MACQUARIE UNIVERSITY

IT PROJECT - ITEC810

Bandwidth Management Framework For IP based Mobile Ad Hoc Networks (MANETs)

Supervisor: Dr. Rajan Shankaran

Written By: KHALID IQBAL (41152190)

June 05, 2009

Abstract

The wireless networks represent a new computing paradigm that has as its' main function the provision of permanent access to mobile users, independently of their physical location. With the decrease in costs of portable devices and with an increase in their capacity, a new concept has arisen. This concept is named as 'Mobile Ad Hoc Network' (MANET).

The demands made by users are increasing as the 'mobile ad hoc networks' became an essential part of this era and call for the use of heavy applications such as multimedia and videoconferencing. These applications require a lot of network bandwidth which is subject to constraint and incur error rates during transmission. To address these problems, this paper presents an in-depth study of the new and emerging research area of quality of service (QoS) support in mobile ad hoc networks (MANETs) and examines the challenges of incorporating scalable bandwidth management scheme in MANET.

This paper also identifies a technique from the available body of literature and discusses how this technique solves bandwidth provisioning problem in mobile ad hoc networks. The objective of this solution is to obtain a better usage of network transmission channels, whilst at the same time accommodating flexible resource allocation.

Acknowledgement

My sincere thanks to my supervisor “Dr. Rajan Shankaran” and course coordinator “Robert Dale” who has given me opportunity of doing this paper work at Macquarie University. My heart-felt thanks to my family, friends and colleagues who have helped me in completion of this work.

TABLE OF CONTENTS

Introduction		06
1.	Mobile Networks	07
1.1	Mobile Ad Hoc Networks (MANET)	07
1.1.1	MANET Features	08
1.1.2	MANET Applications	09
1.2	Quality of Service (QoS)	10
1.2.1	Basic QoS Architecture	10
1.2.2	QoS Identification and Marking	11
1.2.3	Queuing and Scheduling	11
1.2.4	Policy Management	12
1.2.5	End to End QoS Models	12
1.2.6	QoS from a service provider's perspective	13
1.2.6.1	QoS Resource Reservation	13
1.2.6.2	QoS Routing	13
1.2.6.3	Admission Control	13
1.3	Need for QoS in MANETs	13
1.3.1	Key Points	14
2.	MANET Routing Protocols	15
2.1	Proactive or Table-driven Protocols	15
2.1.1	Destination-Sequenced Distance-Vector (DSDV)	16
2.1.2	Optimized Link State Routing protocol (OLSR)	16
2.1.3	Fisheye State Protocol (FSR)	16
2.2	Reactive or On-demand Protocols	17
2.2.1	Dynamic Source Routing Protocol (DSR)	18
2.2.2	Temporally Ordered Routing Algorithm (TORA)	18
2.2.3	Ad hoc On-demand Distance Vector (AODV)	18
2.2.4	Associativity Based Routing (ABR)	19
2.3	Hybrid Approach	19
2.3.1	Zone Routing Protocol (ZRP)	19
3.	Quality of Service Models	20
3.1	Integrated Services (IntServ)	20
3.1.1	IntServ Limitation	21
3.2	Differentiated Services (DiffServ)	21
3.2.1	DiffServ benefit	22
3.3	Why Existing QoS techniques fail in MANET environment	22
3.3.1	IntServ in MANETS	22
3.3.2	DiffServ in MANETS	22

4. QoS models for MANETs 23
4.1 Flexible QoS Model for Mobile Ad Hoc Network: FQMM 23
4.1.1 FQMM limitations 24
4.2 INSIGNIA 24
4.2.1 INSIGNIA limitations 25
4.3 INORA 25
4.4 Adaptive ReServation And Pre-Allocation QoS Architecture 26
4.4.1 Soft/Hard Reservation 26
4.4.2 Adaptive QoS Monitoring 26
4.4.3 ASAP Signaling System 26
4.4.4 Limitations of ASAP scheme 27
4.4.4.1 Flow Restoration Problem 27
4.4.4.2 Reverse Path Problem 27

5. Remedial solutions to address ASAP scheme limitations 28
5.1 Local Repair 28
5.2 Dynamic Virtual Path 29
5.3 Flexible Timing Control 29
5.4 Congestion Control 30
5.5 Conclusion 31

6. References 32

Introduction

The objectives of this paper are 1. Critically analyze different solutions that exist to guarantee QoS in the transmissions in wireless networks, 2. propose a mechanism that warranties QoS for the ad hoc topology wireless networks.

In this paper, a scalable quality of service (QoS) framework referred to as adaptive bandwidth reservation and pre-allocation QoS architecture (ASAP) has been explained and its proposed extensions are analyzed. These extensions provide remedial solution to mobile ad hoc networks. ASAP extended framework includes local repair mechanism which is used to overcome the problems related to the inertness of resource reservations. The reverse path problem is addressed by using dynamic virtual paths. Flexible timing control and congestion control increases the efficiency and performance of ASAP. All these mechanisms together make ASAP scheme scalable to network mobility.

This paper, is divided into five sections. Section 1, provides a brief overview of MANET, QoS and QoS architecture. It then introduces to the reader the needs for QoS in MANETs. Section 2, presents different MANET routing protocols. In section 3, existing QoS models are explained with a detailed discussion of how traditional QoS techniques fail to scale well in MANETs. In section 4, some schemes that have been specifically designed to address the provision of QoS (specifically) bandwidth in MANET are discussed. Finally in section 5, few existing as well as proposed solutions to a QoS signaling protocol “ASAP” for flexible resource allocation are presented with some concluding remarks.

1. Mobile Networks

Mobile networking is one of the most important technologies of modern era. During the last decade, advances in networking techniques have led to the vast use of mobile hosts and wireless networking. In general, there are two distinct approaches for enabling wireless mobile units to communicate with each other:

Infra-structured: [Marco Conti, 2003] Traditionally, wireless mobile networks rely on good infrastructure support, in which mobile devices communicate with access points like base stations connected to the fixed network infrastructure. Examples of this kind of wireless networks are GSM, UMTS, WLAN, etc.

