information about various attacks. Firstly the system is trained using the training data of the normal nodes and then various rules are inferred from that data and is stored in the knowledge base. This information is stored in the knowledge base is further utilized to detect intrusions. Knowledge base is a repository based on which decision is taken to declare a node as a malicious node. One of the most important part of this system is the decision component which generates appropriate alerts to make the system secure.

***Specification Based Intrusion Detection system* (SBID)**

This intrusion detection system is based on the use of specifications. Specifications are the documented form of records of various attacks in the form of constraints. These specifications are used to monitor the unusual transmission of data from source to destination. The main task of SBID is to extract specification based on which intrusions are identified. The specification also provides guidelines to identify normal and correct form of operation of the network. In this we accuracy of specification plays an important role as the whole system relies on the specification to detect intrusion in the network. So the specifications also needs to be verified and then should be used in the system to detect the intrusions.

***Zone Routing Protocol* (ZRP)**

Zone Routing Protocol [4], or ZRP is a hybrid Wireless Networking routing protocol that uses both proactive and reactive routing protocols when sending information over the network. ZRP was designed to speed up delivery and reduce processing overhead by selecting the most efficient type of protocol to use throughout the route.

Intrusion Prevention System

An Intrusion Prevention System (IPS) [10] is a network security/threat prevention technology that examines network traffic flows to detect and prevent vulnerability exploits. Vulnerability exploits usually come in the form of malicious inputs to a target application or service that attackers use to interrupt and gain control of an application or machine.

Intrusion Prevention include following activity:

- Sending an alarm to the administrator.
- Dropping the malicious packets.
- Blocking traffic from the source address.
- Resetting the connection.

***Generalized Intrusion Detection and Prevention* (GIDP):**

GIDP is a hybrid IDP approach that uses a combination of anomaly-based and knowledge-based ID. The architecture of GIDP is shown in Fig.1

A cluster head first gathers data in the form of two matrices: network characteristic matrix (NCM) and a derived matrix (DM). The NCM contains data related to the network routingprotocol for example, NCM with seven parameters:

NCM = RREQ (route request), RREP (route reply), RERR (routeerror), TTL (time to live) values, RREQ src seq, RREQ dest seq, RREP dest seq

The DM consists of parameters which reflects the network performance and can be derived from NCM parameters. For example DM with three parameters:

DM = CPO (control packet overhead), PDR (data packet delivery ratio), CPD (number of control packet dropped)

Then the cluster head employs two phases: training and testing. Fig. 2 shows the time-based operation of GIDP.

When the network is established, the CH continuously gathers NCM and DM information and applies the GIDP training module for N time intervals (TI), resulting in initial training profiles (ITPs) of NCM and DM. The ITPs reflects the normal behavior of the nodes in the network and the expected network performance. In the testing phase the CH applies the testing module after each TI. The testing phase consists of several tasks as shown in fig.1 Firstly it detects intrusion in the network. If there is no intrusion then it updates the ITPs in order to adapt the variation in the

© 2015-19, IJARCS All Rights Reserved

network behavior as time progresses. If there is intrusion, in the second task the CH identifies the attack or attacks using existing information in the knowledge base. In

