ISSN:0975-9646
Preetee K. Karmore et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 2 (4) , 2011, 1774-1779

Detecting Intrusion on AODV based Mobile Ad Hoc Networks by k-means Clustering method of Data Mining

Preetee K. Karmore , Smita M. Nirkhi
Department of Computer Science & Engineering,
G. H. Raisoni College of Engineering ,
Nagpur, India.

***Abstract*- MANET has no clear line of defense so, it is accessible to both legitimate network nodes and malicious nodes. Some of the nodes may be selfish, for example, by not forwarding the packets to the destination, thereby saving the battery power. Some others may act malicious by launching security attacks like black hole or hack the information. Traditional way of protecting networks with firewalls and encryption software is no longer sufficient. Therefore, intrusion detection system is required that monitor the network, detect malicious node and notifies other node in the network to avoid malicious node i. e. IDS detects malicious activities in the networks. We have implemented k-means clustering algorithm of data mining for efficient detection of intrusions in the MANET traffic and also generated black hole attacks in the network. In data mining, clustering is the most important unsupervised learning process used to find the structures or patterns in a collection of unlabeled data. We have used the K-means algorithm to cluster and analyze the data in this paper. The simulation of the proposed method is performed in NS2 simulator and we got the result as we expected.**

***Keywords*- MANET, cluster, Intrusion Detection System, k-means clustering algorithm, data mining, black hole attack.**

I. INTRODUCTION

An ad-hoc network is a self-configuring network of wireless links connecting mobile nodes. These nodes may be routers and/or hosts. The mobile nodes communicate directly with each other and without the aid of access points, and therefore have no fixed infrastructure. They form an arbitrary topology, where the routers are free to move randomly and arrange themselves as required. Each node or mobile device is equipped with a transmitter and receiver. They are said to be purpose-specific, autonomous and dynamic. This compares greatly with fixed wireless networks, as there is no master slave relationship that exists in a mobile ad-hoc network. Nodes rely on each other to established communication, thus each node acts as a router. Therefore, in a mobile ad-hoc network, a packet can travel from a source to a destination either directly, or through some set of intermediate packet forwarding nodes. Routing protocols between any pair of nodes within an ad hoc network can be difficult because the nodes can move randomly and can also join or leave the network. This means that an optimal route at a certain time may not work seconds later.

Ad-hoc networks are highly vulnerable to security attacks and dealing with this is one of the main challenges of developers of these networks today. The main reasons for this difficulty are; "Shared broadcast radio channel, insecure operating environment, lack of central authority, lack of association among nodes, limited availability of resources, and physical vulnerability." Generally, when considering the security of a network, we examine it under the headings; availability, confidentiality, authentication, integrity and non-repudiation. Availability refers to the fact that the network must remain operational at all times. Confidentiality ensures that certain information is never disclosed to certain users. Authentication is the ability of a node to identify the node with which it is communicating. Integrity guarantees that a message is never corrupted when transferred. Non-repudiation states that the sender of the message cannot deny having sent it. An ad-hoc network has extra security requirements caused by its lack of proper infrastructure and the dynamic relationship between the nodes in the network. Because of the lack of infrastructure, accountability is very difficult to determine as there is "no central authority which can be referenced when it comes to making trust decisions about other parties in the network." Intrusion is defined as “any set of actions that attempts to compromise the integrity, confidentiality or availability of resources”. Intrusion detection systems (IDS) are mainly used to detect and call attention to suspicious behavior [3].

Intrusion detection is used in the networks by comparing the set of baselines of the system with the present behavior of the system [3]. Intrusion detection is one of key techniques behind protecting a network against intruders. An Intrusion Detection System tries to detect and alert on attempted intrusions into a system or network, where an intrusion is considered to be any unauthorized or unwanted activity on that system or network [11]. There are two major analytical techniques in intrusion detection, namely misuse detection and anomaly detection. Misuse detection uses the “signatures” of known attacks [7].

