A Glance at Quality of Services in Mobile Ad-Hoc Networks

Zeinalipour-Yazti Demetrios (*csyiazti@cs.ucr.edu*) *

Department of Computer Science
University of California - Riverside
3201 Canyon Crest Dr,
Riverside CA 92507, USA

Abstract. In this report we make a review of the new but rapidly growing area of *Quality of Services (QoS)* in *Mobile Ad-Hoc Networks (MANets)*. Although the application area of MANets was initially proposed for environments of battlefield communications and disaster recovery the evolution of the Multimedia Technology and the commercial interest of companies to reach widely civilian applications has made QoS in such networks an avoidable task. In the last recent years we have also seen the efforts of bringing QoS in wire-based networks becoming a reality. Many ideas inherited from the wire-based networks can be ported to MANets if they take into consideration the *Bandwidth Constrains* the *Dynamic Topology* and the *Constrained Processing and Storing Capabilities* of MANets. In this report we will have a glance at successful QoS Models and Protocols of the IP network such as *IntServ, DiffServ and RSVP* and see how they have affected the evolution of Models and Protocols in the Wireless Ad-Hoc world. The report is mainly concentrated on QoS Models, Signaling Protocols and QoS Routing that have been proposed for MANets. Although QoS MAC Protocols are equivalently important to achieve a complete QoS Architecture they are out of the scope of this report.

1 Introduction

Ad-Hoc network is a dynamic multihop wireless network that is established by a set of mobile nodes on a shared wireless channel. Each mobile host performs *local broadcasts* in order to identify its existence to the surrounding hosts. Surrounding hosts are nodes that are in close proximity to the transmitting host. In that way each mobile hosts becomes potentially a router and it is possible to dynamically establish routes between itself and nodes to which a route exists. Ad-Hoc Networks were initially proposed for military applications such as battlefield communications and disaster recovery, but the evolution of the Multimedia Technology and the commercial interest of Companies to reach widely civilian applications made QoS in MANets an area of great interest. Although much progress has been done in QoS for wire-based networks, there are still many problems. Moreover the problems that exist for QoS in wire-based networks, MANets are facing three new constraints. These constrains are: a)the *Bandwidth Constrains*, since a MANet has usually poor bandwidth resources, b) the *Dynamic Topology* of the MANet, since nodes are continually changing location, connecting and disconnecting from the network making connections many times unreliable, and c)the *Limited processing and Storing capabilities* of mobile nodes, in contrast with routers on the Internet. Due to this constrain we can't design nodes in a complex manner. Although QoS and complexity are terms that usually go together, we have to keep complexity as low as possible since this may also lead to excessive power consumption which is another problem that may arise.

The organization of the rest of this report is as follows. In Section 2 we give a definition for QoS. In Section 3 we will review quickly IP QoS. Section 4 introduces FQMM, the first QoS Model proposed of Ad-Hoc Networks. Section 5 makes a comparison between in-band and out-of-band signaling and introduces INSIGNIA, the first signaling protocol desinged solely for MANets. In Section 6 we will talk about QoS Routing and explain various details of QoS for AODV, which is a sucessful routing protocol for MANets. Finally, we conclude in section 7 with conclusions.

2 What is Quality of Services?

QoS is a term widely used in the last recent years in the area of wire-based networks. QoS stands for Quality of Services and the truth is that there is much debate on what exactly QoS is supposes to mean. Most vendors implement QoS protocols having in mind specific scenarios and taking into consideration different parameters, network topologies and variables. The United Nations Consultative Committee for International Telephony and Telegraphy (CCITT) Recommendation E.800, has defined QoS as: *"The collective effect of service performance which determines the degree of satisfaction of a user of the service"*. This is a widely accepted definition since it doesn't makes any reference to any minimum characteristics, such as `Bandwidth or Delay`, or mechanisms, such as `Admission Control, SLA, Signaling Protocol.`

* Final Research Report for CS260 - Seminar in Mobile Ad Hoc Networks, Fall 2001. Details available online on http://www.cs.ucr.edu/~csyiazti/cs260.html, November 19, 2001