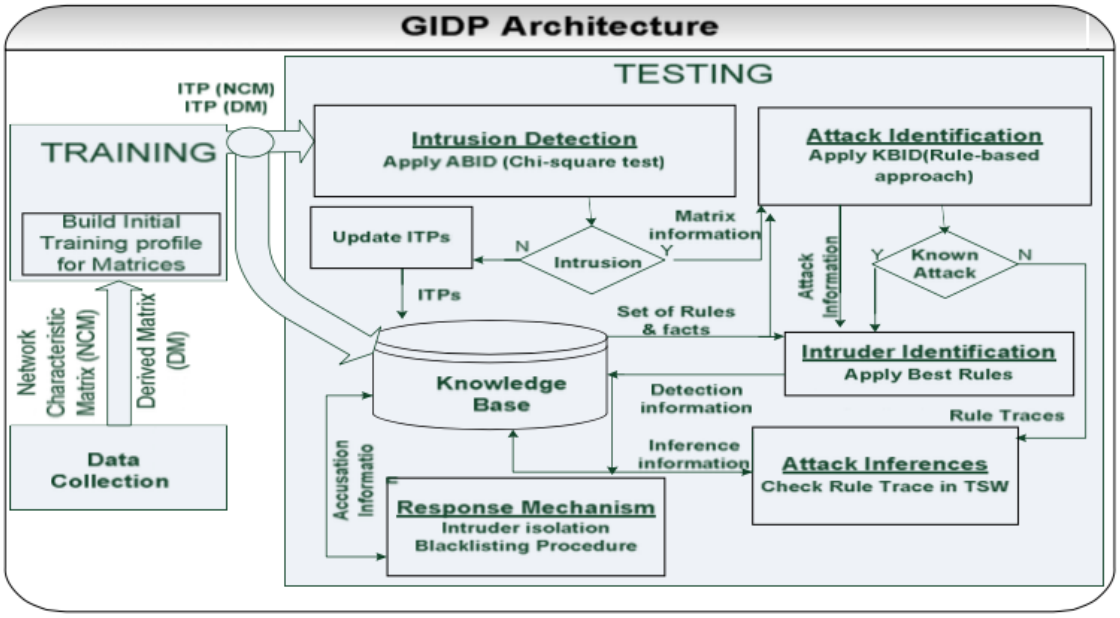

Figure 1: Architecture of GIDP

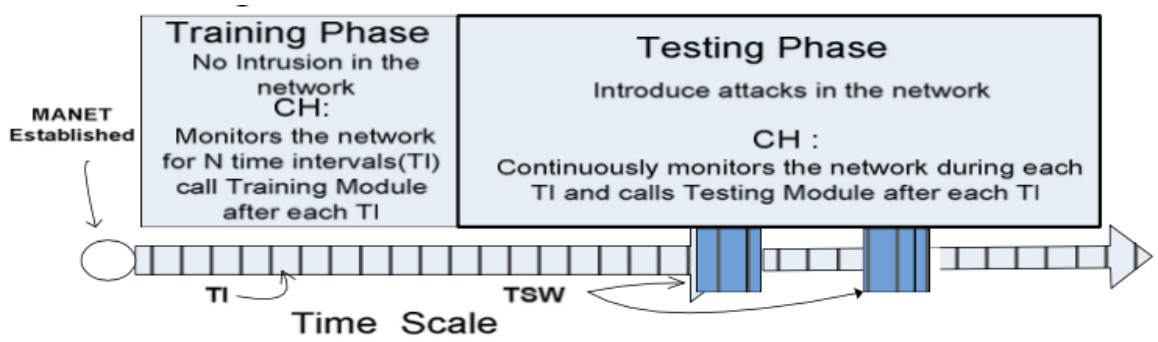

Figure 2: Time based operation of GIDP

the case of known attacks the CH identifies intruding nodes using intruder identification rules specific to the known attack. To optimize the probability of identifying intruders correctly with a low level of false positives, it maintains a test sliding window (TSW) as shown in fig. 2, in which d detections of a node are required in p time intervals (TI). If this detection threshold is passed then the CH will blacklist the node and isolate the node by informing all CNs.

If attack identification detects an attack that does not match the rules for known attack then CH applies the attack inferences. Attack inference stores the rule trace of current TI as Detected Rule Trace and looks for its match in a TSW. If the match is found in a TSW then CH confirms the new attack by constructing adding a rule for the new attack in the set of rules stored in knowledge base.

***Algorithm Technical Details Training* :**

NCM consists of Xi parameters mentioned above, where i=1 to 7 and each Xi =X1, X2, X3,...,XM is a set of random variables from 1 to M, where M is the maximum number of random variables of parameter Xi. For example NCM [Xi] represent the number of RREQ received by all CNs in jth time interval (TI), where M is the maximum number of RREQ received in a TI.

The probability distribution of NCM[Xi] is calculated for the TI. CH then calculates the DM parameters CPO (i.e. Number of control packet / data packet delivered), PDR (i.e. Number of data packet received / data packet originated) CPD (i.e. Number of control packet dropped in establishing maintaining routes in the network) for the jth TI, and this whole process is then repeated for the N time intervals in the training phase. We then calculate mean X i of P(NCM[Xi]) and means of CPO, PDR and CPD for N intervals, which is then stored as an ITP (NCM) and ITP (DM) respectively containing the expected values for that particular network observed for the total time of N*TI seconds.

***Testing* :**

In the testing phase GIDP operates in three stages:

a) Intrusion detection

b) Attack identification and inferences

c) Identification and isolation of intruding nodes

Fig.1 For stage a it employs ABID using chi-square goodness of fit test on NCM and then KBID using a rule-based approach on both matrices NCM DM is applied in stage b and c.