Infra-structure-less: In this approach, the mobile wireless network is commonly known as a mobile ad hoc network (MANET) [Marco Conti, 2003]. A mobile ad hoc network is a self-configuring network of mobile devices connected by wireless links. Nodes in mobile ad hoc network are free to move and organize themselves in an arbitrary fashion. Each user is free to roam about while communicating with others. The path between each pair of the users may have multiple links and the radio between them can be heterogeneous. This allows an association of various links to be a part of the same network.

1.1 Mobile Ad Hoc Network (MANET)

A mobile ad hoc network is a collection of wireless nodes which can dynamically be set up anytime and anywhere without pre-existing network infrastructure. MANET is an autonomous system in which mobile nodes connected by wireless links are free to move randomly and often act as routers as well.
The traffic types in mobile ad hoc networks are different from those in infra-structured wireless network, these types includes:

Peer to Peer: Communication between two nodes which are within one hop.

Remote to Remote: Communication between two nodes beyond a single hop but they maintain a stable route between them. This may be the result of many nodes staying within communication range of each other in one area or possibly moving as a group.

Dynamic Traffic: This occurs when nodes are dynamic and are moving which results in a poor connectivity and degraded network activity.

Following figure shows the architectures of infra-structured and infra-structure-less ad hoc wireless networks.

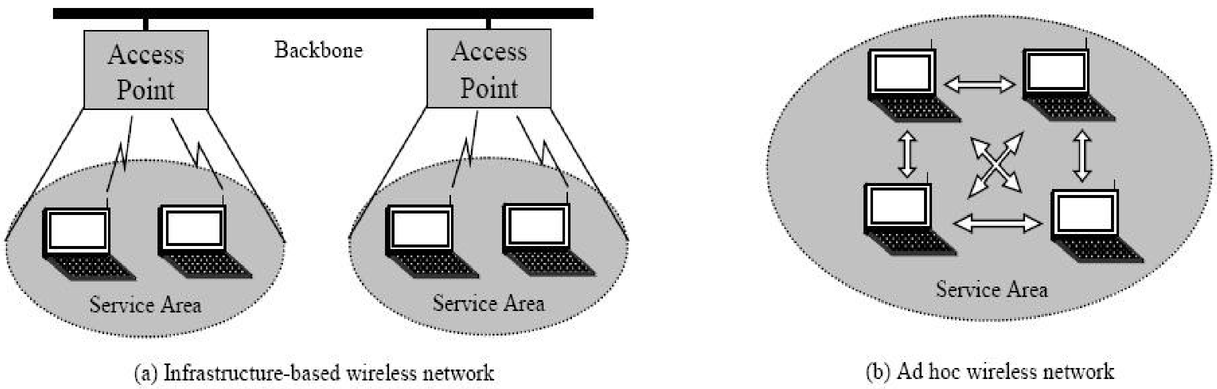

Figure: 1

1.1.1 MANET Features

MANET has the following salient features:

Autonomous terminal*:* In MANET, each mobile node is independent, which may function as both a host and a router. Usually endpoints and switches are impossible to differentiate in MANET.

Distributed operation*:* As there is no central control for network operations therefore the control and management of the network is distributed among nodes. To implement functions the nodes involved should collaborate amongst themselves and each node acts as a relay as required.

Routing Algorithms*:* Basic ad hoc routing algorithms can be single-hop and multi-hop, based on link layer attributes and routing protocols. With lesser functionality and malleability single-hop MANET is simpler than multi-hop in terms of structure and implementation.

Dynamic topology*:* Since the nodes in MANET are mobile, the network topology may change quickly and randomly and the nodes connectivity may vary with time. Mobile nodes dynamically establish routing among themselves as they move around, establishing their own network on the fly. Additionally, a MANET user may require access to a fixed network (e.g. Internet) while communicating within the ad hoc network.

Varying link*:* There is high possibility of bit-error rates in MANETs due to noise, fading, and intervention among channels over which the nodes communicate. It has less bandwidth than a wired network.

Light-weight terminals: MANET nodes are mobile devices with less processing capability, small memory size and low power storage. Such devices require optimized algorithms and mechanisms which implement the computing and communicating functions.

1.1.2 MANET Applications

Ad hoc networking allows mobile devices to maintain connections among them as well as allows adding and removing devices to and from the network. In recent years, ad hoc networking is gaining importance due to the increase in number of widespread applications.

Typical MANET applications include:

Military battlefield: It allows military to take advantage of this self-creating and self-organizing network technology to maintain links between the soldiers, vehicles and military head quarters.

Commercial use: Ad hoc networking can also be very helpful in emergency/rescue operations e.g. in fire, flood or earthquake during disaster relief efforts. Such networks can easily be created where communications infrastructure is damaged and rapid deployment of a communication network is required.

Local level: Ad hoc networks can be set up using notebook computers or palmtop computers to share information e.g. conference or classroom. Another local level application might be a home network where different devices can communicate directly to exchange information.

Personal Area Network (PAN*)*: Short-range MANET between various mobile devices can simplify the intercommunication. Such an ad hoc network can also be extended to any other network.

1.2 Quality of Service (QoS)

Quality of Service (QoS) [Ferguson, et al, 1998] refers to the capability of a network to provide better services to selected network traffic. The main goal of QoS in MANET is to achieve a more reliable network behavior so that information carried by the network can be better delivered with better utilization of network resources.

The United Nations Consultative Committee for International Telephony and Telegraphy (CCITT) Recommendation has defined QoS as: "*The collective effect of service performance which determines the degree of satisfaction of a user of the service*".

Fundamentally, QoS enables to provide better service to certain communication flows. This can be achieved either by raising the priority of a flow or limiting the priority of another flow, usually congestion management tools are used for this purpose. Policing and shaping provides priority to a traffic flow by limiting the throughput of other flows.