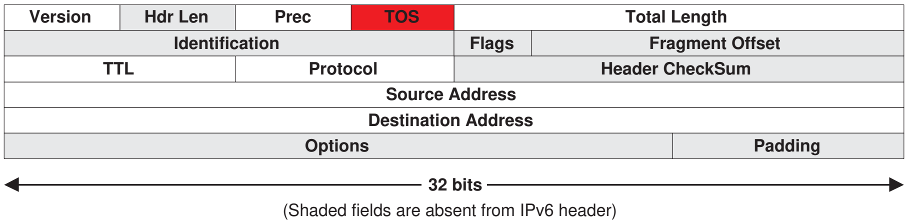

Fig. 1. IPv4 Packet.

3 IP QoS at a glance

Providing QoS in wire-based networks can generally be achieved with the *over-provisioning of resources* and with *network traffic engineering*[6]. With "over-provisioning" we add plentiful capacity in our network making it more resistant to demanding multimedia applications that may run on the top of our network. Resources usually upgraded include data links (e.g. fiber optic), upgrade of routers and network cards. The advantage of this approach is that is easy to be implemented since the upgrade can be done gradually. The main disadvantage of this approach though is that again we remain at 1 service class, since all users have the same priority, and the network may become unpredictable during times of bursting and peak traffic. The main idea of "network traffic engineering", is to classify our users (or their applications) in service classes and handle each class with a different priority. This approach overcomes the problems of the previous approach since everybody is following some rules within the network. Network traffic Engineering has two approaches for achieving QoS which are complementary, designed for use in combination for different network contexts. These are a) *Reservation-based Engineering* and b) *Reservation-less Engineering*.

In Reservation-based Engineering, network resources are apportioned according to an application's QoS request and subject to bandwidth management policy. This approach was used in ATM (Asynchronous Transfer Mode) and is today the method of achieving QoS in RSVP-IntServ.

In Reservation-less Engineering on the other hand no reservation is done within the network. QoS is achieved by the addition of "smart" mechanism into the network such as *Connection Admission Control (CAC), Policy Managers, Traffic Classes, Queuing Mechanisms.* CAC controls which nodes can access the network and it will assure to a node that once it is granted access to the network, it will be served with the QoS parameters it is requesting. Policy Managers ensure that no node will violate the type of service it is pre-assigned. Traffic Classes, such as assured, controlled-load or best-effort services, differentiate the processing priority of data packets. This approach is used in today's DiffServ (Differentiated Services) QoS Architecture where a small bit-pattern in each packet, in the IPv4 TOS octet or the IPv6 Traffic Class octet (see Figure 1), is used to mark a packet to receive a particular forwarding treatment, or per-hop behavior, at each network node. Queuing mechanisms are responsible for dropping the packets with the lowest priority in the case of congestion or to provide explicit feedback to nodes in order to avoid congestion.

4 QoS Models for MANets

The QoS Model specifies the architecture which will enable us to offer services that operate better than the current "best effort" model that exists in MANets. This architecture should take into consideration the challenges of Mobile Ad Hoc Networks e.g. dynamic topology and time-varying link capacity. We have already described the basic concept of QoS Models of the current Wire-based Internet (IntServ/RSVP and DiffServ). Below we analyze the reasons why the above models are not appropriate for MANets and then we introduce the first proposed QoS model for MANets, namely FQMM, which was proposed in [1].

IntServ/RSVP model is not suitable for MANets due to the resource limitations in MANets. There are several factors which prohibit the use of that model over a MANet. 1) *Huge storage and processing overhead* for each mobile host, since they have to build and maintain such information. Moreover the amount of state information increases proportionally with the number of flows, which is also a problem with the current QoS Internet, but which will be fortunately solved with the aggregation of state information on the core routers (DiffServ). 2) The RSVP reservation and maintenance process is a *network consuming procedure*. Thus RSVP signaling packets will grapple with the data packets for resources and more specifically for bandwidth. This happens because RSVP is an out-of-band signaling protocol. In section 5.1 we will explain why in-band signaling protocols are more appropriate for MANets. 3) In order to have a complete QoS Model mechanism such as Connection Admission Control (CAC), classification and scheduling must be provided. These mechanisms though require again a respectable amount of network resources which are usually not available in MANets.