Testing Modules: This module only takes NCM parameters into account and applies chi-square test to identify any intrusion in the network.

a) *Intrusion Detection*

- Do after each TI

– collects NCM(Xi) from all other CNs in TI, for 8i

© 2015-19, IJARCS All Rights Reserved

– calculate the probability distribution P(NCM(Xi))
– calculate average of P(NCM(Xi)) & stores as observed values

- End do
- For 8i Performs Hypothesis Testing by first calculating Chi-computed (X2[i]) for Xi

H_o[i]: Observed distribution of NCM(Xi)fits the expected
H_a[i]: Observed distribution of NCM(Xi)does not fit expected
– If(chi-computed[i](_.d.f[i])) > P-value[i](_.d.f[i])
Reject Ha
– endif

- End for
- Combined Null Hypothesis Testing

Combine H_o: Observed distribution of NCM fits the expected
Combine H_a: Observed distribution of NCM does not fit the expected
– If (combined Ho is rejected)
Perform Attack identification and inferences
– else: Update Expected values NCM(Xi)(i.e ITP(NCM))

- Exit

This module continuously monitors the network. In each TI the CH first performs hypothesis testing for each parameter Xi of NCM at calculated chi-computed values obtain, where Xi is the parameter of NCM and k(1 to M) is the number of random variable in each parameter Xi. The CH then performs combine hypothesis testing of NCM. If the combined Ho is rejected then it assumes intrusion in the TI. Else we update the ITP (NCM) using an exponentially weighted moving average the expected and observed value for update period number(q) respectively. The value of q is incremented in the TI when no intrusion in the MANET is detected. k represents the random variable from 1 to M in each Xi and $\alpha = 2/(q - 1)$ is the weighting factor. As q increases the weighting for older data points decreases exponentially giving more importance to the current observation.

b) ***Attack identification and inferences***

- Reads set of rules
- Set up the interpreter for Rule-based approach
- Interpreter applies Forward-Chaining on set of rules

– If (Any Goal Condition of known attacks are fulfilled)
Apply rules for Intruder Identification and Isolation
– endif
–If(Goal Condition==”POTTENTIALUNKNOWNATTACK”)
Interpreter applies Attack Inference
– endif

- Exit.

Set of rules examples

Rule1:∃*x*(chi-squaretest(NCM[x]))->(checkDerivedMatrix=TRUE)
Rule2:CheckDerivedMatrixΛ∃y(Test(DM[y]))-¿(PotentialAttack=TRUE)
Rule 3 : PotentialAttack->(BestRule=TRUE)
Best Rules for some known attacks.
Rule4:BestRulesΛ(chi-squaretest(NCM[RREQ]))Λ Test(DM[CPO])-> ”SLEEP DEPRIVATION”
Rule 5: BestRules Λ (chi-squaretest(NCM[RREPdestseq]))Λ (Test(DM[PDR])8 lowest(PDR))->”BLACKHONE”
Rule 6 : BestRules Λ (chi-squaretest(NCM[RREPdestseq])) Λ (Test(DM[PDR])->”GREYHOLE”

Attack Inferences

- If(Detected Rule Trace is Empty)
Store Detected Rule Trace=Rule Trace
- Else if(Rule Trace==Detected Rule Trace)
New attack Rule Trace=Rule Trace
Construct a rule for New attack Rule Trace
Append New attack Rule Trace in set of rule trace
Set Detected Rule Trace=Empty
endif
- Endif

c) ***Intruder Identification and Isolation***

a) ***Identifying intruding nodes***

- Obtain known attack Rules for intruder Identification
- For all Goal condition fulfilled
Apply intruder identification rule for each detected known attack add each detected node Vi to List of Nodes Detected(LND)
- end for

b) ***Response Mechanism***

For all nodes Vi in LND

- If(Vi detections in potential Intruder list(PIL)>Detections required to accuse(d))
CH:Blacklist Vi & Broadcast Accusation Packet(AP)
- else: enter Vi in PIL
- endif

