ISSN 1512-1232

Statistical Traffic Pattern Discovery System For Wireless Mobile Networks

C.Karthika [1] ,Dr.M.Sreedhar[2]

[1]Master of Engineering, Embedded System Technologies ,Velalar College of Engineering and Technology
Erode, India karthika.c8@gmail.com
[2]Associate professor, EEE Dept Velalar College of Engineering and Technology Erode, India
callsreedhar@gmail.com

Abstract

This paper deals with discovering of raw traffic data for the given network area by using sensor nodes based on statistical characteristics. This is used to analyses the communication anonymity of mobile adhoc networks based on packet decryption. It is capable of calculating the probability of source/destination nodes and end to end communication relationship. In existing statistical traffic analysis method, is fails to discover sensitive information from the statistical characteristics of the network. It does not provide a method to identify the actual source and destination nodes. In the proposed method there are two approaches are used namely time slicing technique used to build point to point traffic matrices, and heuristic approach used to identify the actual source and destination nodes. By using these techniques, the hidden traffic patterns can be discovered in good accuracy and traffic delay can be restricted.

Keywords: *traffic matrix; point-to-point matrix; statistical traffic pattern*

1. INTRODUCTION

Military organizations and other similar organizations are concerned about the security of information transfers, have always heavily relied on secure exchanges of messages. Mobile network is a new wireless networking pattern for mobile hosts. Unlike conventional mobile wireless networks, ad hoc networks do not rely on any fixed infrastructure. Instead, hosts rely on each other to keep the network as be linked. The military tactical and other security-sensitive operations are the main applications of Mobile networks, although there is a tendency to adopt ad hoc networks for commercial uses due to their distinctive properties. Ad hoc networks are a new pattern of wireless communication for mobile hosts.

In a MANETs, there is no fixed infrastructure such as base stations or mobile switching centers. Mobile nodes are communicate directly within each other's radio range via wireless links, while those that are far apart rely on other nodes to convey messages as routers. Node mobility in an ad hoc network causes numerous changes of the network topology. As per T. He et al (2008), anonymity in MANETs includes identification of the location anonymity of sources, destinations as well as route. Same anonymity of sources and destinations means it is difficult to possible for other nodes to identify the correct locations of the sources and destinations. Anonymity routing protocols have two types of hop by hop encryption and redundant traffic. In an network, there are three wireless nodes are present i.e., node1, node2, and node3. Node 2 is located in the transmission range of node 1, and node 3 is located in the transmission range of node 2 but not the transmission range of node 1.In that two consecutive packets are detected i.e., node1 broadcasts a packet and then node 2 broadcasts a packet simultaneously.

M. Reed et al (2002) introduces that, hop by hop routing encryption can be divided into onion routing and hop-by-hop authentication. In onion routing, packets are encrypted in the source node and decrypted layer by layer (i.e., hop by hop) along the routing path. The MANETs have several prominent characteristics such as Dynamic topologies, Bandwidth-constrained, variable capacity

ISSN 1512-1232

links, Energy-constrained operation, Limited physical Security. Due to these features, mobile ad hoc networks are particularly vulnerable to denial of service attacks launched through compromised nodes.

M. Wright et al (2004) Evidence based statistical traffic analysis model particularly for MANETs. In this model, every captured packet is taken as evidence based point-to-point (one-hop) transmission between the sender and the receiver. To derive end-to-end (multihop) relations, the sequences of point-to-point traffic matrices are used. This approach provides a realistic attacking framework against MANETs but still leaves significant information about the communication patterns as undiscovered. Initially, this scheme fails to deal with several essential constrains (e.g., maximum hop-count of a packet) when deriving the end-to-end traffic from the point-to-point evidences. Then, it does not provide a technique to identify the real source and destination nodes. Moreover, it only uses a accumulative traffic ratio to infer the end-to-end communication relations which incurs a lot of inaccuracy in the derived probability distributions of source and destinations.

This paper aims to obtain the source/destination probability distribution, i.e., the probability of each node, and the end-to-end link probability distribution, i.e., the probability for each couple of nodes to be an end-to-end communication pair. To achieve its goals it includes two key steps: 1)Build point-to-point traffic matrices using the time-slicing technique, and then derive the end-to-end traffic matrix with help of traffic filtering rules; and 2) Apply a heuristic approach to identify the real source and destination nodes, and then compare the source nodes with their corresponding destination nodes.

