AN ENHANCED GLOBEL STATISTICAL PATTERN DISCOVERY SYSTEM FOR MANET

Dr.K.Prabha[1], K.Nirmaladevi[2]

[1]Assistant Professor, Computer Science & Periyar University PG Extension Centre, Dharmapuri, Tamilnadu, India
[2]Ph.D Research Scholar, Computer Science & Periyar University PG Extension Centre, Dharmapuri, Tamilnadu, India
nirmaladevi0291@gmail.com

Abstract — MANET routing relies on techniques such as re-encryption on each hop to hide end-to-end communication relations. However, passive signal detectors and traffic analyzers can still retrieve sensitive information from PHY and MAC layers through statistical traffic analysis. In this paper, propose a Statistical Traffic Pattern Discovery System (STPDS). STPDS intends to find out the sources and destinations of captured packets and discover the end-to-end communication relations. The proposed approach does not require analysers to be actively involved in MANET transmissions or to decrypt the traffic. Specifically, it proposes a local estimation model to approximate the readings of sensor nodes in subsets, and prove rated error-bounds of data collection using this model. In the process of model-based data collection and formulate the problem of selecting the minimum subset of sensor nodes into a minimum dominating set problem which is known to be NP-hard, and propose a greedy heuristic algorithm to find an approximate solution. It further proposes a monitoring algorithm to adaptively adjust the composition of node subsets according to changes of sensor readings.

***Keywords*— STPDS, NS2, Node Reality, Node Lifetime.**

I. INTRODUCTION

Mobile Ad-Hoc Networks (MANETs) has become one of the most prevalent areas of research in the recent years because of the challenges it pose to the related protocols. MANET is the new emerging technology which enables users to communicate without any physical infrastructure regardless of their geographical location, that's why it is sometimes referred to as an-infrastructure less network. The proliferation of cheaper, small and more powerful devices make MANET a fastest growing network. An ad-hoc network is self-organizing and adaptive.

Device in mobile ad hoc network should be able to detect the presence of other devices and perform necessary set up to facilitate communication and sharing of data and service. Ad hoc networking allows the devices to maintain connections to the network as well as easily adding and removing devices to and from the network. Due to nodal mobility, the network topology may change rapidly and unpredictably over time. The network is decentralized, where network organization and message delivery must be executed by the nodes themselves. Message routing is a problem in a decentralize environment where the topology fluctuates. While the shortest path from a source to a destination based on a given cost function in a static network is usually the optimal route, this concept is difficult to extend in MANET.

Table-driven (proactive) routing: This type of protocols maintains fresh lists of destinations and their routes by periodically distributing routing tables throughout the network. The main disadvantages of such algorithms are: Respective amount of data for maintenance and slow reaction on restructuring and failures. On-demand (reactive) routing: This type of protocol finds a route on demand by flooding the network with

Route Request packets. The main disadvantages of such algorithms are: High latency time in route finding and excessive flooding can lead to network clogging.

Network Simulator (NS2):

NS2 is simply an event driven simulation tools that has proved useful in the dynamic nature of communication networks. Simulation of wired as well as wireless network functions and protocols legacy, routing algorithms, TCP and UDP.

NS2 Components:

TCL is open script language which is used to program NS2. NAM – Network Animator, its consists of visual demonstration of NS of output. Its declared the pre-processing and post analysis. Trace analysis using PERL/ MATLAB.

Hybrid (both proactive and reactive) routing: his type of protocol combines the advantages of proactive and reactive routing. The routing is initially established with some proactively prospected routes and then serves the demand from additionally activated nodes through reactive flooding. The choice of one or the other method requires predetermination for typical cases.

Hierarchical routing protocols: With this type of protocol the choice of proactive and of reactive routing depends on the hierarchic level in which a node resides. The routing is initially established with some proactively prospected routes and then serves the demand from additionally activated nodes through reactive flooding on the lower levels. The choice for one or the other method requires proper attribution for respective levels. The main disadvantages of such algorithms are: Advantage depends on depth of nesting and addressing scheme. Reaction to traffic demand depends on meshing parameters.