End for

c) ***Accusation Packet(AP) Handling***

- Each CN Vi maintain its local BlacklistTable(BLT)
- if CN Vi receives an AP for CN Vi
- If CN Vi has node Vj in its BLT then ignore AP
- Else: CN adds node Vj to its BLT and rebroadcast AP
- endif
- endif

d) ***Isolating Intruding Nodes***

- if node Vi receives packet form node Vi
- if node Vj is in node Vi BLT
Ignore packet and drop all packets queued from Vj
- Else: handle & process packet
- endif
- endif

ZRP BASED MANET BY EFFECTIVE K-MEANS CLUSTERING

ZRP

The Zone Routing Protocol (ZRP) [4] aims to address the problems of both proactive and reactive routing, by combining the best properties of both approaches as shown in fig. 4.1. In ZRP, Intra Zone Routing Protocol (IARP) is the proactive part and the reactive part is the Inter Zone Routing Protocol (IERP).

© 2015-19, IJARCS All Rights Reserved

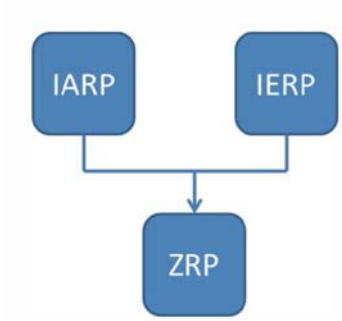

Figure 3: Combine IARP and IERP

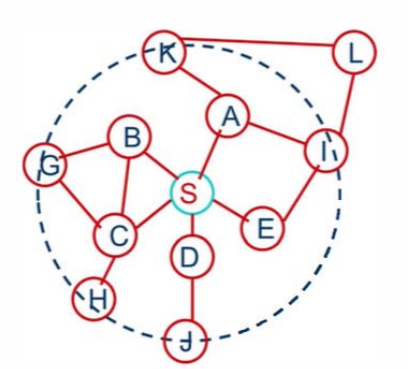

Figure 4: Initial link is established using any link state algorithm

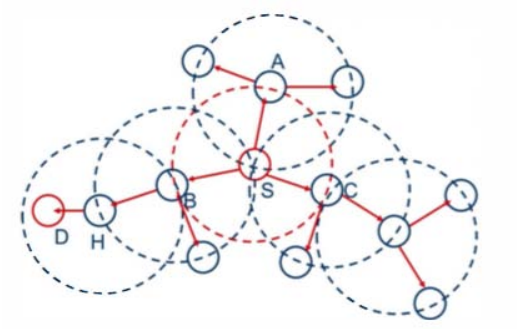

Figure 5: multicast trees

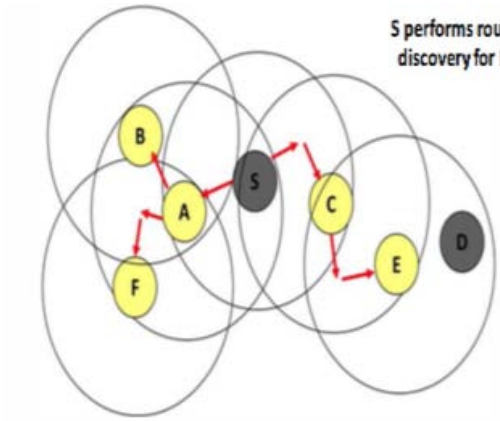

Figure 6: Route discovery

K-Means Clustering

Clustering [4] is a division of data into groups of similar objects. Representing the data by fewer clusters necessarily loses certain fine details, but achieves simplification. It models data by its clusters. From a machine learning perspective clusters correspond to hidden patterns, the search for clusters is unsupervised learning.