1.2.1 Basic QoS Architecture

The basic QoS architecture introduces following fundamental pieces for QoS implementation:

1. Packet/flow identification and marking techniques
2. Scheduling and queue management in network elements.
3. Policy management to control and manage traffic across a network.
4. End to end QoS based framework that supports QoS based protocols.

Following figure shows three main components of QoS implementation.

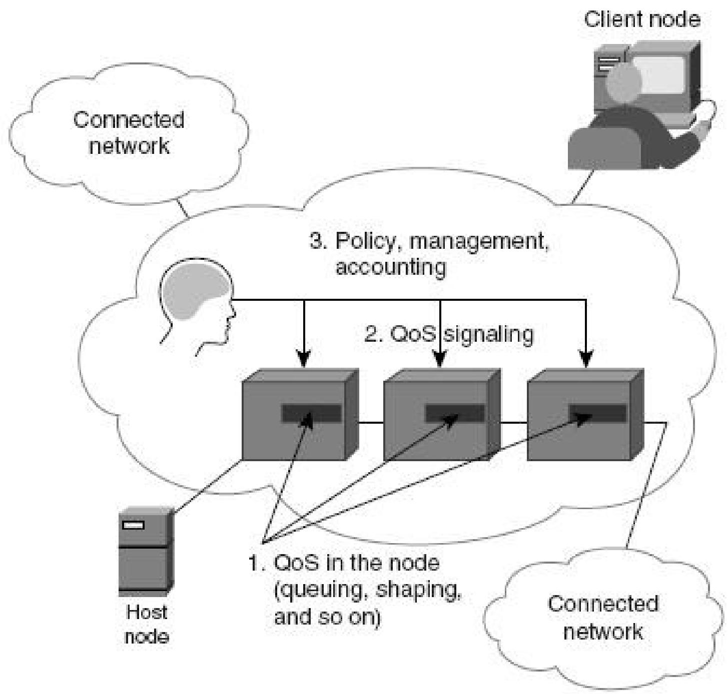

Figure: 2

1.2.2 QoS Identification and Marking

To provide preferential service to specific traffic, firstly it must be identified and secondly that packet may or may not be marked. These two tasks make up packet classification. When the packet is identified but not marked classification is said to be on a per hop basis. This is when the classification pertains only to the router that it is on and not passed to the next router. This happens with *priority queuing* and *custom queuing*. Common methods of identifying flows include access control lists, policy-based routing, committed access rate and network-based application recognition.

1.2.3 Queuing and Scheduling

Queues are not of infinite size, they can overflow. When a queue gets full additional packets would be dropped. So, a mechanism is necessary which can do two things:

1. Make sure that the queue does not fill up, so that there is room for high priority packets.

2. Allow some sort of criteria for dropping packets that are of lower priority before dropping higher priority packets

1.2.4 Policy Management

QoS management helps to set and evaluate QoS policies and goals. Following steps are involved in QoS management:

- Baseline the network with devices such as RMON probes. This helps in determining the network traffic characteristics. Also, applications targeted for QoS should be base lined.
- Deploy QoS techniques when the traffic characteristics have been obtained and applications are identified for increased QoS.
- Evaluate the results by testing the response to see whether the QoS goals have been achieved.

1.2.5 End-to-End QoS Models

Service levels refer to the actual end-to-end QoS capabilities which determine the ability of a network to deliver service needed by specific network traffic from end to end.

Three basic levels of end-to-end QoS can be provided across a heterogeneous network,

Best-effort service: Best-effort service is basic connectivity with no guarantee. **Differentiated service:** Some network traffic is treated better than the rest. This is provided by traffic classification and also called soft QoS.

Integrated service: This is an absolute reservation of network resources for specific traffic. Also called hard QoS.

All these three QoS models are explained with further details in **Section 3.**

1.2.6 QoS from a service provider's perspective

From the service provider's perspective, the provision of QoS will involve the followings:

1.2.6.1 QoS Resource Reservation

Each flow with some QoS guarantee will have to be allocated some resources to ensure that as soon as the packet of that particular flow arrives, it will not have to wait for some resource to be freed and it will be transmitted to the next node instantly. Another issue is which resources should be reserved for a given set of QoS parameters and how these resources will be freed when the flow is terminated.

1.2.6.2 QoS Routing

QoS routing needs are to find a route to the destination which satisfies QoS requirements and should be provided by the network layer. Such routing is also known as "QoS-aware routing."

1.2.6.3 Admission Control

Admission control is a network Quality of Service procedure; it determines how bandwidth and latency are allocated to streams with various requirements. Admission control schemes need to be implemented between network edges. An application that uses the network to transport traffic with QoS must first request a connection, which involves informing the network about the characteristics of the traffic and the QoS required by that application. The network judges whether it has sufficient resources available to accept the connection and then either accepts or rejects the connection request. This is known as Admission Control.

1.3 Need for QoS in MANETs

Without QoS policies, each packet is given equal access to network resources. In order to efficiently utilize network resources, it is necessary to identify which network traffic is critical and allocate appropriate network resources to support those traffic streams. Secondly, different Applications require different network performance, based on bandwidth needs and latency sensitivity, therefore; QoS is needed in mobile ad hoc networks. If the latency sensitivity is high, higher will be the bandwidth requirement. The data transfers can have zero tolerance for packet loss and high tolerance for delay and jitter. Voice and Video conferencing applications require both high bandwidth and high latency sensitivity.

As mentioned above, traffic behavior and QoS requirements for different applications vary from application to application; following table is to present different applications' requirements in mobile ad hoc networks.