A node discards a route request packet that it has already seen. The route request packet uses sequence numbers to ensure that the routes are loop free and to make sure that if the intermediate nodes reply to route requests, they reply with the latest information only. When a node forwards a route request packet to its neighbors, it also records in its tables the node from which the first copy of the request came. This information is used to construct the reverse path for the route reply packet. AODV uses only symmetric links because the route reply packet follows the reverse path of route request packet. As the route reply packet traverses back to the source, the nodes along the path enter the forward route into their tables. The Wireless Routing Protocol (WRP): The Wireless Routing Protocol (WRP) is a table-based distance-vector routing protocol. Each node in the network maintains a Distance table, a Routing table, a Link-Cost table and a Message Retransmission list.

II. APPLICATIONS OF MANET

Collaborative Work:

For some business scenarios, the need for collaborative computing might be more important outside office environments than inside a building. After all, it is often the case where people do need to have outside meetings to cooperate and exchange information on a given project;

Crisis-Management Applications:

These arise, for example, as a result of natural disasters where the entire communications infrastructure is in disarray (for example, Tsunamis, hurricanes etc.). Restoring communications quickly is essential. By using ad hoc networks, an infrastructure could be set up in hours instead of days/weeks required for wire-line communications;

Vehicular Area Network:

An ad hoc network is specifically useful in forming networks among different vehicles on the road and can propagate information like accidents, congestion etc. It can also help determining close by facilities in the neighborhood such as gas station, restaurants, hospitals, and other facilities. Personal Area Networking – A Personal Area Network (PAN) is a short-range, localized network where nodes are usually associated with a given person. These nodes could be attached to someone's cell phone, laptop, television set, and so on.In these scenarios, mobility is only a major consideration when interaction among several PANs is necessary, illustrating the case where, for instance, people meet in real life. In the occasion where there is a group effort required, the MANET plays a major role in wireless communication and provides effective communication. At the time of disaster, it is easy to develop a wireless network rather than a wired network. The places where wired network may be affected by the disasters, MANET can be implemented. As far as the Personal Area Networks (PAN) is concerned, they need not much coverage. They just need the coverage of very limited area. MANETs server the purpose in such situations.

Data Monitoring and Mining:

MANETs can be used for facilitating the collection of sensor data for data mining for a variety of applications such as air pollution monitoring and different types of architectures can be used for such applications. It should be noted that a key characteristic of such applications is that nearby sensor nodes monitoring an environmental feature typically register values. This kind of data redundancy due to the spatial correlation between sensor observations inspires the techniques for in-network data aggregation and mining. By measuring the spatial correlation between data sampled by different sensors, a wide class of specialized algorithms can be developed to develop more efficient spatial data mining algorithms as well as more efficient routing strategies.

The Distance table of a node x contains the distance of each destination node y via each neighbor z of x. It also contains the downstream neighbor of z through which this path is realized. The Routing table of node x contains the distance of each destination node y from node x, the predecessor and the successor of node x on this path. It also contains a tag to identify if the entry is a simple path, a loop or invalid. Storing predecessor and successor in the table is beneficial in detecting loops and avoiding counting-to-infinity problems. The Link-Cost table contains cost of link to each neighbor of the node and the number of timeouts since an error-free message was received from that neighbor.

III. ANONYMOUS ROUTING PROTOCOL FOR MANET

MANETs are more vulnerable to both active and passive attacks. Wireless transmissions are easy to capture remotely and undetected, while the lack of central management and monitoring make network nodes susceptible to active attacks. Providing security for MANETs is a challenging task, and many researchers have engaged in designing protocols for diverse security related task such as key management, authentication, confidentiality, etc. Recently researchers have also tackled the problem of anonymity in wireless networks. It is clear that providing anonymity in ad hoc networks is important as users may wish to hide the fact that they are accessing some service or communicating with another user. Another application is hiding the location of users participating in the network. Hiding nodes that participate in the network also makes it more difficult for an adversary to focus his attack as he will not be able to identify and locate the more active nodes within the network.

RAM: Code privacy can be protected by using a tamper resistant cryptographic processor. The protocol is such that an outside party looking at the memory accesses (reads and writes) can't gain any information about what is being computed and how it is being computed. The code's privacy is protected which could be useful to prevent reverse engineering and software piracy.