DiffServ on the other hand is a lightweight model for the interior routers since individual state flows are aggregated into set of flows (see Figure 2). This makes routing a lot more easily in the core of the network. Thus this model could be a potential model for MANets. In MANets though there is no clear definition of what is a core, ingress or egress router because of the dynamic topology of the network. This drawback would again take us back to the IntServ model where several separate flow states are maintained, causing a heavy storage cost in every node. Moreover the concept of the *Service Level Agreement (SLA)*, defined in Wire-based QoS models is not more applicable. SLA basically defines the contract between the customer (e.g. ISPs) and the clients. The charging model in MANets has still a long way to go and could be characterized as a "gray area". Generally speaking if someone acquires QoS parameters and he pays for such parameters then of course there must be some Entity which will assure or at least give him assured parameters of service. In a completely ad-hoc topology where there is no concept of service provider and client and where there are only clients it would be quite difficult to innovate QoS, since there is no obligation from somebody to somebody else making QoS almost infeasible.

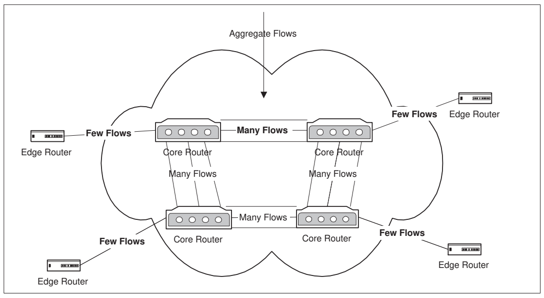

Fig. 2. In DiffServ interior flows are aggregated.

Flexible QoS Model for MANets (FQMM)[1], is the first QoS Model proposed for MANets in 2000 by Xiao et al. The idea of the paper is to combine knowledge from the solutions offered in the wire-based networks and apply them to a new QoS Model which will take into consideration the characteristics of MANets. The basic idea of that model is that it uses both the *per-flow state* property of IntServ and the *service differentiation* of DiffServ. In other words, this model proposes that highest priority is assigned per flow provisioning and other priority classes are given per-class provisioning. This model is based on the assumption that not all packets in our network are actually seeking for highest priority because then this model would result in a similar model with IntServ where we have per-flow provisioning for all packets. The FQMM hybrid model defines three types of nodes, exactly as in DiffServ a) ingress, b) core and c) egress (see Figure 3). The difference though is that in FQMM the type of a node has nothing to do with its physical location in the network, since this wouldn't make any sense in a dynamic network topology. A node is characterized as ingress if it is transmitting data, core if it is forwarding data and egress if it is receiving data.

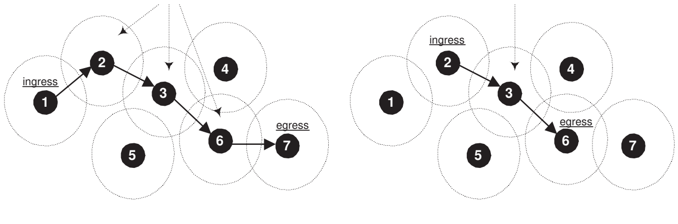

Fig. 3. In the first scenario nodes 2,3,6 act as interior (core) nodes and in the second scenario 2 only node 3 act as a core node.