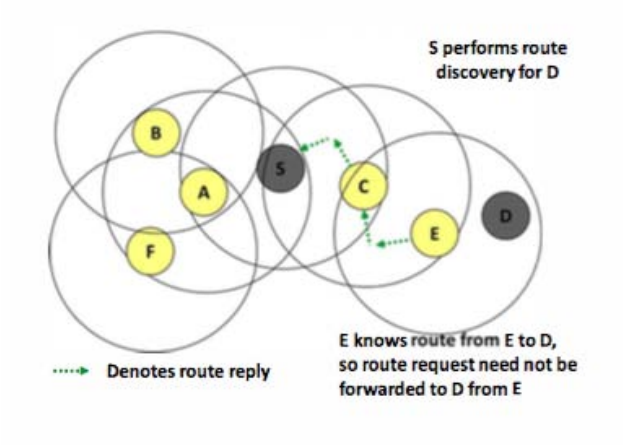

Figure 7: Route Reply

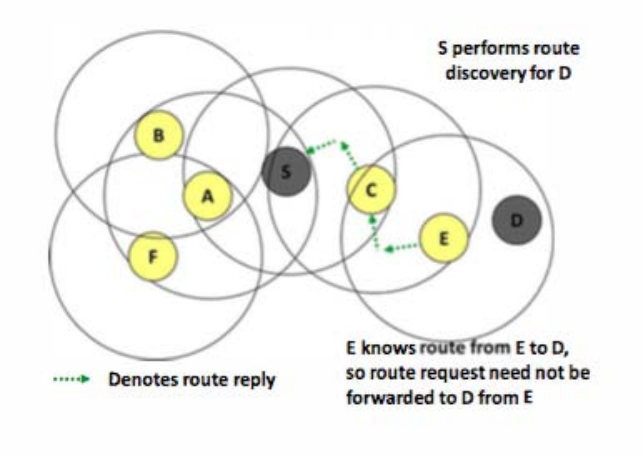

Figure 8: Packet send

Effective K-Means Approach For Intrusion Detection

In our IDS, we have used the new proposed Effective K-means algorithm [4]. The centroids of the clusters are constructed using the algorithm. The desired node features can be picked from the trace file which is obtained on running the simulation in Network Simulator-2. We have assumed the value of K=2 because, we want to obtain two centroids of highly dense segments.

One of these dense segments consists of nodes with normal behaviour and the other consists of abnormal or intrusive behaving nodes. The Effective K-means algorithm provides a data set which is represented by two centroids of highly dense segments. The IDS is host based and monitors every node in the MANET. If any event is generated by a node, then the selected features (as given in Table) of that particular node is fetched. Then the mean square error is calculated and Euclidean distance from the previously constructed centroids is checked. If the result is close to the normal segment centroid then IDS assumes the node to be normal and allows it to proceed with its normal events. Else, it will not allow the node to proceed with its events. The IDS will simply drop any event from the queue, which is generated by the node which has been detected as a malicious node. The above process Will be continued until all the nodes showing intrusive behaviour are detected and separated from the normal nodes.

Thus malicious nodes can be separated from the nodes working properly and as a result, our MANET can again get back to its normal functioning i.e. routing packets properly.