A. Traffic Analysis

Traffic analysis attacks against the static wired networks (e.g., Internet) have been well investigated. The brute force attack in existing methodologies tries to track a message by enumerating all possible links a message could traverse. In node flushing attacks the attacker sends a large quantity of messages to the targeted anonymous system (which is called a mix-net). Since most of the messages modified and reordered by the system are generated by the attacker, the attacker can track the rest a few (normal) messages. The timing attacks as proposed in existing studies focuses on the delay on each communication path. If the attacker can monitor the latency of each path, it can correlate the messages coming in and out of the system by analyzing their transmission latencies. The message tagging attacks require attackers to occupy at least one node that works as a router in the communication path so that they can tag some of the forwarded messages for traffic analysis. By recognizing the tags in latter transmission hops, attackers can track the traffic flow. The watermarking attacks are actually variants of the message tagging attacks. They reveal the end-to-end communication relations by purposely introducing latency to selected packets.

B. Network Communication using Manet

MANET protected by anonymity enhancing techniques, it is a difficult task itself to identify an actual destination node as the target due to the ad hoc nature. That is, destinations are indistinguishable from other nodes (e.g., relays) in a MANET. In fact, they usually act as relay nodes as well, forwarding traffic for others. The adversaries are not able to determine whether a particular node is a destination depending on whether the node sends out traffic. This is totally different from the situation in traditional infrastructural networks where the role of every node is determined.

A statistical disclosure attack often targets a particular given source node and intends to expose its corresponding destinations. It is assumed that the packets initiated by the source are sent to several destinations with certain probability distribution. The background traffic also has certain probability distribution. After a large number of observations, the attackers are able to figure out the possible destinations of the given source. Nonetheless, the statistical disclosure attacks cannot be applied to MANETs either, because the attackers cannot easily identify the actual source nodes in MANETs.

TABLE I
PARAMETERS OF SIMULATION

Parameters	Value

Area	100*1000 m
No. of nodes	5 to 50
No. of repetition	5 times
Physical/Mac layer	IEEE 802.11g
Pause time	30 sec
Mobility model	Random direction model
Node movement	5 – 35 m/s
Data sending rate	2 Mbps

STPDS as the Generalized STPDS (GSTPDS). To perform GSTPDS, the adversaries only need to monitor the nodes beside the boundaries of the super nodes. The traffic inside each super node can be ignored, since it will not affect the inter-region traffic patterns. In addition, GSTPDS does not need the signal detectors to be able to precisely locate the signal source. They are only required to determine which super node (region) the signals are sent from. Moreover, in STPDS, 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 GSTPDS because most potential receivers of a packet will be contained within one or a few super nodes.

C. Derive the accumulative traffic matrix R

Given a sequence of point-to-point traffic matrices W|1 * K, their goal is to derive the end-to-end traffic matrix R = (r(i, j) N*N where r(i, j) is the accumulative traffic volume from node i to node j, including both the point-to-point traffic captured directly and multi-hop traffic deduced from the point-to-point traffic.

Input: f (W|1 * K)
1: R =W1
2: for e = 1 to K - 1 do
3: R = g(R+1) + +1
4: end for
5: return R.

D. Figures and Tables

A node can be either a sender or a receiver within this time interval. But it cannot be both. 2) Each traffic matrix must correctly represent the one-hop transmissions during the corresponding time interval.