5 QoS Signaling

5.1 In-band VS Out-of-Band Signaling

Signaling is used in QoS networks to reserve and release resources. In this section we are going to discuss general QoS Signaling terminology and then we are going to describe INSIGNIA, the first signaling protocol designed exclusively for MANets. In order to achieve "correct" QoS Signaling there are two prerequisites: a) Reliable transfer of signals between routers and b) Correct interpretation and activation of the appropriate mechanism to handle the signal. In simple words that means that the signaling that is sent by routing nodes within our network has to be understandable and implemented by the rest nodes. The transfer of signals between routers can be divided into "in-band signaling" and "out-of-band signaling". In-band signaling refers to the fact that any network control information is encapsulated into the data packets making the signaling approach easy and "lightweight". Out-of-band signaling on the other hand refers to the approach that uses explicit control packets. This approach is characterized[3] "heavyweight" because extra information is carried in the network and consumes more network bandwidth. Moreover in out-of-band signaling systems, signaling packets must have higher priority than data packets in order to achieve on-time notification. This approach can lead to a complex system though where performance will degrade substantially. On the other hand this approach is characterized[3] as more scalable since the control messages don't rely on the transmission of data packets. Furthermore the supported services can be rich and powerful. RSVP is an example of out-of-band signaling. In MANets, bandwidth and power constrains is an important issue. MANets can't tolerate complex signaling protocols. We instead seek for a lightweight and simple signaling protocol that can be afforded by the MANet architecture. The direct mapping of existing signaling protocols is also not feasible. RSVP is the de-facto signaling protocol for IntServ and although it can perform relatively well in small-scale wire-based networks, it does not take into consideration the distinct characteristics of MANets. In RSVP bandwidth and power constrains are not a point of concern. Furthermore, it is not adaptive for time-varying topology because it has no mechanism to rapidly respond to the topology change in MANets. We have to recall that RSVP is a "soft-state" protocol where resources are released if a signal does not arrive in intervals from the time of the reservation. Although the soft-state property makes RSVP a robust protocol, since it ensures that no resources will remain allocated, the definiton of the *interval* is a trade-off between performance and adaptation to topology changes.

5.2 INSIGNIA Signaling Protocol

As mentioned before INSIGNIA is the first signaling protocol designed explicitly for MANets in 1998 by Ahn et al.[2]. INSIGNIA supports fast flow reservation, restoration and adaptation algorithms that are specifically designed to deliver adaptive real-time service in MANets environments. INSIGNIA can be characterized as an *in-band RSVP signaling protocol* since it encapsulates control signals in the IP option of every IP data packet (see Figure 4a), which is now called *INSIGNIA option*. INSIGNIA also maintains flow state information for the real-time flows on an end to end basis, informing the source nodes for the status of their flow. This configuration makes INSIGNIA a "lightweight" signaling protocol appropriate for MANets.

INSIGNIA Option field Figure 4a illustrates the INSIGNIA option which is encapsulated in an IP packet. *Reservation Mode* is a 1 bit field which indicates if this packet is currently seeking for a reservation (REQ) or if this package has already reserved resources (RES). In the case it is in a REQ mode the package is forwarded to the INSIGNIA module for further treatment. The INSIGNIA module may then either grant the resources, which would mean that the Service Type field would be set to real time (RT), or deny granting the resources, which would mean that the Service Type field would automatically degraded to best effort (BE). In both cases the packet will be forwarded to the next immediate node which means that if there is a route to the destination host, finally the INSIGNIA option will indicate either BE or RT. The bandwidth request is a 16 bit field which indicates the minimum and maximum amount of bandwidth requested by a package. Based on this field the INSIGNIA module may determine what amount of resources should be given away. The bandwidth Indicator field is again a single bit field which is used by the receiver to determine if the Maximum requested bandwidth has been satisfied.

Bottleneck Node in INSIGNIA Figure 4b illustrates the case where there is bottleneck node during the flow reservation process. We can see that the INSIGNIA option field has degraded from Real Time/Maximum Bandwidth (RT/MAX) to Real time/ Minimum Bandwidth (RT/MIN). It is also important to mention that if M2 couldn't meet the Request criteria's the out coming package could have a Best Effort/ Minimum Bandwidth (BE/MIN) field.

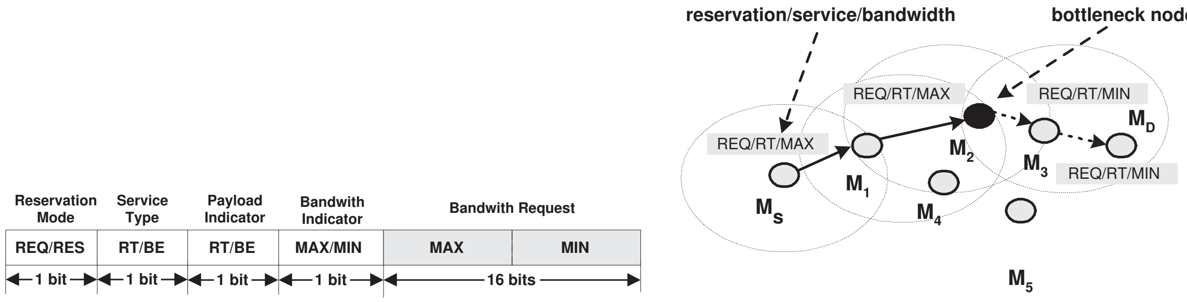

Fig. 4. a) INSIGNIA IP Option fields. b) A flow setup & the use of Bandwidth Indicator.