Applications	Traffic Behavior	QoS Requirements
Electronic Mail (SMTP) File Transfer (FTP) Remote Terminal (Telnet)	Traffic Behavior Small, batch file transfers	QoS Requirements Very tolerant of delay B/W requirement: low **Best effort**
HTML Web Browsing	Series of small, bursty file xfer	Tolerant of moderate delay B/W requirement: varies **Best effort**
IP-based Voice (VoIP) Real Audio	Constant or variable bit rate	Very sensitive to delay/jitter B/w requirement: low **Requires predictable delay/loss**
Streaming Video	Variable bit rate	Very sensitive to delay/jitter B/w requirement: High, variable **Requires predictable delay/loss**

Figure: 3

1.3.1 Key Points

Summarizing, a QoS framework for MANETs would achieve the followings:

- Intelligent buffer handling and queuing
- Prioritization of routing control messages
- User traffic prioritization
- Efficient resource usage
- QoS aware routing

2. MANET Routing Protocols

In this section key routing strategies are described which can be deployed in a MANET environment. These routing strategies are as follows:

- Proactive or Table-driven protocols
- Reactive or On-demand protocols
- Hybrid protocols

2.1 Proactive or Table-driven Protocols

Networks using proactive protocols; maintain the knowledge of every route by exchanging route information periodically regardless of whether the routes are being used or not. Each node maintains the necessary routing information, and propagates topology updates in response to instantaneous connectivity changes in the network.

These protocols tend to suffer from wasted bandwidth due to the large control overhead in maintaining unused routes, especially when frequent changes in network topology occur.

The advantage of proactive protocols is that route information to the destination is available immediately and no delay is experienced whenever an application sends packets. In case of interactive applications it is really useful.

Following mechanisms are adopted in proactive protocols:

- Topology information stored at each node.
- Varying dynamically the size of route updates and update frequency
- Optimizing flooding
- Combining Distance Vector (DV) and Link State (LS) features

Examples of proactive or table-driven protocols include;

- Destination-Sequenced Distance-Vector Routing (DSDV)
- Optimized Link State Routing (OLSR)
- Fisheye State Rou ting (FSR)

2.1.1 Destination-Sequenced Distance-Vector: DSDV

DSDV [Perkins C. and Bhagwat P. 1994] is based on the classical Bellman-Ford routing algorithm. Every node maintains detailed information of all destinations. Each entry is marked with a sequence number and to reduce network traffic generated by route updates, it uses full dump or incremental packets. An aging mechanism based on increasing sequence numbers indicates updated route to avoid routing loops and the count-to-infinity problem. It also provides readily available routing information regardless of whether the source-node requires route or not.

2.1.2 Optimized Link State Routing protocol: OLSR

The Optimized Link State Routing (OLSR) [Clausen, T., et al, 2009] protocol is an optimization over the classical link state protocol, developed for mobile ad hoc networks. It is proactive protocol and exchanges topology information with other nodes regularly. It is suitable for large networks and where the traffic is random. It uses hop-by-hop routing which means that every node uses its local information to route packets. OLSR reduces the flooding overhead by using only selected nodes called MPRs (multipoint relay). This technique reduces the number of retransmissions required to flood a message to all network nodes.

2.1.3 Fisheye State Protocol: FSR

Fisheye State Protocol (FSR) [Grilo, A, et al, 2005] is a proactive protocol which is based on “fisheye” technique. It was proposed to reduce the size of information required to represent graphical data. The fisheye approach is to maintain accurate distance and path quality information of the immediate neighborhood, with progressively less detail as the distance increases. FSR maintains topology details at every node. Link state packets are not flooded in fisheye state protocol instead nodes maintain a table having link state information based on information received from neighboring nodes and periodically exchange this information with local neighbors only.

2.2 Reactive or On-demand Protocols

Second approach for routing in wireless mobile ad hoc environment is routing on-demand protocols. This approach is characterized by the elimination of routing tables at nodes and consequently the need of route updates to track changes in the network topology. On-demand routing protocols calculate path before data transmission. If data is not generated by nodes, then the routing activity is completely absent. For this reason, they are also called reactive protocols. A reactive protocol is characterized by the following procedures used to manage paths:

- Path discovery
- Path maintenance
- Path deletion (optional)

Data forwarding in reactive protocols is accomplished according to two main techniques:

- Source routing
- Hop-by-hop

These protocols build routes only when required for carrying traffic, therefore route discovery process is pre-requisite to establish communication between two nodes and then, that route is maintained until the communication finishes. Reactive protocols also tend to generate large overheads and suffer packet loss as topology changes become more frequent. However, these protocols perform better than proactive protocols, especially when topology changes are not very frequent.

Examples of on-demand protocols include:

- Dynamic Source Routing (DSR)
- Temporally Ordered Routing Algorithm (TORA)
- Ad Hoc On -demand Distance Vector Routing (AODV)
- Associativity Based Routing (ABR)

2.2.1 Dynamic Source Routing Protocol: DSR

The DSR [Broch, J., et al, 2001] protocol is very efficient and simple routing protocol which is specifically designed for use in multi-hop wireless ad hoc networks. By using this protocol, the network is completely self-organizing and self-configuring, requiring no administration or network infrastructure. All routing is automatically determined and maintained by the DSR Routing Protocol whenever nodes in the network join or leave the network or any wireless transmission conditions change.

This protocol is composed of two mechanisms called *Route discovery* and *Route maintenance* that work together to allow the discovery and maintenance of routes in mobile ad hoc networks.

2.2.2 Temporally Ordered Routing Algorithm: TORA

TORA [Vincent D, et al, 1997] is designed to react efficiently against any topological changes and to deal with network partitions. TORA belongs to one of a family of algorithms referred as “link reversal”. The main objective is to find stable routes that can be quickly and locally repaired. This protocol is highly adaptive, efficient, scalable and very well suited for use in large mobile networks. TORA provides routing by a completely different approach from those described so far.

In this protocol functioning is divided into three phases: route discovery, route maintenance and route deletion. A key feature of TORA is that, if one of the outgoing links of a node breaks, the node has at least one other downstream node and the destination can still be reachable, and thus no repairing activities are required.