In anomaly detection, profiles of normal behavior of systems, usually established through automated training, are compared with the actual activity of the system to flag any significant deviation. Anomaly detection can detect unknown attacks, but often at the price of a high false alarm rate. We have implemented the k-means clustering algorithm for efficient detection of intrusion in the adhoc networks. The IDS architecture for mobile ad hoc network is shown in fig 1 [23]. The rest of the paper is organized as follows, Section 2 contains related work, Section 3 describes our proposed approach, and Section4 describes the Results and Section 5 describes conclusion.

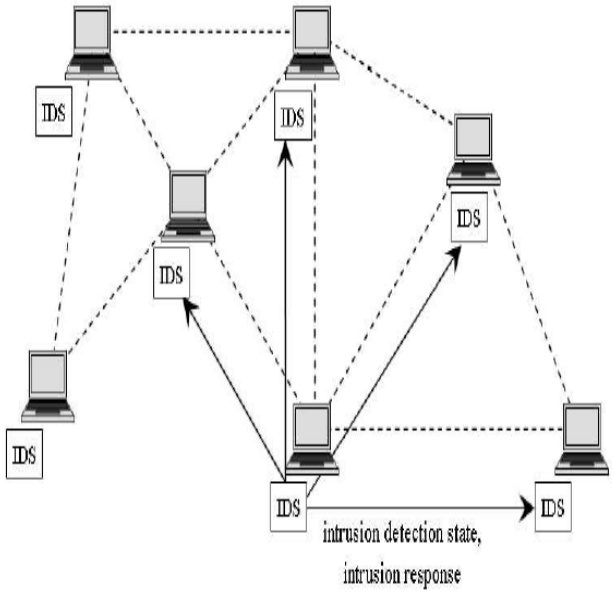

Fig 1: IDS Architecture for Mobile Ad hoc Network.

II. RELATED WORK

The first intrusion detection model was developed in 1987 in which Denning proposed a model based on the hypothesis that security violations can be detected by monitoring a system check records for abnormal patterns of system usage [3]. Many techniques have been discussed to prevent attacks in a wireless adhoc networks as follows. Ricardo Puttini et al [15], propose design and development of the IDS are considered in 3 main stages. A parametrical mixture model is used for behavior modeling from reference data. The associated Bayesian classification leads to the detection algorithm [14].

Yongguang Zhang et al [16], propose new intrusion detection and response mechanisms are developing for wireless ad- hoc networks. Many authors proposed different techniques to prevent attacks in wireless ad hoc networks. But all these methods reported to have a lot of pros and cons of its own proposal. The authors mainly classified their mechanism as signature method, anomaly method. In [19, 20], a work reported by Hu et al., digital signatures approach was used to detect rushing and wormhole attacks. Huang et al.'s reported work can detect type and source of the attack [21].

In [22], author reported a survey on intrusion and its detection measures. In this paper author explained vulnerability of ad hoc networks, then explained routing protocols in MANET then explained different types of attacks in ad hoc networks and at last security schemes in mobile ad hoc networks. In this security schemes they reported 4 approaches to detect intrusions. These four approaches are: Profile Based Intrusion Detection Approach, the Enhancement of Intrusion Detection System for Ad Hoc Network (EIDAN), Agent Based Efficient Anomaly Intrusion Detection System in ad hoc networks, and Intrusion Detection based on K-means clustering. In this paper, data mining method have been used to provide solution against security issues in MANET networks.

III. PROPOSED METHOD

Initially, mobile ad-hoc network is created, which consist of number of normal nodes and malicious nodes. For node communication AODV routing protocol has been used. Black hole attack has been implemented to denote malicious node in the network which will be act as intruder in the MANET. Our IDS will detect that malicious node in the network and avoid the effect of it. So that performance of the network will be improved. After detecting the node as malicious, the sequence number field has been decremented and hop count field has been incremented, so that the event (i. e. send, receive, forward, drop) generated by that malicious node will not be processed in any ways. In this way we have detected an intrusion in ad hoc network and improved the performance of the network. Clustering based Data mining technique have been employed, for intrusion detection in our approach which can improve variant detection rate, control false alarm rate and reduce false dismissals.

A. Data Mining Technique to detect intrusion

Data mining is the process of analyzing data from different perspectives and summarizing it into useful information. Data mining can be divided into four types: association analysis, sequence analysis, classification analysis and cluster analysis. Classification algorithm about Data Mining can be used to construct classifier, after the invasion of a large number of data sets being trained. Classifier can be used for intrusion detection [12].