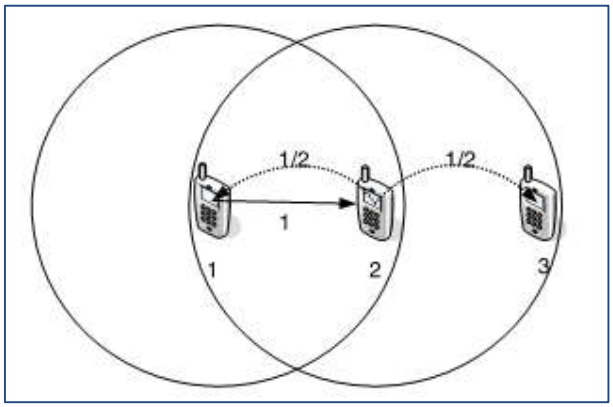


Fig. 1 A simple wireless ad hic network

End to End Traffic Matrix

From the above point to point matrix details end-to-end traffic matrix is found out using this module. Given a sequence of point-to-point traffic matrices Wj 1xK, the goal is to derive the end-to-end traffic matrix R = (r(i, j) NxN, where r(i, j) is the accumulative traffic volume from node i to node j, including both the point-to-point traffic captured directly and multi-hop traffic deduced from the point-to-point traffic use the accumulative traffic matrix and end-to-end traffic matrix interchangeably. The algorithm function f (W|1xK)) as the inputs to derive the accumulative traffic matrix R.

E. Source/Destination Probability Distribution

In this phase, the actual source and destination probability Distribution are denoted respectively, as two vectors S =(s(1), s(2), ... , s(N) and D =(d(1), d(2), ... , d(N)), where s(i) and d(i) (i = 1 to N) represent the probability for node i to be an actual source and destination, respectively. Note that if the total number of source nodes is m, then should have for S. However, since only care about the relative order among all possibilities (to know which nodes are more possible to be the actual sources) but not the total number m, can always assume m = 1.

During the distribution finding, the vector space similarity (or cosine similarity) of two vectors V and U is defined as follows:

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

End-to-End Link Probability Distribution

This module derives a probability distribution matrix P = (p(i, j)) N*N, in which each entry p(i, j) represents the probability of the i j linkability (i.e., node i and node j are a pair of actual source and destination). Again, note that only the relative order among these entries is of interest, since aim at discovering the most possible communication links. As described above, the probability for node i to be a destination depend on two factors: the traffic from each node j to node i and node j's probability to be a source.

Suppose j - i is an actual source-destination pair. If set the total traffic coming out from j to zero, the probability for i to be a destination will decrease. Similarly, if set the incoming traffic to node i to zero, the probability for node j to be a source will also decrease.Thus, identify a source-destination (S-D) pair by evaluating the significance of the probability reduction due to the elimination of the traffic sent by the source or received by the destination.

For instance, in the example scenario, to identify the most possible destination of node 1, it can erase all traffic sent by node 1 from the point-to-point traffic matrices, base on which compute the destination probability distribution D-. By comparing D- with D (obtained using the original point-to-point matrices), can find out the node whose destination probability drops most significantly due to elimination of the traffic sent by node 1. This node is most possible to be the destination of node 1. Functions are used to remove the traffic sent by node i. as well as to remove the traffic received by node j.

WIRELESS AD HOC NETWORK APPLICATIONS

The decentralized nature of wireless ad-hoc networks makes them suitable for a variety of applications where central nodes can't be relied on and may improve the scalability of networks compared to wireless managed networks, though theoretical and practical limits to the overall capacity of such networks have been identified. Minimal configuration and quick deployment make ad hoc networks suitable for emergency situations like natural disasters or military conflicts. The presence of dynamic and adaptive routing protocols enables ad hoc networks to be formed quickly. Wireless ad-hoc networks can be further classified by their applications.

Area Applications:

- ❖ Mobile ad hoc networks (MANET) - continuously self-configuring, infrastructure-less network of mobile devices connected without wires
- ❖ Vehicular Ad hoc Networks (VANETs) - communication between vehicles and roadside equipment.
- ❖ Intelligent manners during vehicle-to-vehicle collisions, accidents.
- ❖ Smart Phone Ad hoc Networks (SPANs) - wireless access points, or traditional network infrastructure.
- ❖ Mobile ad hoc networks (iMANETs) - Persistent System's Cloud Relay
- ❖ Military / Tactical MANETs - emphasis on security, range, and integration with existing systems.