2.2.3 Ad hoc On-demand Distance Vector: AODV

Ad hoc On-demand Distance Vector protocol [Charles. E. Perkins, et al, 2000] builds on the DSDV algorithm. AODV provides loop free routes while repairing broken links; because the protocol does not require global periodic routing advertisements. AODV uses symmetric links between neighboring nodes. It does not try to follow paths between nodes where one of the nodes cannot hear the other one. Nodes do not lie on active paths; they neither maintain any routing information nor exchange any periodic routing table information. At intermediate nodes AODV relies on dynamically establishing route table entries.

2.2.4 Associativity Based Routing: ABR

ABR [Toh, C-K., 1997] is designed for mobile ad hoc networks where mobile nodes act as routers in a wireless environment with no base stations. ABR is based on the concept of associativity. The protocol is source initiated, thus; there is no need for periodic routing updates. The main idea of associativity is that there is no point in choosing a route based on the shortest path when the route is going to be broken or invalidated due to node's mobility. Every node in the network learns its 'associativity' with its neighboring nodes to determine the best route. Association in ABR takes up a few metrics such as link delay, signal strength, power life and route relaying load etc. Routes are chosen when they have a high degree of associativity.

A detailed survey containing the communication complexities of both proactive and reactive routing protocols is presented by [Royer, E.M. and Toh, C.-K. 1999].

2.3 Hybrid approach

The hybrid approach combines the features of both proactive and reactive routing. Example of such approach is Zone Routing Protocol.

2.3.1 Zone Routing Protocol: ZRP

ZRP [Haas, Z.J. and Pearlman, M.R. 1999] allows the use of proactive routing within a local zone while reactive routing scheme is used globally. Its performance depends on zones organization within the network and the traffic patterns. Routes are immediately available within the zone and for destinations which are outside the zone; ZRP uses route discovery procedure.

The aim of this protocol is to address the problems by combining the best properties of both proactive and reactive approaches.

3. Quality of Service Models:

The QoS model specifies the architecture in which certain services can be provided in the network. Major internet QoS models are described in this section. These models are named as;

- Integrated Services (IntServ)
- Differentiated Services (DiffServ)

3.1 Integrated Services (IntServ): [Braden R., et al, 1994]

Internet offers only point-to-point best-effort data delivery. To control network congestion, routers use a simple FIFO service policy, which reply on buffer management and packet discarding. Typically, an application does not have any knowledge whether all its data transmitted over the network will be delivered to the receiving end until explicitly notified by the network. New service architecture is needed to support real-time applications, with various QoS requirements. It is referred as Integrated Services (IntServ).

The concept is very simple, distinguishable stream of related datagrams that require the same quality of service, to provide QoS to that stream different service class requires the network and routers to explicitly manage their bandwidth and buffer resources. This implies resource reservation, admission control, packet scheduling, and buffer management which are also key building blocks of integrated services.

Moreover, this requires a flow specific state in the routers. The flow specific state includes bandwidth requirement, delay bound, and cost of the flow. As the Internet is connectionless therefore a soft-state approach is adopted to refresh flow rates periodically using a signaling system, such as resource reservation protocol (RSVP).

There are two service classes defined within the IntServ:

- Guaranteed Service (GS)
- Controlled load Service (CLS)

Guaranteed Service traffic is characterized by peak rate, token bucket parameters and maximum packet size. GS needs traffic access control at the user side and fair packet queuing (FPQ) at routers to provide a minimum bandwidth. To guarantee no packet loss, correct buffering can be provided at each router.

Controlled load Service is for applications that can tolerate variance in packet delays, as well as a minimal loss rate that must closely approximate the basic packet error rate of the transmission medium. The CLS does not accept specific target values for delay or loss. It uses loose admission control and simple queue mechanisms and is basically for adaptive real time communications.

3.1.1 IntServ Limitations:

Integrated services require buffer management and packet scheduling on a per flow basis, therefore, it becomes very difficult and costly for the routers to provide IntServ if the number of flows and line rate increase.

3.2 Differentiated Services (DiffServ): [Blake, S. 1998]

Differentiated Services (DiffServ) is another solution, which can provide QoS control on a service class basis. DiffServ is more feasible and cost-effective than the IntServ. It provides scalable service discrimination in the Internet without the need of per flow state and signaling at every hop. The DS approach is to provide QoS in networks employs a well-defined set of building blocks from which a variety of services may be built. These building blocks are listed below;

• Setting bits in the type of service (ToS) byte at network edges and administrative boundaries

• Using those bits to determine how packets are treated by the routers.

• Conditioning marked packets at network boundaries according to the requirements of each service.

In this model, network traffic is classified and conditioned at the entry point and assigned to different behavior aggregates. Each aggregate is assigned a DS codepoint. Different DS codepoints signify that the packet should be differently handled by the interior routers. Each different type of processing that is provided to the packets is called a different per-hop behavior (PHB). In the network, packets are forwarded according to the PHBs. The PHB is indicated by a DiffServ codepoint (DSCP) in the IP header of every packet. The DSCP markings are applied by a trusted customer or by the boundary routers on entry to the DS network.

3.2.1 DiffServ Benefit:

The advantage of DiffServ is that many traffic streams can be aggregated to one which is forwarded using the same PHB at the routers, thereby simplifying the processing. In DiffServ QoS is invoked on packet-by-packet basis, there is no signaling and no other processing is required in the core of the DS network.

3.3 Why Existing QoS techniques fail in MANET environment

3.3.1 IntServ in MANETS: [Xiao H., et al, 2003]

The IntServ model is not suitable for MANETs because of following shortcomings;

Scalability: IntServ provides per flow granularity so the amount of state information increases proportionally with the number of flows and this results as storage and processing overhead on routers, which is the scalability problem of IntServ.