Clustering analysis algorithm can be used to construct the network model of normal behavior, or intrusion behavior model. Association analysis algorithm can be used to describe the invasion of behavior patterns of association rules, through these rules intrusion detection can come [12]. We have used clustering method of data mining to detect an intrusion so that security of MANET will be improved.

Clustering is the method of grouping objects into meaningful subclasses so that the members from the same clusters are quite similar, and the members from different clusters are quite different from each other [13]. Therefore clustering methods can be useful for classifying log data and detecting intrusions.

Clustering algorithms can be categorized into four main groups: Partitioning algorithm, hierarchical algorithm, density-based algorithm and Grid-based algorithm. We have used Partitioning algorithm. Partitioning algorithms construct a partition of a Database of N objects into a set of K clusters [13]. K-means clustering is the partitioning method of data mining. We have used this k-means algorithm for constructing the clusters of data. Brief description of this algorithm is given below.

1. K-Means Approach for Intrusion Detection

In proposed system, k-means algorithm is used to construct the centroids of clusters. The features of nodes are given as input to the k-means algorithm which are shown table I. These features are selected from the trace file which is generated by running the simulation. We have taken k=2, because we have to construct two clusters out of these two clusters one cluster consist of normal behavior of node and other consist of intrusive behavior or abnormal behavior of node. The flow diagram of this k-means algorithm is shown in figure 2.The

result of k-means algorithm is clustered data set that is represented by 2 centroids of clusters. Figure 3 shows proposed flowchart for detection of intrusion. Our proposed IDS system is present at every node. That is our IDS system is host based which monitor each and every node in the network whether any node in the network generates any events or not. If any, then the features of that node is extracted, calculates the mean square error and then check Euclidean distance from the centroids which have been constructed previously. If it is nearest to normal cluster centroid then IDS will assume that the node is normal and it will allow to proceed its events normally, if it is nearest to abnormal or malicious cluster centroid then it will not allow to proceed, that is our IDS will drop an event from the queue which is generated by malicious node. In this way we have detected malicious nodes in mobile ad-hoc networks and avoided the effect of it. We have used data mining technique in order to improve the efficiency and effectiveness of the mobile ad-hoc network nodes.

TABLE I: FEATURES OF NODE

Sr. no	Features	Description
1	tRREQ	Total no. of Route Requests 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	tSent	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.
8	tAvgSendTime	Average time required for sending packets by each node.
9	tAvgRecvTime	Average time required for receiving packets by each node.
10	tAvgDropTime	Average time required for dropping the packets by each node.
11	tAvgForwardTime	Average time required for forwarding the packets by each node.