2. RELATED WORK

As per J. Raymond (2001), Traffic analysis models have been widely investigated for static wired networks. For example, the attack to track a data is to list all possible links a message could traverse, i.e., the brute force attack. In recent times, statistical traffic analysis attacks have fascinated broad interests due to their submissive nature, i.e., attacker's needs only to collect information and perform analysis without changing the network activities. As per Jojy Saramma John,and R.Rajesh (2014) introduced node flushing attack the attacker sends a huge amount of messages to the particular anonymous system. Since most of the messages modified and reordered by the attacker, the attacker can track the remaining few messages. K.P.Manikandan et al (2014) presented surveys of attacking systems, in that the Routing Table Overflow Attack (RTOA) attack is basically happens to proactive routing algorithms, which update routing information periodically. It can simply transmit excessive route advertisements to overflow the target system's routing table. In Routing Table Poisoning (RTP) method the compromised nodes in the networks send fictitious routing updates or modify real route update packets sent to other uncompromised nodes. RTP may result in suboptimal routing, congestion in portions of the network, or even make some parts of the network inaccessible. In this Packet Replication attack, an adversary node replicates stale packets and it consumes additional bandwidth and battery power resources available to the nodes and also causes unnecessary confusion in the routing process. C. E. Perkins et al (2000) introduced Rushing Attack has On-demand routing protocols(ODRP) that uses replica suppression during the route discovery process are vulnerable to this attack. In the Ad-hoc On-demand Distance Vector (AODV) the attacker may advertise a route with a smaller distance metric than the original distance and also advertise a routing update with a large sequence number and invalidate all routing updates from other nodes. Dynamic Source Routing (DSR) protocol is alike to AODV .That is also forms route on-demand but the main difference is that it uses source routing instead of relying on the routing table at each intermediate node.

Some advanced attacks are Wormhole attack, Black hole attack, Rushing Attack, Byzantine attack, Resource Consumption Attack and the Location disclosure attack..Many forms of the attacks mentioned above, but statistical traffic analysis is different from that and it intends to discover

ISSN 1512-1232

responsive information from the statistical characteristics of the network traffic, for example, the traffic volume, maximum hop counts, number of message packet dropping etc.,. The adversaries typically do not change the network behavior (such as injecting or modifying packets). The only thing is to quietly collect traffic information and perform statistical calculations.

Due to the exceptional characteristics of MANETs, very restricted investigation has been conducted on traffic analysis in the circumstance of MANETs. He et al (2008) proposed a timing-based approach is to trace down the possible destinations for given known source nodes. In this technique, assuming the transmission delays are bounded at each relay node and they estimate the flow rates of transmission paths using packet matching. Based on the estimated flow rates, a set of nodes that partitioning the network into two parts, one part is to which the source can communicate in adequate rate and the other to which it cannot that are identified to estimate the possible destinations. In Liu et al (2010) intended a traffic inference algorithm (TIA) for MANETs based on the supposition that the difference between data frames, routing frames, and MAC control frames .It is visible to the passive adversaries, so that they can be familiar with the point-to-point traffic using the MAC control frames, identify the end-to-end flows by tracing the routing frames, and then deduce the actual traffic pattern using the data frames. The TIA achieves good accuracy in traffic inference, while the process firmly fixed to a particular anonymous routing protocol but it is not a common approach for traffic pattern discovery.

3. TRAFFIC ANALYSIS MODEL

A mobile ad hoc network is protected by secrecy enhancing techniques such that all information flows are encrypted. The traffic analyzer cannot decrypt the information flows, and neither can they disclose the multihop communication relations from the routing layer and above. However, the adversaries capture every packet transmitted in the network using location tracking systems, at any given time. It aims to deanonymize the network communications on a per-flow basis.

The attacks on MANETs challenge the mobile infrastructure in which nodes can join and leave easily with dynamics requests without a static path of routing. Schematics of various attacks as described on individual layer are as under:

- Application Layer: Malicious code, Repudiation
- Transport Layer: Session hijacking, Flooding
- Network Layer: Sybil, Flooding, Black Hole, Grey Hole. Worm Hole, Link Spoofing, Link Withholding, Location disclosure etc.
- Data Link/MAC: Malicious Behavior, Selfish Behavior, Active, Passive, Internal External.
- Physical: Interference, Traffic Jamming, Eavesdropping.