© 2015-19, IJARCS All Rights Reserved

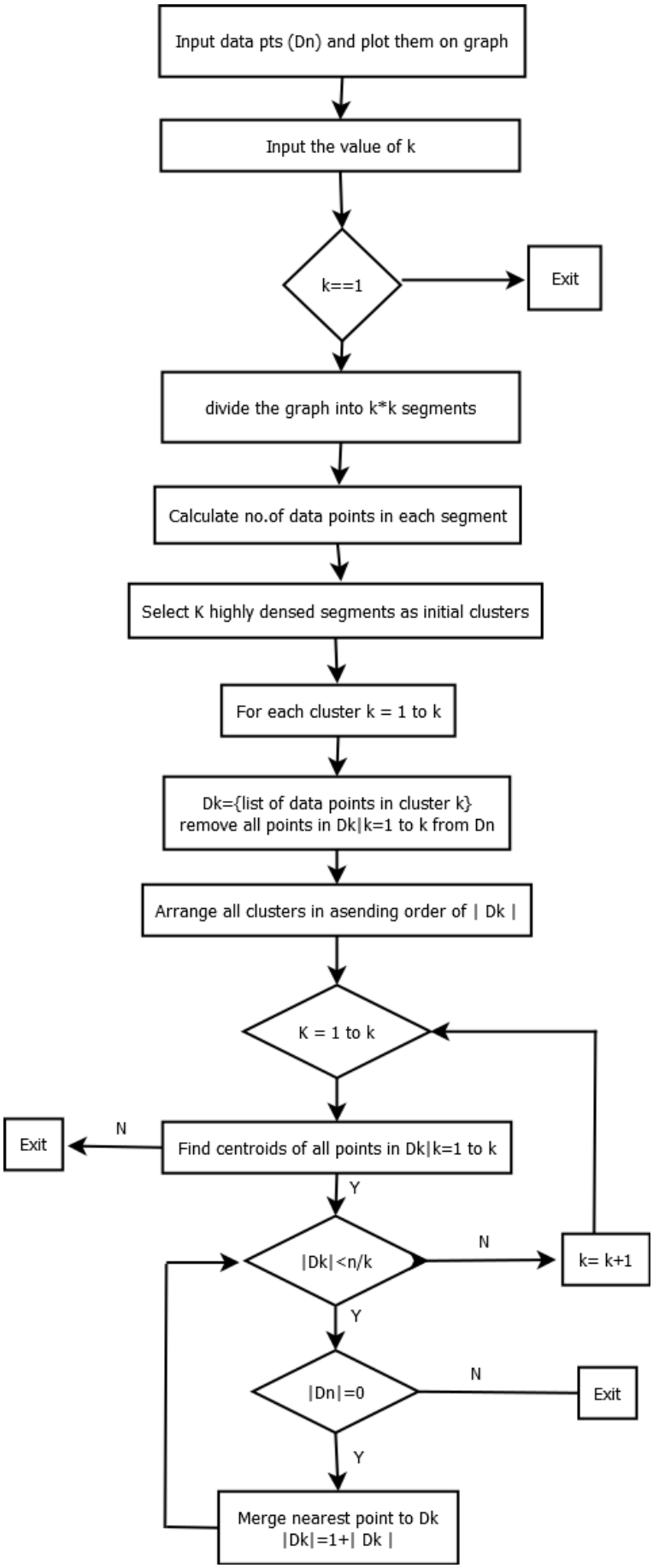

Figure 9: Effective K-means clustering

© 2015-19, IJARCS All Rights Reserved

Table 1: Features of nodes [4]

Sr. No.	Features of Node	Description
1	tREQ	Total no. of route request sent by each node
2	tRREP	Total no. of route reply received by each node
3	tRERR	Total no of route error received by each node
4	tSend	Total no of packets sent by each node
5	tReceive	Total no of packets received by each node
6	tDrop	Total no of packets dropped by each node
7	tForward	Total no of packets forwarded by each node

SMART IDS

Intrusion Detection System

Intrusion Detection System (IDS) is a system which is used to prevent attacks in MANET. Attacks in MANET are classified as active and passive attacks. A passive attack do not disturb the actual functioning of the network whereas the active attack affects the actual functioning of the network as in this unauthorized person tries to extract data being exchanged.

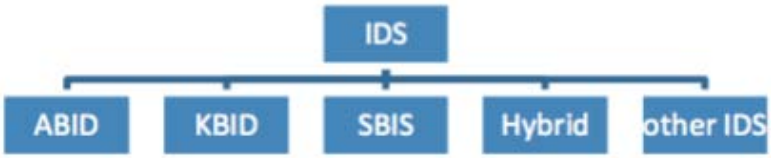

Figure 10: classification of IDS based on ID technique

Anomaly based intrusion Detection : This system is based on the observation of the deviation from the normal traffic. This system basically consists of two parts namely testing and training. Training phase is to capture the knowledge from the existing normal traffic and testing phase is further used to test the present scenario against the anomalies. This method is workable for selected number of attacks and those attacks should occur due to abnormal functioning of the network. It involves training from the operation of non malicious nodes. Good quality training is a key to success of the anomaly based intrusion system.

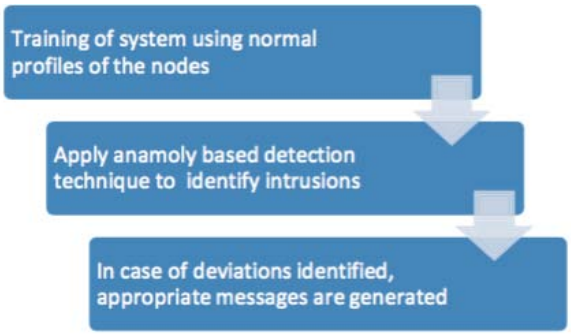

Figure 11: Anomaly based intrusion detection process

Knowledge based intrusion Detection:

It is based on knowledge base which is nothing but a repository where we store information about various attacks. Firstly the system is trained using the training data of the normal nodes and then various rules are inferred from that data and is stored in the knowledge base. This information is stored in the knowledge base is further utilized to detect intrusions. Knowledge base is a repository based on which decision is taken to declare a node as a malicious node. One of the most important part of this system is the decision component which generates appropriate alerts to make the system secure.