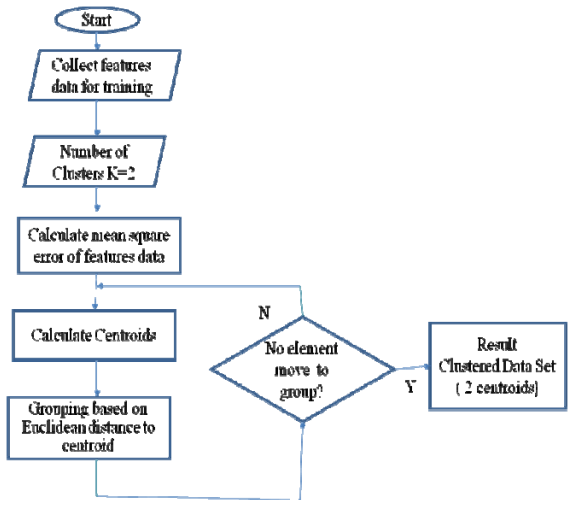

Fig 2: Flowchart of K-means clustering algorithm for Intrusion Detection

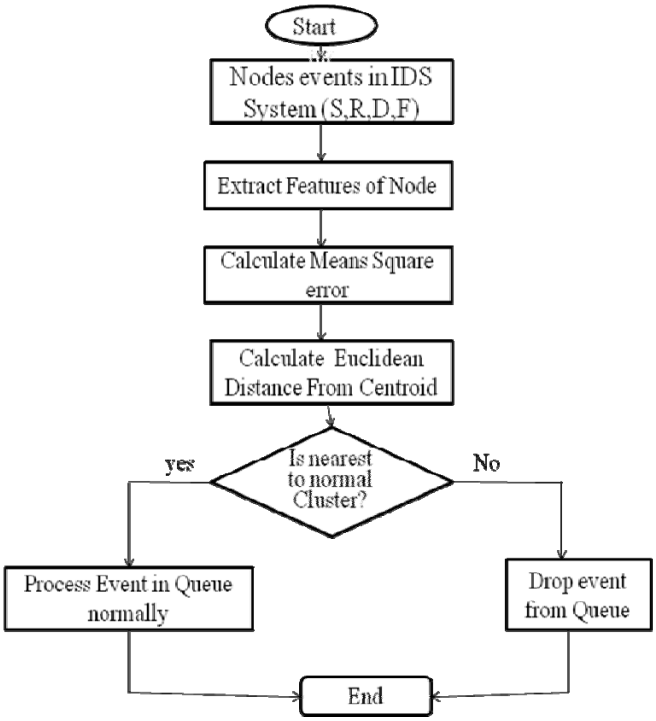

Fig 3: Proposed flowchart for Detection of Intrusion

IV. EXPERIMENTAL RESULTS

A. *Simulation Setup*

The simulation is carried out in NS-2 simulator installed on Linux machine. The experimental set up consists of 20 wireless nodes. All the nodes use AODV as a routing protocol within the area of 500m x500m. AODV protocol is a suitable approach for mobile networks due to low message overhead. The simulation is run for 500 seconds. The simulation statistics is shown in table II. We have used packet size of 512 bytes which starts sending packets at around 0.20 sec. We have used UDP traffic as underlying transport protocols. UDP connections are established between even numbered nodes (zero included) and odd numbered nodes.

In the scenarios, even numbered nodes (Node 0 - Node 16) are the sending nodes and odd numbered nodes (Node 1 - Node 17) are the receiving nodes and the even numbered nodes send the packets to the next odd numbered nodes, for example Node 0 to Node 1, Node 2 to Node 3, Node 4 to Node 5 etc. Thus, we could count the sent and received packets between any 2 nodes. In the scenarios, UDP agents are attached to the even numbered nodes and NULL agents are attached to odd numbered nodes. During simulation UDP data traffic is sent in bytes/sec by the source node to the destination node. We attach the CBR (Constant Bit Rate) application that generates constant packets through the UDP connection. We have used black hole attack on the normal traffic in this scenario.

We run two different simulations, one with the black hole attack scenario, in which attacker node captures packets sent by source node and do not forward anywhere, which is shown in figure 5 and figure 6 and other scenario is intrusion detection scenario, in which we have avoided the effect of black hole attack.

In detection scenario, the packets transmitted by source node arrive properly to the destination node. That is there is no packet capturing by the attacker node. The existing node model is modified at its routing protocol for detecting intrusive activities because the IDS check the behavior of the nodes while sending or receiving the data.

TABLE II: SIMULATION STATISTICS

Parameters	Values
Simulation Time	500 Seconds
Simulation Area	500m * 500m
Traffic Type	CBR (UDP)
Packet Size	512 bytes
Number of nodes	20 (18 normal, 2 malicious)

B. Evaluation of results

We have used AODV routing protocol in 20 mobile nodes. For evaluation purpose, we mostly consider source, destination and attacker node whereas other nodes assist in routing of the packets and have their own purpose. The basic node scenario is shown in figure 4. In figure 5 and figure 6 node 18 and node 19 are malicious or black hole nodes. Here the source node 6 wants to send data packets to the destination node 7, so it will initiate route discovery process by sending route request packets in the network.

Whenever malicious node receives route request packet it will immediately send false route reply message with highest destination sequence number and minimum hop count to the source node 6. When the route reply reaches first to the source node, the source node assumes that the route discovery process is completed and it will avoid the reply from other nodes. Source begins to sends all the data packets to the malicious node. So malicious node, consumes all the data packets passed by source node. Similarly, node 19 consumes all the data packets sent by the source node 0.