3.1 Brute Force Attack

Brute force attack is very instructive due to it helps to determining how much, when and where to use the dummy traffic. These dummy messages are sent through the network that makes difficult the adversary's process. These adversary in a setting in which each mix node waits until it receives messages before flushing them. In addition, assume each message goes through exactly to mix nodes.

3.2 Node Flushing Attacks

The flush attack is very effectual and mounted by an active external attacker. If the nodes wait until they have t messages before "flushing", an adversary can send t messages and easily associate t-1 messages leaving the node with those having entered. It can be noted that the observer will be able to match his inputs with the messages leaving from the node. The observer cannot discriminate

ISSN 1512-1232

the dummy traffic from valid message. Authenticate each message and detect flushing attempts could be computationally infeasible. But these attacks are only suitable for wired networks.

3.3 Timing Attacks

The system could be susceptible to timing attacks because different routes can be taken different amounts of time. When it uses mix nodes the attack is very effective. Messages passed through the mix nodes with a variable amount of time, before flushing messages. This attack can be carried out by passive observer. This attack is only suitable for wired networks.

These are the existing attacks for wired networks. The proposed system is for wireless networks. The Statistical Disclosure Traffic Pattern Discovery System uses the heuristic approach that is used to find out the hidden traffic pattern in MANETs. It performs the traffic pattern analysis based on statistical characteristics of captured raw traffic. This method is used for the passive observer observes the real source and destination nodes, and then correlates the source to their equivalent destination. It reused the evidence-based model and then derived the source/destination probability distribution and multi-hop probability distribution used to find the traffic pattern. All previous methods are used for limited attack, they cannot detect both the source and destination at the same time for any given network. This attacking system detects all the source and destinations, also traces their link between them.

4.TRAFFIC PATTERN DISCOVERY MODEL

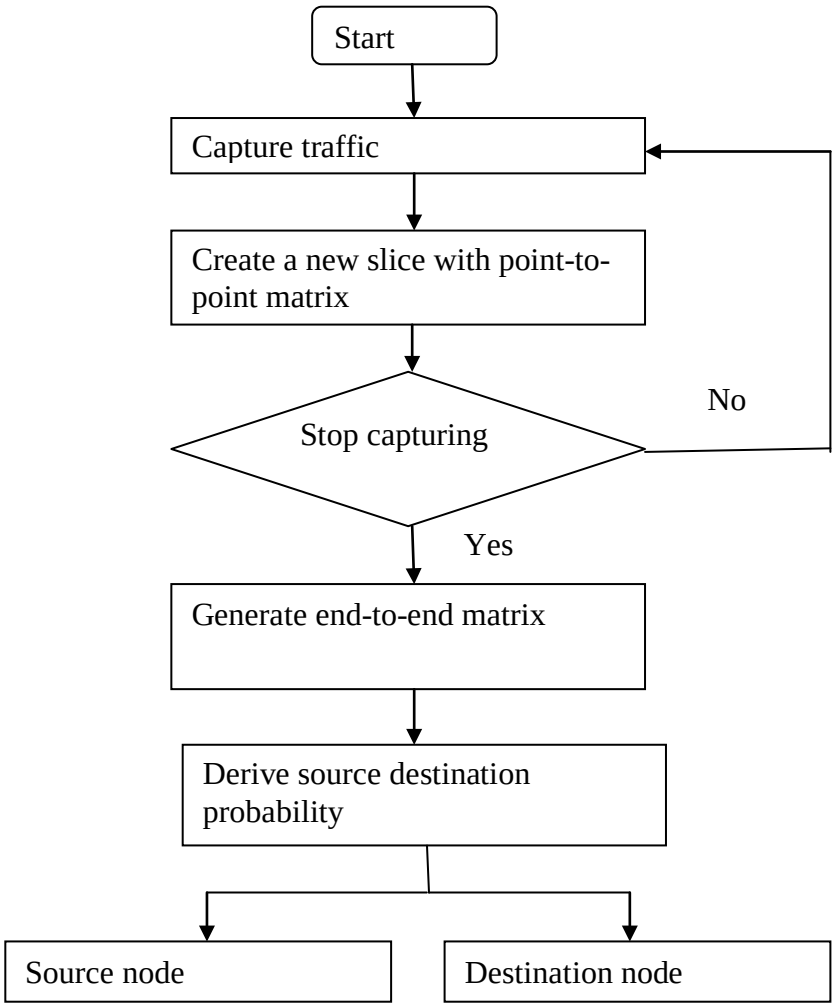


Fig. 1: Work flow model

ISSN 1512-1232

To discover the hidden traffic patterns in a MANET system, this model includes two steps. Initially, it captures the raw traffics to construct point-to-point traffic matrices and then derives the end-to-end traffic matrices. Second it calculates the probability for each node to be a source/destination, for further analyzing the end-to-end traffic matrices. From that each pair of source and destination nodes of end-to-end probability distribution can be obtained.