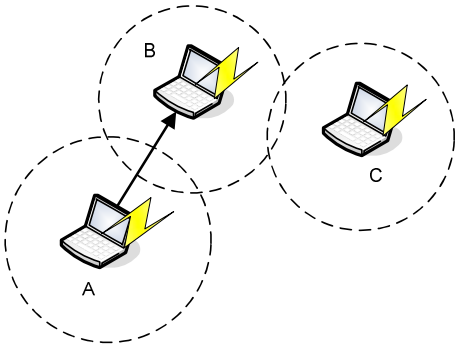
AD HOC ROUTING PROTOCOLS

An ad hoc routing protocol is a convention, or standard, that controls how nodes decide which way to route packets between computing devices in a mobile ad hoc network. In ad hoc networks, nodes are not familiar with the topology of their networks. Instead, they have to discover it: typically, a new node announces its presence and listens for announcements broadcast by its Neighbours. Each node learns about others nearby and how to reach them, and may announce that it too can reach them.

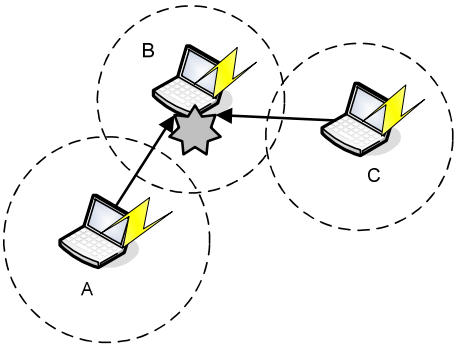
ROUTING IN MANETs

Unlike wired networks, routing in MANETs poses unique challenges. Designers of routing protocols for MANETs need to address several issues. In this chapter these issues are identified and the routing protocols available for MANETs are classified. Then working principle of a few protocols such as DSDV, DSR, AODV, etc. are explained. Their pros and cons are also identified. This chapter concludes with a summary of routing in MANETs.

Bandwidth constraint: The nodes in the network have a relatively low bandwidth when compared to traditional wired networks. This is an important issue to consider when designing routing protocols for MANETs since the utilization of bandwidth by the routing protocol in the network must be minimized.
Error prone broadcast channel: The nodes in the MANET broadcast the information to all the neigh boring nodes on the wireless channel. The channel itself is prone to several errors such as attenuation, multi-path fading, etc. Thus the routing protocol itself must be designed taking into consideration these issues.