Figure 7 and figure 8 shows the simulation results generated after detecting black hole attack and avoiding the effect of it through our implemented IDS detection technique. When the simulation is compiled, we saw that sending node is sending the messages to receiving node properly. Figure 7 shows that CBR packets are reaching from source node 0 to the destination node 1 as expected. With our intrusion detection technique, we have nullified the effect of intrusion. So even if malicious node 19 present nearest to source node 0, it is not able to capture the packets passing through its neighbor. Similarly, figure 8 shows that CBR packets are reaching from source node 6 to the destination node 7 as expected.

In this way, we have detected an intrusion in ad hoc network and avoided its effect in the network. As we have nullified the effect of black hole in the network, the performance of the network is improved. Figure 9, 10 and 11 shows the graphs which are generated by implementing our solution for intrusion in ad hoc network. The graphs show that we have improved the packet loss, delay, and throughput of the network.

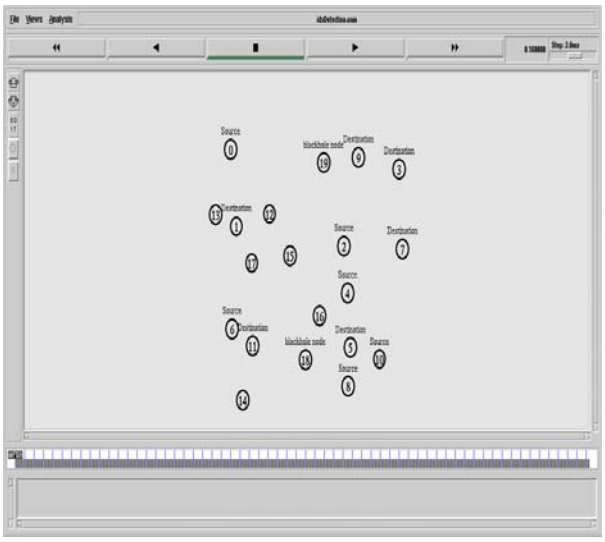

Figure 4: Simulation scenario

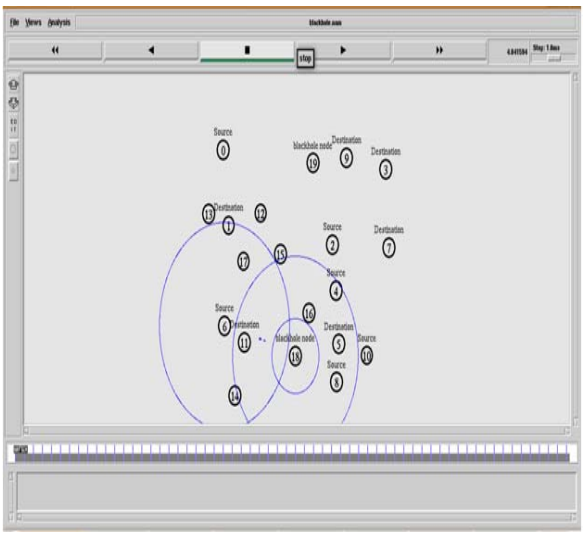

Fig 5: Node 18 (Black Hole Node) absorbs the connection Node 6 to Node 7

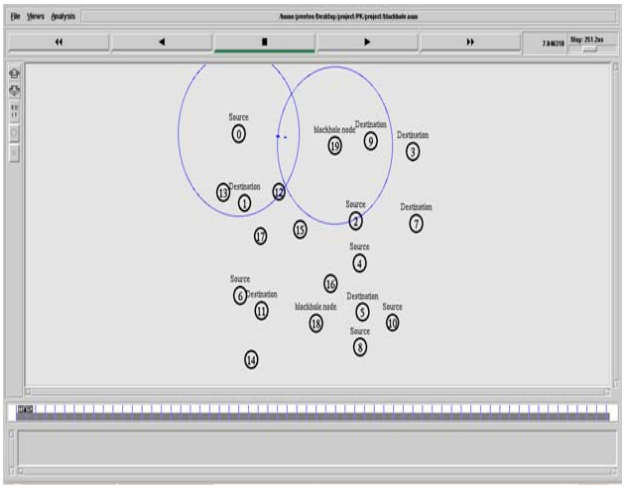

Fig 6: Node 19 (Black Hole Node) absorbs the connection Node 0 to Node 1