4.1. Traffic Matrix Construction

Initially it needs to build point-to-point traffic matrices such that each traffic matrix only contains independent one-hop packets. If two packets captured at different time could be the same packet appearing at different locations, such as the two packets sent by node 1 and node 2 one after the other, so they are dependent on each other. To avoid a single point-to-point traffic matrix, apply a “time slicing” technique for two dependent packets arrival. From the sequence of point-to-point traffic matrices our goal is to derive the end-to-end traffic matrices. It includes both the point-to-point traffic captured directly and multihop traffic deduced from the point-to-point traffic.

4.2. Probability Distribution

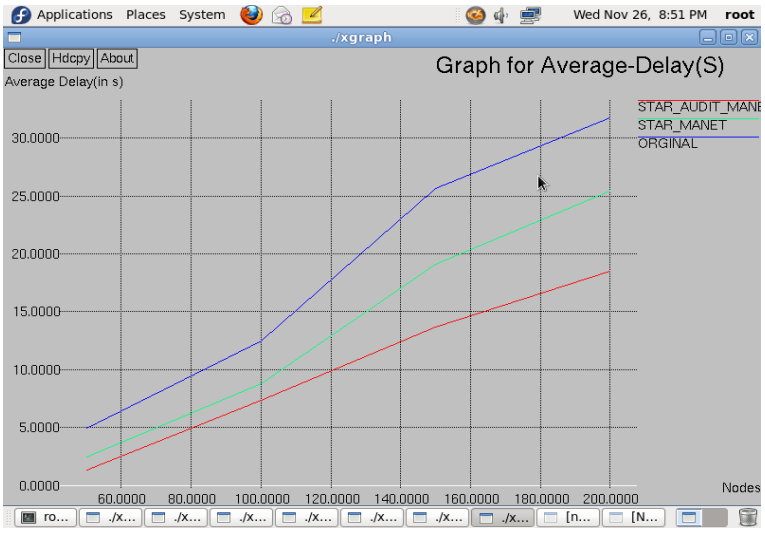
To identify the source and destination by calculating the source/destination probability distribution. Source probability distribution and destination probability distributions are derived from the capturing raw traffic data. It needs the algorithms for finding the actual source and their corresponding destinations. By introducing the vector space similarity assessment, can ensure that, two nodes with higher probability to be neighbors have less impact on each other’s source/destination probability distribution, which reasonably reduces the neighborhood noise. To reduce the neighborhood noise, can utilizes the vector space similarity assessment. The vector space similarity (or cosine similarity) of two vectors V and U is defined as follows:

$$Sim(V,U)=V.U/(|V||U|)$$

where V and U denotes the dot product of V, and U, |V|, and |U| denote the norm of V and U. If two nodes have similar outgoing and incoming traffic vectors (in the end-to-end traffic matrix), they are likely to be neighboring nodes (relays of each other), and so they should have less impact on the source/destination probability distribution of each other.

5. DEMONSTRATION

The network environment is simulated using NS2. A network simulator is software that predicts the behavior of a computer network. In simulators, the computer network is typically modeled with devices, links, applications etc. and the performance is analyzed.