The INSIGNIA module is involved every time an IP packet, which contains the INSIGNIA option, is received. In coordination with the Admission Control module, it then allocates bandwidth to the flow if the resource requirement can be satisfied. Otherwise if the resource request can't be satisfied the packet is set to best effort service and forwarded if necessary.

It is important to keep in mind that INSIGNIA is just the signaling protocol and that there is a necessity to moreover include a routing protocol, such as DSR[7], AODV[8] or TORA[9], which will track changes in the ad-hoc topology and which will make updates to the routing table of each node, the need of admission control module which will allocate the requested bandwidth after it determines that such resources are available and the need for other components such as packet forwarding module, packet scheduling module and medium access controller module. As a whole INSIGNIA is an effective signaling protocol since it is an in-band signaling protocol and since the allocation of resources is "soft-state". The major drawback of INSIGNIA though is that flow state information should be kept in the mobile hosts, which could become a scalability problem as the number of flow states increases. Moreover INSIGNIA design basically enables the existence of only two classes of services real time (RT) and best effort (BE).

6 QoS Routing in MANets & QoS for AODV

QoS Routing in MANets is an essential component to realize a complete QoS MANet Architecture. The QoS Routing procedure can inform a source node of the bandwidth and QoS availability to destination node in the network. This knowledge enables the establishment of QoS connections within the network and the efficient support of real-time multimedia traffic. There are generally many proposed solutions for QoS routing in MANets such as [10][12][5]. In this section we are going to explore the QoS version of AODV, which was proposed in 2000 by Perkins et al.

The AODV is an on-demand routing protocol[8] which is based on the idea both of DSDV[11] and DSR[7]. AODV minimizes the number of required broadcasts by creating routes on an on-demand basis, as opposed to maintaining a complete list of routes as in the DSDV algorithm. It inherits this property from the DSR (Dynamic Source Routing) protocol. We already know that in AODV, when a source node desires to send a message to some destination node and do not already have a valid route to that destination; it initiates a Path Discovery process to locate the other node. It broadcasts a *route request (RREQ)* packet to its neighbors, which then forward the request to their neighbors, and so on, until either the destination or an intermediate node with an "updated" route to the destination is located. AODV also utilized *destination sequence* (from DSDV) to ensure that all routes are loop-free and contain the most updated route information.

The main idea of making AODV QoS-enabled is to add extensions to the route messages (`RREQ, RREP`) during the phase of route discovery. A node which receives a RREQ with a *quality of service extension* must be able to meet the service requirement in order to either rebroadcast the RREQ (if doesn't have an updated route in its cache), or unicast a RREP to the source. If, after establishment of such a route, any node along the path detects that the requested Quality of Service parameters can no longer be maintained, that node must originate an `ICMP QOS_LOST` message back to the source (that had originally requested the now unavailable parameters.

As we have mentioned before several extensions are needed in the routing table structure and the RREQ and RREP messages for supporting QoS routing. Below we are describing the route table extensions as well as the RREQ and RREP extensions described in the Internet Draft for QoS for AODV. The "plain" AODV router table[8] contains the following fields *Destination Sequence Number, Interface, Hop Count,* *Next Hop, List of Precursors*. Moreover to the above router table fields, QoS for AODV defines 4 more elements, which are added to the properties of each particular route. These extensions are *Maximum Delay, Minimum Available Bandwidth, List of Sources Requesting Delay Guarantees* and *List of Sources Requesting Bandwidth Guarantees.*