Signaling: Signaling protocols have three phases: connection establishment, connection maintenance and connection teardown. In highly dynamic networks such as MANETs, this is not promising approach since routes may change very quickly and the adaptation process of protocols using a complex handshake mechanism would just be too slow. Moreover, the signaling overhead is a potential problem as well.

3.3.2 DiffServ in MANETS: [Xiao H., et al, 2003]

The main drawbacks of DiffServ approach in MANETs are as follow:

SLA (Service Level Agreement): DiffServ is based on the concept of SLA's. SLA is a sort of contract between a customer and its Internet Service Provider (ISP) which specifies the forwarding service the customer will receive. In other words, these are the traffic conditioning rules which are placed at the nodes from where the traffic originates. They are responsible for re-marking the traffic streams, discarding or shaping packets according to the traffic profile. In ad hoc topology there is no obvious scheme for the mobile nodes to negotiate the traffic rules therefore making SLAs is difficult in MANETs.

Ambiguous core network: In DiffServ, traffic classification and conditioning has to be done at the boundary nodes. This makes QoS provisioning much easier in the core of the network. In MANETs there is no obvious core network because every node is a potential sender, receiver and router.

4. QoS models for MANETs

In the previous section it was discussed that MANETs propose different requirements to quality of service infrastructures than wired networks do and finally concluded that neither a pure IntServ nor a DiffServ approach is satisfying.

Following section describes few other QoS technologies which are specifically designed for mobile ad hoc networks and are based on the concepts identified in the last section.

4.1 Flexible QoS Model for Mobile Ad Hoc Network: FQMM

FQMM [Xiao, H., et al, 2000] is designed for small to medium sized MANETs, using a flat non hierarchical topology. It defines three kinds of nodes exactly as in DiffServ:

- An ingress node is a mobile node that sends data.
- Interior nodes are the nodes that forward data to the other nodes.
- An egress node is a destination node.

For example, in following figure there are eight mobile nodes and a route is established for communication from node M1 to M6. When data is sent from M1 to M6, M1 behaves as an ingress node; classifying, marking, and policing packets. Nodes M3, M4, and M5 behave as interior nodes to forward data. M6 is the egress node which is also the destination as well.

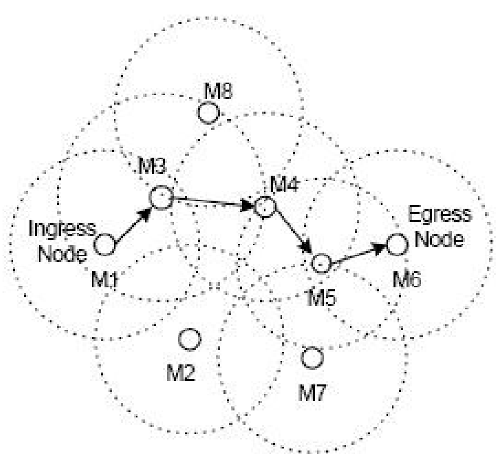

Figure: 4

The basic idea of FQMM is that it uses a hybrid per-flow and per-class provisioning policy. In this scheme, traffic of the highest priority is given per-flow provisioning while other priority classes are given per-class provisioning. Like DiffServ, FQMM has service differentiation. The per-class granularity can be improved to per-flow granularity for certain classes of high priority traffic. However, it is difficult to offer per flow granularity to all the traffic in a MANET due to bandwidth limitation and other constraints.

This QoS model is flexible because of the following features:

- Nodes have dynamic roles.
- The provisioning policies are hybrid and flexible.
- FQMM can be combined in a flexible manner to meet different network conditions and QoS requirements.

4.1.1 FQMM limitations

FQMM is an attractive solution for MANETs; however it still has some major problems:

- Without an explicit control on the number of services with per-flow granularity, the scalability problem still exists.
- FQMM may not be able to satisfy hard QoS requirements, due to its DiffServ behavior in ingress nodes,

4.2 INSIGNIA

INSIGNIA [Farkas K., et al, 2006] signaling protocol is specially designed for mobile ad hoc networks. INSIGNIA supports algorithms like fast flow reservation, restoration and adaptation; which are designed to deliver adaptive real time service. INSIGNIA implements an in-band approach by encapsulating some control signals in the IP option of every data packet, which is called INSIGNIA option. Moreover flow state information is kept in every node and this information is refreshed periodically.

Working: When a source node requests to establish a reservation to a destination node it sets the reservation (RES) mode bit in the INSIGNIA IP option service mode of a data packet and forwards that data packet toward the destination node. The bandwidth request field allows source node to specify its maximum and minimum bandwidth requirements. On reception of a RES, intermediate routing nodes execute admission control to accept or deny the request. When a node accepts a request, resources are dedicated and packets are scheduled accordingly. If the reservation is denied, packets are treated as best effort mode packets.

4.2.1 INSIGNIA limitations

Although INSIGNIA presents very promising approach to QoS support in MANETs but the system still lacks a few essential features:

- The most significant drawback of INSIGNIA is scalability problem due to the flow state information which is kept within the nodes of a certain path.
- Bandwidth usage in INSIGNIA is not efficient. The extra reservation on the path from the sending node to the bottleneck is a waste of bandwidth.
- INSIGNIA imposes a major processing overhead on the network as there is no mechanism to dynamically change the frequency by which control signals are inserted into the data packets.
- INSIGNIA offers only two bandwidth levels to be used, MINIMUM and MAXIMUM. A better approach would be needed in order to satisfy application requirements and to fully utilize the available resources.

4.3 Intelligent Optimization Self-Regulated Adjustment: INORA

INORA [Dharmaraju D., et al, 2002] represents a QoS signaling approach in a loosely coupled manner. The idea is based upon to provide multiple routes between a source and destination. INORA also provides feedback to the routing protocol on per hop basis to direct the flow along the route that is able to satisfy the QoS requirements of the flow.

The concept of loosely coupling QoS signaling and routing is a very promising approach. However, the interface for signaling to access routing should be as generic as possible in order to guarantee portability.