ISSN 1512-1232

Fig.2: Graph for average delay

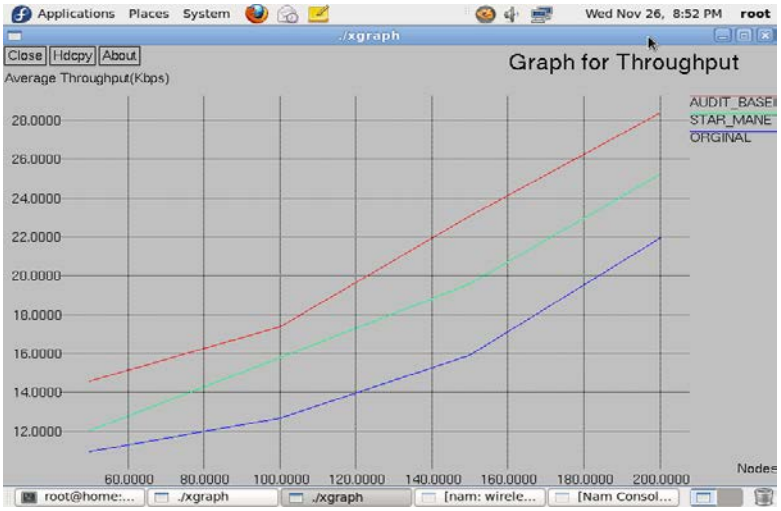


Fig.3: Graph for throughput

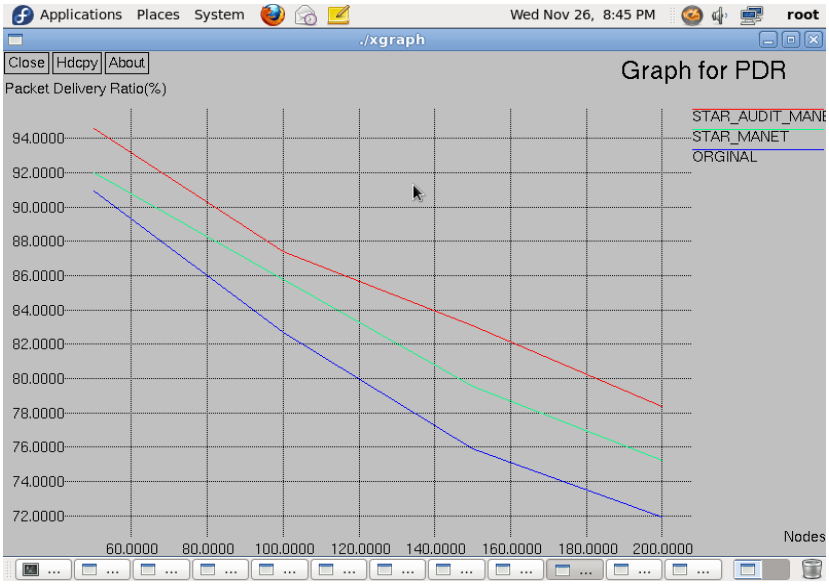


Fig.4: Graph for packet delivery ratio.

These graphs were simulated for average delay,false positive rate,message drops,throughput,and packet delivery ratio.Graphs are depends upon the number of nodes used for simulation.In this all graphs, star audit manet is the simulated output for this experiment.Star manet and original graphs were compared to this experiment. While comparing these three graphs our audit manet gives the best result.Number of message drops were reduced by this experiment.So performance and throughput of this model has raised.

6. DISCUSSION AND FUTURE WORK

The adversaries can globally monitor the traffic across the entire network region. This assumption is conventional from the network users' point of view. Usually, it is difficult for the attackers to perform such a global traffic detection. However, even though the adversaries are not able to monitor the entire network, they can monitor several parts of the network simultaneously.

ISSN 1512-1232

For example, an attacker can deploy sensors (signal detectors) around some particular mobile nodes to track their movements and eavesdrop all of their traffic. These sensors may even move accordingly. With the restricted capabilities, the attacker can take advantage of STARS to perform traffic analysis.

To perform Generalized statistical traffic pattern discovery system, the adversaries only need to monitor the nodes beside the boundaries of the supernodes. The traffic inside each supernode can be ignored, since it will not affect the inter-region traffic patterns. In addition, Generalized traffic pattern does not need the signal detectors to be able to precisely locate the signal source. They are only required to determine which supernode (region) the signals are sent from. Moreover, in statistical traffic pattern system, the actual receiver of a point-to-point transmission is not identifiable among all the potential receivers within the sender's transmitting range. This inaccuracy can be mitigated in Generalized traffic pattern, because most potential receivers of a packet will be contained within one or a few supernodes. Genaralized traffic pattern discovery system will be the direction of our future research.

7. CONCLUSION

This method is basically an attacking system, which only needs to capture the raw traffic from the PHY/MAC layer without looking into the contents of the intercepted packets. From the captured packets, it constructs a sequence of point-to-point traffic matrices to derive the end-to-end traffic matrix, and then uses a heuristic data processing model to reveal the hidden traffic patterns from the end-to-end matrix. It demonstrates that the existing MANET systems can achieve very restricted communication anonymity under the attacks. It achieves the higher throughput, reduced delay, good packet delivery ratio and less overhead.