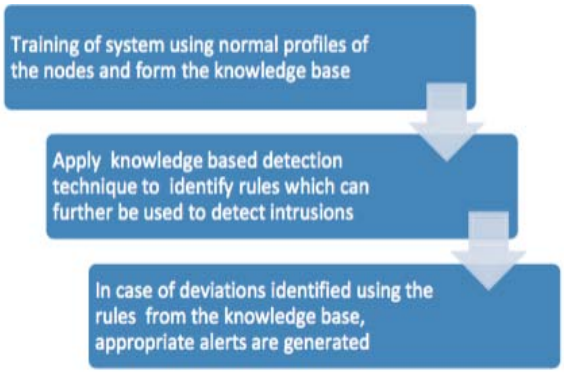

Figure 12: Knowledge based intrusion detection process

Specification based intrusion Detection:

This intrusion detection system is based on the use of specifications. Specifications are the documented form of records of various attacks in the form of constraints. These specifications are used to monitor the unusual transmission of data from source to destination. The main task of SBID is to extract specification based on which intrusions are identified. The specification also provides guidelines to identify normal and correct form of operation of the network. In this we accuracy of specification plays an important role as the whole system relies on the specification to detect intrusion in the network.

So the specifications also needs to be verified and then should be used in the system to detect the intrusions.

© 2015-19, IJARCS All Rights Reserved

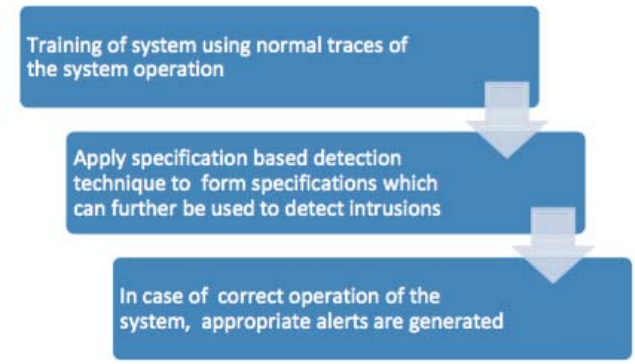

Figure 13: Specification based intrusion detection process

5.2 Classification Of Nodes By Ids

In MANET, while many intrusion detection systems can be used to detect various types of attacks exist, still they suffer from lack of accurate detection of intrusion with respect both to the network size and to the node movement pattern. Intrusion detection systems also lack in performance and reliability when the size of the network increases. It has been a challenging task to design a intrusion detection system that is able to detect vast variety of attacks at various layers.

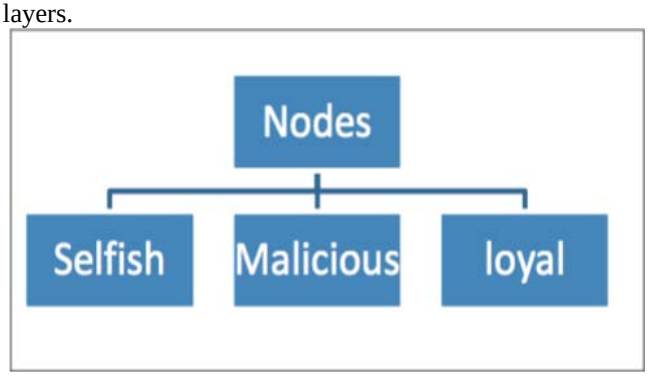

Figure 14: Classification of nodes by IDS

Recently, The Clustering Problem in MANET have been evolved consists of selecting the loyal nodes in a given MANET topology as cluster heads, and ensuring that regular nodes are connected to cluster heads such that the lifetime of the network is maximized.[7] Besides the common performance metrics, other factors can also be used to distinguish loyal nodes from that of selfish and malicious nodes. MANET should also be taken into consideration. Selfish nodes are those nodes which do not forward packets for their self interest. The interest could be to preserve their energy also. On the other hand malicious nodes are those nodes which may or may not pass the packet forward but it can impose threat to the network either by capturing some data or information or message contained in the packet being transmitted. Initially we form a network in MANET consisting of mobile nodes. The nodes comprises of selfish nodes, malicious nodes and normal nodes. The empty circles in the figure denote the loyal nodes which can be treated as loyal nodes. For node communication a hybrid protocol such as ZRP is used. Assuming black hole attack, The IDS divides the network into several clusters and will differentiate the loyal nodes from the malicious nodes and selfish nodes using data mining technique such as clustering. This will result in more secure network.