Maximum Delay, indicates the maximum number of seconds allowed for a transmission from a source (or from an intermediate node forwarding the RREQ) to the destination. Every time a node receives a RREQ it subtracts (from the Delay indicated in the RREQ) the NODE_TRAVERSAL_TIME, which is the time required by this node to process the RREQ. The NODE_TRAVERSAL_TIME is by default set to 40 milliseconds, but could have a different value, in the case this node has more or less processing power. If the NODE_TRAVERSAL_TIME is bigger than the delay time indicated in the RREQ the node will simply discard the RREQ and not process any further. The graphical representation in Figure 5 shows how RREQ1 would be forwarded by the intermediate (core) routers during the path discovery process. At every step the delay field in the RREQ message is reduced by the Traversal_Time of the router. At the end egress D will reply a RREP message which will have a starting delay value of 0. This delay value will be added to the Traversal_time of each node and registered (cached) in the Routing Table for future RREQs. The caching of the delay value will make the future discovery of that route a trivial task. So for example a future RREQ2 done by another node A will be directly dropped by core node B since the demanded delay (10ms) can't be met (since it is 80ms).

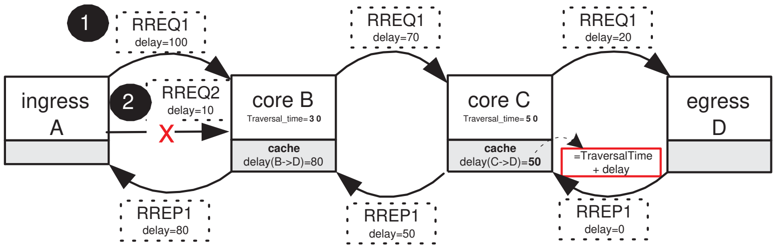

Fig. 5. RREQ forward or drop is based on the delay demanded.

Of course in the future a node, such as core C, may have increased load which would change it NODE_TRAVERSAL_TIME from 50ms to 100ms. This change would affect all depending nodes such as B and A. For this reason node C will forward an ICMP QOS_LOST message to all potentially nodes affected by the QoS parameter. This is also the reason why each node had initially stored a list of depending nodes, the "List of Sources Requesting Delay Guarantees". The ICMP QOS_LOST message is quite short and it is sent recursively to all nodes affected.

The main disadvantage of the QOS_LOST message approach, and which is not mentioned anywhere in the "QoS for AODV" internet draft, is that there are no reservations which makes the protocol handling violation of QoS parameters an a' posteriori approach instead of reserving the "promised" parameters which would be an a' priory approach. In my personal point of view, a node should not take over more jobs after identifying that it would violate it existing QoS parameters.

Minimum Available Bandwidth, is a field which indicates the requested amount of bandwidth for a specific link (route). Every time a node receives a RREQ it must compare its available link capacity with the capacity of bandwidth requested in the RREQ. If the requested bandwidth is not available then the node will again, as the delay example, discard the RREQ and not process any further. If the bandwidth is available then the request will process until the egress router D is reached. At that point the egress router D, will respond with a RREP message which will be initialized with a bandwidth value equal to infinitive (a very large number). Each node forwarding the RREP compares the bandwidth field in the RREP and its own link capacity and maintains the minimum of the two in the Bandwidth field of the RREP before forwarding the RREP. This bandwidth value will be registered (cached) in the Routing Table bandwidth value for future RREQs. The caching of the bandwidth value will again make the future discovery of that route a trivial task. We can see again that RREQ2 (which is route discovery message from A to D) request won't be satisfied because its request is 80Kbps exceeds the available 50Kbps which is now cached in nodes' B cache and hence be directly dropped by core node B.

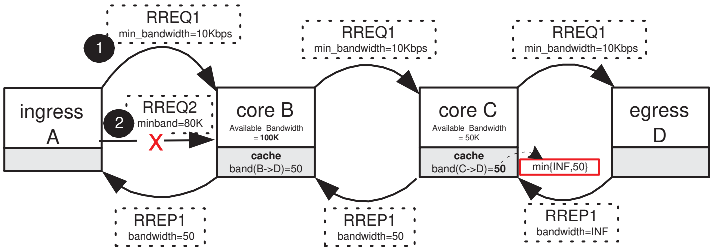

Fig. 6. RREQ forward or drop is based on the bandwidth demanded.