A transmits data to B but C doesn't hear it



C transmits to B Collision

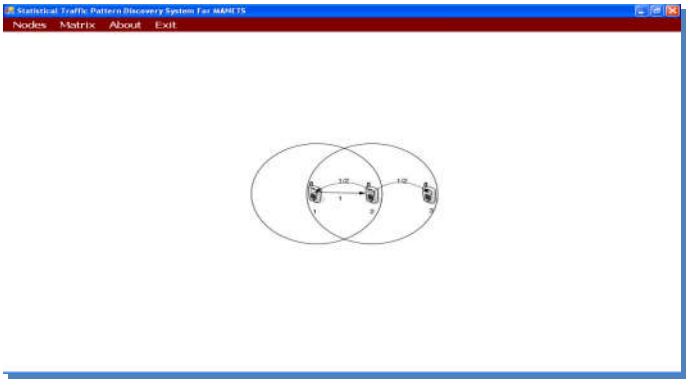


Fig:1 Input form design

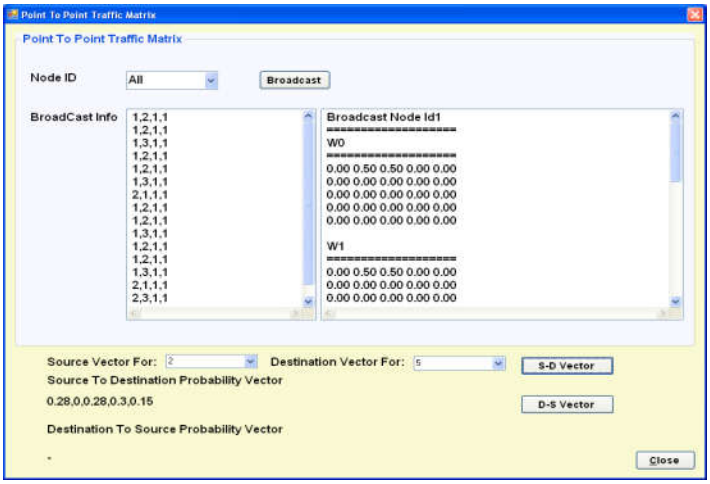
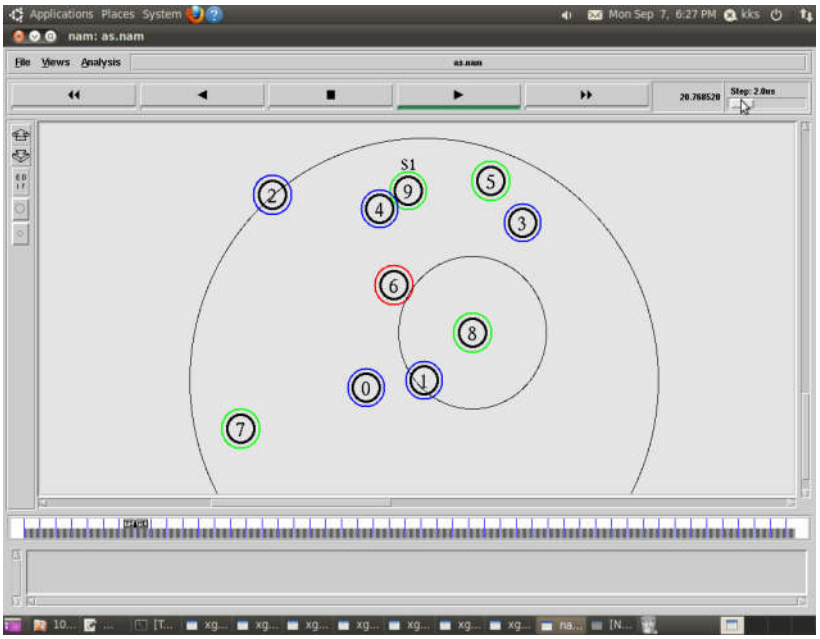


Fig. 2 Using NS2 at Node Identifying



IV. CONCLUSIONS

This research work proposes a novel STARS for MANETs. STARS 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, STARS 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.

The empirical study demonstrates that the existing MANET systems can achieve very restricted communication anonymity under the attack of STARS. In addition, the STPDS is extended as GSTPDS which divides the entire network into multiple regions geographically; and deploys sensors along the boundaries of each region to monitor the cross-component traffic.Its commonly used in Network Simulator tools. Also it treats each region as a super node and use GSTPDS to figure out the sources, destinations, and end-to-end communication relations.

REFERENCES

[1] Seys and B. Preneel. Arm: Anonymous routing protocol for mobile ad hoc networks. In IEEE AINA, 2006.
[2] S. Seys and B. Preneel, "ARM: Anonymous Routing Protocol for Mobile Ad Hoc Networks," Proc. IEEE 20th Int'l Conf. Advanced Information Networking and Applications Workshops (AINA Work- shops '06), pp. 133-137, 2006.
[3] 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.
[4] Y. Zhang, W. Liu, W. Lou, and Y. Fang, "MASK: Anonymous On-Demand Routing in Mobile Ad Hoc Networks," IEEE Trans. Wireless Comm., vol. 5, no. 9, pp. 2376-2385, Sept. 2006.
[5] Y.-C.Hu,A.Perrig,andD.B.Johnson, "Ariadne: A secure on-demand routing protocol for ad hoc networks," in ACM MobiCom, Atlanta, GA, Sep. 2002.
[6] Y. Qin and D. Huang, "OLAR: On-Demand Lightweight Anonymous Routing in MANETs," Proc. Fourth Int'l Conf. Mobile Computing and Ubiquitous Networking (ICMU '08), pp. 72-79, 2008.
[7] Y. Zhang, W. Liu, and W. Lou. Anonymous communications in mobile ad hoc networks. In INFO-COM, 2005.
[8] 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. 1-92010.

ABOUT AUTHORS

(1) Dr.K.Prabha, she was working Assistant Professor of Computer Science in Periyar University PG Extension Centre, Periyar University, Dharmapuri, Tamilnadu, India. Her research interests include Network Security and Data Mining.
(2) K.Nirmaladevi, she is doing her Ph.D Computer Science in Periyar University PG Extension Centre, Periyar University, Dharmapuri, Tamilnadu, India. Her research interests include Networking in MANET.