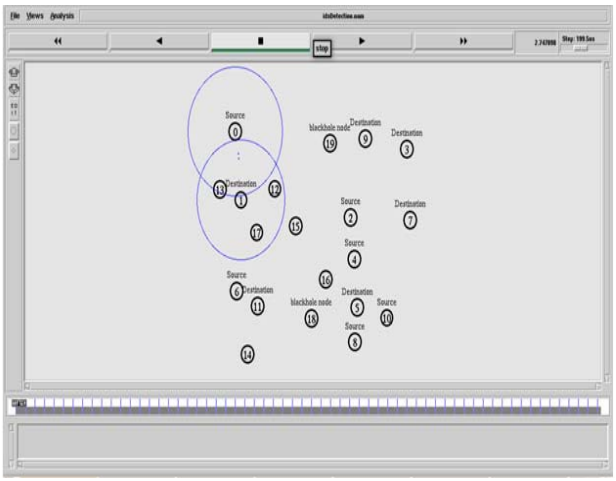

Fig 7: CBR packets are reached from source node 0 to destination node 1 properly

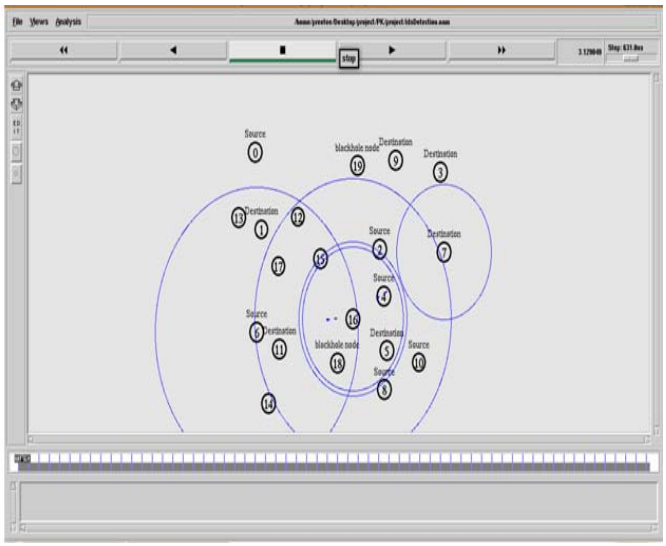

Fig 8: CBR packets are reached from source node 6 to destination node 7 properly

Fig 9 shows the packet loss graph. Red colored graph shows the average packet loss in % when there is attack in the network and green colored graph shows improved average packet loss after applying our intrusion detection technique. The percentage of packet loss is reduced with our technique. Fig 10 shows the throughput graph. Red colored graph shows the throughput when there is attack in the network and green colored graph shows the improved throughput after applying our intrusion detection technique. The throughput is increasing as the number of normal increases with our technique. Increase in throughput improves the network performance. Fig 11 shows the Delay graph. Red colored graph shows the average Delay in milliseconds when there is attack in the network and green colored graph shows the improved average Delay in milliseconds after applying our intrusion detection technique. The packet delay is reduced with our technique. In this way, we have improved three network parameters namely: throughput, packet delay and packet loss. By improving these three parameters of the network, we enhanced the network performance.