As in the Delay case, a node in the future may have a drop in link capacity which will lead in the generation of an ICMP QOS_LOST message to all potentially nodes affected by the QoS parameter. The list of nodes that are affected by this property is now stored in the "List of Sources Requesting Bandwidth Guarantees.

7 Conclusions

QoS in MANets is a new but rapidly growing area of interest. This great research and market interest is firstly because of the rising popularity and necessity of multimedia application and secondly because of the potential commercial usage of MANets. Thus QoS support in MANets has become an unavoidable task. In this report we have tried to give a brief introduction to QoS issues in networks. I have started with a review of the current trends and solutions given in the wire-based IP network, where much more progress has been done. Although the knowledge and ideas of the QoS Models, Signaling, Routing protocols can not be directly mapped to MANets, because of the bandwidth constrains and dynamic topology of such networks, many ideas have been adapted for these networks. After that we have analyzed main issues on QoS Models, Signaling, Routing and Mac protocols. These issues are complicated because most experimentation and simulations don't reveal accurate knowledge [4] since the experimentation is done without taking into consideration various real conditions, such as quality of radio links and availability of the nodes and their resources. Much more work remains to be done in this area until it reaches the public in a simple form.

References

1. H.Xiao, K.Chua, W.Seah and A.Lo : A Flexible Quality of Service Model for Mobile Ad-Hoc Networks, (http://www.cs.ucr.edu/~csyiazti/downloads/cs260/qosModels/xiao.pdf).
2. Lee and Campbell : INSIGNIA: In-Band Signaling Support for QoS In Mobile Ad Hoc Networks, (http://www.ececs.uc.edu/~guptanis/research/papers/QoS/insignia98.ps).
3. Kui Wu and Janelle Harms : QoS Support in Mobile Ad Hoc Networks, (http://www.cs.ualberta.ca/~wkui/research/QoSReview.ps).
4. Satyabrata Chakrabarti and Amitabh Mishra : QoS Issues in Ad Hoc Wireless Networks, (http://www.ececs.uc.edu/~guptanis/research/papers/QoS/QoSAdHoc.Mishra.pdf).
5. C.R.Lin and J.Liu : QoS Routing in Ad Hoc Wireless networks, IEEE J. Sel. Areas Commun., vol. 17 (8), p. 1426, August 1999., (http://www.ececs.uc.edu/~guptanis/research/papers/QoS/QoSAdHoC.LinCR.pdf).
6. Demetris Zeinalipour, Stella Aristeidou, Sofia Kazeli : IP Quality of Services (in Greek), (http://www.cs.ucr.edu/~csyiazti/downloads/papers/ip_qos/papers/ip-qos.pdf), 1999.
7. J. Broch, D.B. Johnson and D.A Maltz : The Dynamic Source Routing Protocol for Mobile Ad-Hoc Networks, IETF Internet Draft, draft-ietf-manet-dsr-01.txt, December 1998 (Work in Progress).
8. C.E. Perkins and E.M Royer : Ad Hoc On Demand Distance Vector (AODV) Routing, IETF Internet Draft, draft-ietf-manet-aodv-02.txt, November 1998 (Work in Progress).
9. V.Park, S Corson : Temporally-Ordered Routing Protocol (TORA), IETF Internet Draft, draft-ietf-manet-tora-epc-00.txt, November 1997 (Work in Progress).
10. E.M Royer and C.E. Perkins : Quality of Servie for Ad Hoc On Demand Distance Vector (AODV) Routing, IETF Internet Draft, draft-ietf-manet-aodvqos-00.txt, July 2000 (Work in Progress).
11. C.E. Perkins : Highly Dynamic Distance Vector (DSDV) Routing for Mobile Computers, ACM SIGMOD 1994 Conference on Communications Architectures, Protocols and Applications p224-234 1994.
12. R.Sivakumar, O.Sinha, V.Bharghavan : CEDAR: a Core-Extraction Distributed Ad Hoc Routing Algorithm, IEEE Journal on Selected Areas in Communications, Special Issue on Ad Hoc Networks, Vol17, No8, 1999