4.4 Adaptive Reservation and Pre-allocation QoS Architecture: ASAP

ASAP [Jianbo Xue 2003] provides adaptive QoS support to real time applications in infrastructure based wireless IP networks. The goal of this analysis is to extend the ASAP framework so that it could be deployed in mobile ad hoc networks.

4.4.1 Soft/Hard Reservation

In ASAP architecture, a new reservation concept, soft/hard reservation is introduced for efficient resource allocation and better performance. Soft reservation can be considered as the claim of a traffic flow for a certain bandwidth to be used in future. Hard reservation enables a traffic flow to exclusively reserve some bandwidth.

The actual reservation mechanism is two pass based. When a new real time flow is about to start, a soft reservation request is sent first. If there are enough resources available, the requested bandwidth will be soft reserved for that flow. After a soft reservation is established, the end node sends a hard reservation message requesting the same amount of bandwidth. This hard reservation will remove all the traffic occupying the corresponding soft reserved bandwidth. So after a hard reservation, the QoS traffic can immediately start running with its necessary QoS support.

4.4.2 Adaptive QoS Monitoring

QoS monitoring packets periodically investigates the QoS situation on every node within a certain path. Hard reservation messages are sent whenever the end-to-end QoS changes. Monitoring interval can be changed dynamically. For example more frequent monitoring is needed, if the network is unstable, in order to adapt bandwidth fluctuations. If the network is stable, processing overhead can be saved by keeping the monitoring rate low.

4.4.3 ASAP Signaling System

ASAP also provides efficient in-band signaling for resource reservation, management, adaptation and releasing. This signaling is designed to produce minimum possible overhead and to provide maximum flexibility.

4.4.4 Limitations of ASAP scheme

Although ASAP makes use of in-band signaling and fast adaptation but the protocol still fails to meet some MANET specific demands. Few problems of ASAP in a mobile ad hoc environment are described below.

4.4.4.1 Flow Restoration Problem

If no reservation is made at any node for data transmission then the flow will be transmitted using best effort. During transmission if node starts moving and gets out of the network range, routing then finds a new path. This state remains until the soft reservation (SR) message detects the missing reservation and triggers a hard reservation message, which will repair the missing reservation problem. The repairing mechanism has few major shortcomings.

1. The latency for flow restoration can be quite long.
2. Detecting a missing reservation.

4.4.4.2 Reverse Path Problem

Hard reservation message is supposed to follow the reverse path which was previously established during soft reservation. This can be hard to achieve because of following reasons.

1. Path established during soft reservation may be outdated.
2. Routes may not be symmetric which means a soft reservation is sent along one path but the corresponding hard reservation message takes another way.

5. Remedial solutions to address ASAP scheme limitations

The purpose of these solutions is to redesign ASAP signaling protocol in order to provide fast flow restoration and short reaction time to topology changes. To achieve this objective following new mechanism are developed:

- Local Repair
- Dynamic Virtual Path
- Flexible Timing Control
- Congestion Control

5.1 Local Repair

Purpose: In mobile ad hoc networks, the flow path might be frequently broken, as a result of topology and routing information changes. From the point where the path is broken to the receiver end, no QoS can be guaranteed. Local repairing mechanism will help to re-establish reservation state on this path as fast as possible.

Working: A local repair [Xue J., et al, 2003] is triggered by a soft reservation message. Upon receiving a soft reservation message the node reserves some bandwidth within the specified range and update table entries as usual. Before passing the message to the next hop the node checks whether its actual hard reservation corresponds to the hard reservation specified within the received message or not. If both values are equal nothing has to be done and the soft reservation message is sent along the path. In case, if the actual hard reservation specified is smaller than its own hard reservation for that flow, the node releases additional reservation, otherwise, if the specified value is greater than its own, the node tries to allocate additional resources. Both releasing and allocation of bandwidth has to be done with respect to the range and the amount of soft reserved bandwidth for that flow.

Implementation: How this scheme is implemented is given below;

```
receiving_message (SR ) ;
s o f t _r e s e r v i ng _ an d_ up da t i n g _ t a b l e s (SR.MinBW , SR .MaxBW) ;
i f ( HardResvBW < SR . HardBW) {
        i f ( SoftResvBW > = SR. HardBW _ HardResvBW ) {
                h a r d _ r e s e r v i n g (SR. HardBW _ HardResvBW ) ;
```

```
        }
        e l s e i f ( SoftResvBW > 0 ) {
                h a r d _ r e s e r v i n g ( SoftResvBW ) ;
        }
}
e l s e {
        ha rd _r e l e a s i n g ( HardResvBW _ SR . HardBW ) ;
}
```

5.2 Dynamic Virtual Path

Purpose: Unsymmetrical routes and links causing difficulties for hard reservation messages to follow the reverse path, established during soft reservation can be addressed by dynamic virtual path mechanism [Xue J., et al, 2003].

Working: This can be achieved by adding a hop count field to the hard reservation message. This hop count field will be refreshed on every node and incremented on every other. In addition to this, a global MAXHOPCOUNT constant is defined. If, this hop count reaches MAXHOPCOUNT on a certain node the hard reservation message is not processed any further and will be sent to the source node specified within the flowID of the message. As the hard reservation message will be received by the source node, the node updates its QoS entry for that flow by setting actualSoft and actualHard to the corresponding values within the message. The subsequent soft reservation message will then trigger a local repair in order to switch any soft reserved bandwidth into hard reservation on all the remaining nodes.

5.3 Flexible Timing Control

Purpose: In ASAP, there are two timing parameters which are related to the efficiency and performance of ASAP; the SR (Soft Reservation) sending interval and the soft state time-out period. The interval between SR messages is critical in determining the speed with which ASAP adapts changes. This time interval may cause excessive messages and high processing load. The selection of this interval can be made flexible according to the dynamics of the network.