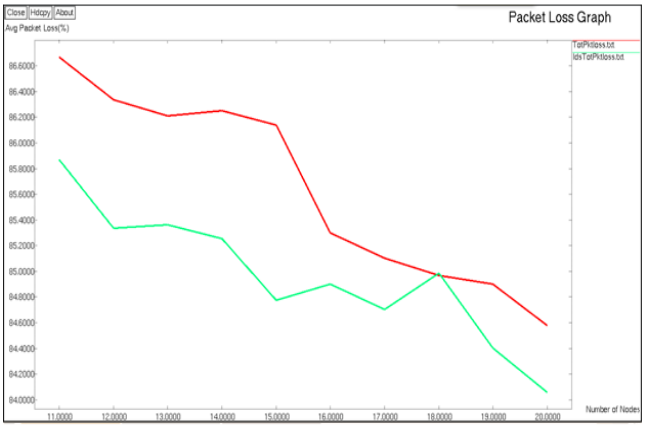

Fig 9: Packet loss graph

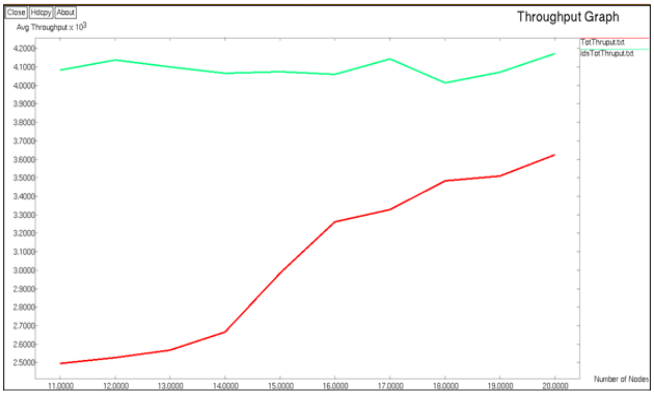

Fig 10: Throughput graph

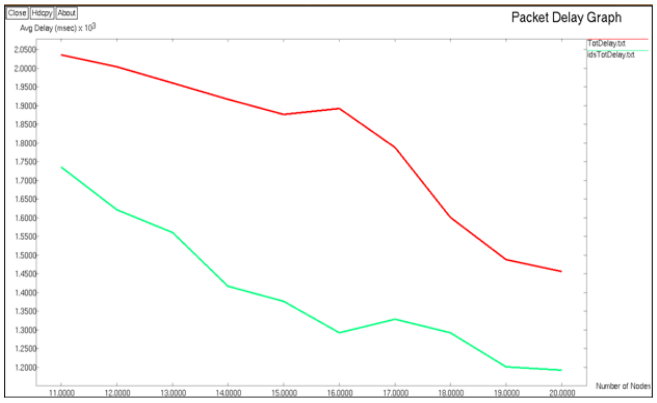

Fig 11: Delay Graph

V. CONCLUSION

Intrusion detection is a hot field of the network security research, and it is a new kind of defense technology of the network security. Hence, a better intrusion detection mechanism is presented in this paper utilizing cluster data mining technique. We have implemented the proposed architecture with K-means clustering algorithm and done the Simulation and analyzed the result. Our proposed intrusion detection architecture is designed to detect black hole attack. The aim is to improve the detection rate and decrease the false alarm rate.

REFERENCES

[1] R. Nakkeeran, T. Aruldoss Albert and R. Ezumalai, "Agent BasedEfficient Anomaly Intrusion Detection System in Adhoc Networks", IACSIT International Journal of Engineering and Technology Vol. 2, No.1, February, 2010.

[2] Yan Yu,"A Novel Intrusion Detection Approaches Based on Data Mining", 978-1-4244-6349-7/10/$26.00 c_2010 IEEE.

[3] L. Prema Rajeswari, R. Arockia Xavier Annie, A. Kannan, "ENHANCED INTRUSION DETECTION TECHNIQUES FOR MOBILE AD HOC NETWORKS", IET-UK International Conference on Information and Communication Technology in Electrical Sciences (ICTES 2007), Dec. 20-22, 2007. Pp.1008-101.

[4] Peyman Kabiri and Mehran Aghaei, "Feature Analysis for Intrusion Detection in Mobile Ad-hoc Networks", International Journal of Network Security, Vol.12, No.2, PP.80–87.

[5] B. Sun, K. Wu, and U. Pooch, "Routing Anomaly Detection in Mobile Ad Hoc Networks", Proceedings of the 12th IEEE Int'lConf. On Computer Communications and Networks (ICCCN'03), Dallas, TX, Oct. 2003.

[6] Baolin Sun, Hua Chen, Layuan Li, "An Intrusion Detection System for AODV", Proceedings of the 10th IEEE International Conference 2005, pp. 358-365.