COMPARISON

We have discussed few architectures of MANET Based IDS in the previous chapter. MANET is a wireless infrastructure where each node travels independently within the zone.
Here we compared K-nearest clustering and smart IDS with parameters like Transmission rate, Detection of attacks, Information authentication, Packet loss, Throughput. Both the systems identifies the intrusion in the system. In k-nearest clustering only outing protocol is used and in smart IDS intrusion detection mechanism is used i.e. Anomaly based intrusion detection, Knowledge based intrusion detection, Specification based intrusion detection.

Table 2: Comparison table

Sr. No.	Parameters	K-nearest clustering[4]	Smart IDS[1]
1	Transmission rate	Here the transmission rate of packet flow is slow because it will transmit loyal as well as suspicious packets	Here the transmission rate of packet is high because it will drop the suspicious packets
2	Detection of attacks	Only applicable of detection of attacks	All known and unknown attacks are detected and discarded by IDS like ABID, KBID, SBID
3	Information authentication	Does not guarantees information authentication	Guarantees information Authentication
4	Packet loss	High	Low
5	Throughput	Low	High

CONCLUSION

Intrusion Detection Systems (IDS) plays an important role in achieving survivability of information system and preserving their safety from attacks. Intrusion detection systems (IDSs) attempt to identify computer system and network intrusions and misuse by gathering and analyzing data. IDSs have traditionally been developed to detect intrusions and misuse for wired systems and networks. More recently, IDSs have been developed for use on wireless networks. These wireless IDSs can monitor and analyze user and system activities, recognize patterns of known attacks, identify abnormal network activity, and detect policy violations for Wireless network. Wireless IDSs gather all local wireless transmissions and generate alerts based either on predefined signatures or on anomalies in the traffic. Here we try to increase the attack detection capability of the intrusion detection technique.

© 2015-19, IJARCS All Rights Reserved

REFERENCES

[1] Syed Zohaib Hussain Zahidi, Fadi Aloul, Member, IEEE, Assim Sagahyroon, and Wassim El-Hajj, "Optimizing Complex Cluster Formation in MANETs Using SAT/ILP Techniques" IEEE Sensors journal, VOL. 13, NO. 6, JUNE 2013.

[2] Peng Zhao, Xinyu Yang, Wei Yu, and Xinwen Fu, "A Loose-Virtual- Clustering-Based Routing for Power Heterogeneous MANETs" IEEE Transactions on vehicular technology, VOL. 62, NO. 5, JUNE 2013

[3] Shengming Jiang, Dajiang He, and Jianqiang Rao, "A Prediction-Based Link Availability Estimation for Routing Metrics in MANETs" IEEE/ACM Transactions on networking, VOL. 13, NO. 6, DECEMBER 2005

[4] Srivastava Sumit ,Deepankar Mitra, Devanshi Gupta, "Proposed Intrusion Detection on ZRP based MANET by Effective K-means Clustering Method of Data Mining" International Conference on Reliability, Optimization and Information Technology –ICROIT 2014, India, pp No. 156-160, Feb 6-8 2014

[5] Lung-Chung Li and Ru-Sheng Liu, "Securing Cluster-Based Ad Hoc Networks with Distributed Authorities" IEEE transactions on wireless communications, VOL. 9, NO. 10, pp-3072-3081, OCTOBER 2010

[6] Nicklas Beijar, "Zone Routing Protocol (ZRP)" (Networking Laboratory, Helsinki University of Technology).

[7] Preetee K. Karmore et al., "Detecting Intrusion on AODV based Mobile Ad Hoc Networks by k- means Clustering method of Data Mining" (UCSIT) International Journal of Computer Science and Information Technologies, Vol. 2(4), 2011,1774-1779.

[8] Jabas, A., R. M. Garimella, and S. Ramachandram, "MANET Mining: Mining Temporal Association Rules" in International Sympossium on Parallel and Distributed Processing with Applications. 2008. Sydney, NSW, Australia.

[9] Adnan Nadeem, Michael P. Howarth, "A Survey of MANET Intrusion Detection & Prevention Approaches for Network Layer Attacks" IEEE communications surveys & tutorials, Vol. 15, No. 4, Fourth quarter 2013

[10] Adnan Nadeem, Michael Howarth, "A Generalized Intrusion Detection & Prevention Mechanism for Securing MANETs" Centre for Communication Systems Research University of Surrey, United Kingdom.

© 2015-19, IJARCS All Rights Reserved