REFERENCES

[1] Yang Qin, Dijiang Huang, Senior Member, IEEE, and Bing Li"STARS: A Statistical Traffic Pattern DiscoverySystem for MANETs", 2013
[2] D. Chaum, "Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms," Comm. ACM, vol. 24, no. 2, pp. 84- 88, 1981.
[3] Qinghua Wang, "Traffic Analysis & Modeling in Wireless Sensor Networks and Their Applications on Network Optimization and Anomaly Detection," Network Protocols and Algorithms ISSN 1943- 3581 2010, Vol. 2, No.1.
[4] Jojy Saramma John, R.Rajesh, "Efficient Anonymous Routing Protocols in Manets: A Survey," International Journal of Computer Trends and Technology (IJCTT) – volume 11 number 1 May 2014.
[5] G. Danezis, "Statistical Disclosure Attacks: Traffic Confirmation in Open Environments," Proc. Security and Privacy in the Age of Uncertainty (SEC '03), vol. 122, pp. 421-426, 2003.
[6] T. He, H. Wong, and K. Lee, "Traffic Analysis in Anonymous MANETs," Proc. Military Comm. Conf. (MILCOM '08), pp. 1-7,2008.
[7] M. Wright, M. Adler, B. Levine, and C. Shields, "The Predecessor Attack: An Analysis of a Threat to Anonymous Communications Systems," ACM Trans. Information and System Security, vol. 7, no. 4, pp. 489-522, 2004.
[8] M. Reed, P. Syverson, and D. Goldschlag, "Anonymous Connections and Onion Routing," IEEE J. Selected Areas in Comm., vol. 16, no. 4, pp. 482-494, May 2002.

ISSN 1512-1232

[9] Raymond, "Traffic Analysis: Protocols, Attacks, Design Issues, and Open Problems," Proc. Int'l Workshop Designing Privacy Enhancing Technologies: Design Issues in Anonymity and Unobservability, pp. 10-29, 2001.
[10] X. Wang, S. Chen, and S. Jajodia, "Network Flow Watermarking Attack on Low-Latency Anonymous Communication Systems," Proc. IEEE Symp. Security and Privacy, pp. 116-130, 2007.
[11] D. Figueiredo, P. Nain, and D. Towsley, "On the Analysis of the Predecessor Attack on Anonymity Systems," technical report, Computer Science, pp. 04-65, 2004.
[12] G.Danezisand A.Serjantov, "Statistical Disclosure or Intersection Attacks on Anonymity Systems," Proc.SixthInformation Hiding Workshop (IH '04), pp. 293-308, 2004.
[13] G. Danezis, C. Diaz, and C. Troncoso, "Two-Sided Statistical Disclosure Attack," Proc. Seventh Int'l Conf. Privacy Enhancing Technologies, pp. 30-44, 2007.
[14] C. Troncoso, B. Gierlichs, B. Preneel, and I. Verbauwhede, "Perfect Matching Disclosure Attacks," Proc. Eighth Int'l Symp. Privacy Enhancing Technologies, pp. 2-23, 2008.
[15] Y. Liu, R. Zhang, J. Shi, and Y. Zhang, "Traffic Inference in Anonymous MANETs," Proc. IEEE Seventh Ann. Comm. Soc. Conf. Sensor Mesh and Ad Hoc Comm. and Networks (SECON ' 10), pp. 1-9,2010.
[16] J. Raymond, "Traffic Analysis: Protocols, Attacks, Design Issues,and Open Problems," Proc. Int'l Workshop Designing Privacy Enhancing Technologies: Design Issues in Anonymity and Unobservability, pp. 10-29, 2001.
[17] Huang, D. (2008): Unlinkability Measure for IEEE 802.11 Based MANETs. IEEE Trans. Wireless Comm., vol. 7, no. 3, pp. 1025-1034.
[18] C. E. Perkins and P. Bhagwat, "Highly dynamic destination sequenced distance vector routing (DSDV) for mobile computers," Proceedings of the ACM SIGCOMM '94 Conference, pp. 234– 244, Aug. 1994.
[19] C. E. Perkins, E. R. Royer, and S. R. Das, "Ad hoc on-demand distance vector (AODV) routing," IETF MANET Working Group,Internet-Draft, Mar. 2000.
[20] K.P.Manikandan, Dr.R.Satyaprasad, and Dr.K.Rajasekhararao," A Survey on Attacks and Defense Metrics of Routing Mechanism in Mobile Ad hoc Networks," in *(*IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 2, No.3, March 2011.

Article received: 2015-04-15