Working: Larger intervals are preferable when the network is relatively stable and has enough resource. A smaller interval is helpful in case of high mobility of network nodes. The soft state time-out period must be coherent with the SR interval. The SR interval should not be smaller than the timeout period in order to keep the flow path alive. And the timeout period should not go too far beyond the SR interval. The reason is that too large a soft state time-out will keep reservations on, out of date or broken paths alive for a long time, thereby preventing other traffic from accessing those resources.

5.4 Congestion Control

Purpose: In mobile ad hoc networks for better bandwidth management, congestion control is very important. When congestion occurs it needs to be detected so that transmission rate can be controlled accordingly.

Working: In MANETs, the queue length is not a valid indication of congestion. We need to detect congestion in a node's neighborhood by monitoring the wireless channel utilization ratio. This information can be obtained by defining a threshold value, and when the channel utilization ratio is larger than this threshold, we can assume that this node's neighborhood is getting congested.

By examining the channel usage of a node, we are able to take into account the activities of both the node itself and its surrounding neighbors and therefore obtain a good approximation of the bandwidth usage. Depending on the bandwidth usage we can increase or decrease transmission throughput so that the network never get congested.

5.5 Conclusion

The challenge of bandwidth management in mobile ad hoc networks has been addressed in this paper. After investigating some of the existing QoS models and protocols it was concluded that none of these technologies are able to meet the MANET requirements. Therefore a new signaling protocol referred as ASAP has been explained. After critical analysis of ASAP, two major problems of using ASAP in MANETs were determined with each one's explanation. Then the remedial solutions to address these limitations has been described .that are based on well-defined design principles. A local repair mechanism is used to overcome the problems related to the inertness of two pass based reservations like the big latency during flow restorations. The reverse path problem is addressed by using dynamic virtual paths while flexible timing control and congestion control helped in better bandwidth management and efficient usages of network resources.

6. References

Blake, S. (1998), *An Architecture for Differentiated Services*, IETF RFC2475, http://www.ietf.org/rfc/rfc2475.txt

Braden, R., Clark, D., and Shenker, S., (1994), *Integrated Services in the Internet Architecture* — An Overview, IETF RFC1633, http://www.ietf.org/rfc/rfc1633.txt

Broch, J., Johnsson, D. B. Maltz , D. A. (2001), *The Dynamic Source Routing for Mobile Ad Hoc Networks*. Internet draft, draft-ietf-manet-dsr-05.txt.

Charles. E. Perkins, E. M. Royer, and S.R. Das, (2000), *Quality of Service for Ad Hoc on Demand Distance Vector Routing*, IETF Internat Draft, draft-ietf-manet-aodvqos-00.txt.

Clausen, T., Dearlove, C., Jacquet, P. (2009), *Optimized Link State Routing Protocol,* IETF MANET Working Group Internet Draft: http://tools.ietf.org/html/draft-ietf-manet-olsrv2-08

Dharmaraju, D.; Roy-Chowdhury, A.; Hovareshti, P.; Baras, J.S.; (2002) *INORA - a unified signaling and routing mechanisms for QoS support in mobile ad hoc networks.* Parallel Processing Workshops, International Conference, pages86-93

Farkas, K.; Budke, D.; Plattner, B.; Wellnitz, O.; Wolf, L.; (2006), *QoS Extensions to Mobile Ad Hoc Routing Supporting Real-Time Applications*, Computer Systems and Applications, 2006. IEEE International Conference, pages 54 – 61.

Ferguson, Paul, and Huston, Geoff. (1998), *Quality of Service: Delivering QoS on the Internet and in Corporate Networks*. New York: John Wiley & Sons, pages 1-6.

Grilo, A.; Macedo, M.; Sebastiao, P.; Nunes, M.; (2005) *Stealth optimized fisheye state routing in mobile ad-hoc networks using directional antennas.* Vehicular Technology Conference,. IEEE 61st, Volume 4, pages 2590 – 2596.

Haas, Z.J. and Pearlman, M.R. (1999), *Determining the optimal configuration for the zone routing protocol*, IEEE Journal, Volume 17, Issue 8, pages 1395 – 1414.

Jianbo Xue (2003), *Adaptive QoS-supporting architecture for real-time application in wireless ip network,* http://www.iks.inf.ethz.ch/asap/asap03.pdf

Marco Conti, (2003), *Body, Personal, and Local Ad Hoc Wireless Networks*, The handbook of ad hoc Wireless networks, CRC press, Chaper 1, page 17

Perkins C., and Bhagwat P. (1994) *Highly Dynamic Destination-Sequence Distance Vector Routing (DSDV) for Mobile Computers*, ACM SIGCOMM Computer Communication Review, Vol. 24, pages 234-244.

Royer, E.M. and Toh, C.-K. (1999) *A Review of Current Routing Protocols for Ad Hoc Mobile Wireless Networks*, IEEE Personal Communications, Volume 6, Issue 2, pages 46 – 55.

Toh, C-K. (1997), *Associativity Based Routing For Ad Hoc Mobile Networks*, Wireless Personal Communications Journal, Vol. 4.

Vincent D. Park and M. Scott Corson, (1997), *A highly distributed routing algorithm for mobile wireless networks*, in Proc. IEEE INFOCOM'97, Kobe, Japan.

Xiao H., Chua Kee Chaing, Seah Khoon Guan Winston (2003), *Quality of Service Models for Ad Hoc Wireless Networks*, pages 499-453

Xiao H., WKG Seah, A Lo, KC Chua, (2000), *A flexible quality of service model for mobile ad-hoc networks.* Vehicular Technology Conference Proceedings, Tokyo. IEEE 51st, Vol.1, pages 445-449.

Xue, J.; Stuedi, P.; Alonso, G.; (2003), *ASAP: An adaptive QoS protocol for mobile ad hoc networks*, 14th IEEE Proceedings, Vol. 3, pages 2616 – 2620.