[7] Y. Huang, W. Fan, W. Lee, and P. S. Yu, "Cross-Feature Analysis for Detecting Ad-Hoc Routing Anomalies", Proceedings of the 23rd IEEE International Conference on Distributed Computing Systems, 2003, pp. 478-487.

[8] Abhay Kumar Rai, Rajiv Ranjan Tewari & Saurabh Kant Upadhyay, "Different Types of Attacks on Integrated MANET-Internet Communication", International Journal of Computer Science and Security (IJCSS) Volume (4): Issue (3).

[9] S. Madhavi, "An Intrusion Detection System in Mobile Adhoc Networks", 2008 International Conference on Information Security and Assurance, 978-0-7695-3126-7/08 $25.00 © 2008 IEEE DOI 10.1109/ISA.2008.80.

[10] Peyman Kabiri and Mehran Aghaei, "Feature Analysis for Intrusion Detection in Mobile Ad-hoc Networks", International Journal of Network Security, Vol.12, No.2, PP.80–8.

[11] Oleg Kachirski, Ratan Guha, "Intrusion Detection Using Mobile Agents in Wireless Ad Hoc Networks", Proceedings of the IEEE Workshop on Knowledge Media Networking (KMN'02) 0-7695-1778-1/02 $17.00 © 2002 IEEE.

[12] Li Bo and Jiang Dong-Dong, "The Research of Intrusion Detection Model Based on Clustering Analysis", 2009 International Conference on Computer and Communications Security, 2009 IEEE.

[13] Meng Jianliang Shang Haikun Bian Ling, "The Application on Intrusion Detection Based on K-means Cluster Algorithm", 2009 International Forum on Information Technology and Applications, 978-0-7695-3600-2/09 $25.00 © 2009 IEEE DOI 10.1109/IFITA.2009.34.

[14] Hang Yu Yang, Li-Xia Xie, 'Agent based Intrusion Detection for a Wireless Local Area Network', Proceedings of the IEEE third International Conference on Machine Learning and Cybermatics, 2004, pp. 2640-2643.

[15] Ricardo Puttini, Maíra Hanashiro, Javier García-Villalba, C. JBarenco, "On the Anomaly Intrusion-Detection in Mobile Ad Hoc Network Environments", Personal Wireless Communications ,Volume 4217/2006, Springerlink, September 30, 2006.

[16] Yongguang Zhang, Wenke Lee, "Intrusion detection in wireless ad-hoc networks", Pages: 275 - 283 Year of Publication: 2000 ISBN: 1-58113-197-6, ACM, 2000.

[17] C. Siva Ram Murthy, B.S. Manoj, "Ad Hoc Wireless Networks Architectures and Protocols", Pearson Education.

[18] Ms. Preetee K. Karmore, Ms. Smita M. Nirkhi, "Data Mining: A Novel Approach for Intrusion detection in Ad hoc Networks", ICNCC International Conference on Network Communication and Computer, march 2011.

[19] Y. C. Hu, A. Perrig, and D. B. Johnson, "Rushing attacks and defense in wireless ad-hoc network routing protocols," Proceedings of the 2nd ACM workshop on Wireless security, pp. 30-40, USA, 2003.

[20] Y. C. Hu, A. Perrig, and D. B. Johnson, "Wormhole attacks in wireless networks," IEEE Journal on Selected Areas in Communications, vol. 24, no. 2, pp. 370-380, IEEE, 2006.

[21] Y. A. Huang, and W. Lee, "A Cooperative Intrusion Detection System for Ad Hoc Networks," Proceedings of the ACM Workshop on Security of Ad Hoc and Sensor Networks (SASN' 03), pp. 35-147, USA, 2003.

[22] Preetee Karmore, Sonali Bodkhe, "A Survey on Intrusion in Ad Hoc Networks and its Detection Measures", International Journal on Computer Science and Engineering (IJCSE), Chennai, India, 2011.

[23] Yongguang Zhang, Wenke Lee," Intrusion Detection in Wireless AdHoc Networks", Proceedings of the Sixth Annual International Conference on Mobile Computing and Networking (MobiCom'2000), August 6–11, 2000.