Overview of Cross-Layer QoS Support in Mobile Ad Hoc Networks*

LAMIA ROMDHANI and CHRISTIAN BONNET

Institut Eurécom 2229 route des Crêtes, BP, 193

06904 Sophia Antipolis Cedex.

Firstname.Name@eurecom.fr

Abstract: Mobile Ad hoc NETworks (MANETs) have several new characteristics that should be carefully studied to be able to define and introduce Quality of Service (QoS) support in these networks. Moreover, there are many difficulties to use directly the works that have been done for wired network to provide QoS enhancement. In fact, the frequent changes in network topology, as well as the lack of wireless resources, makes mobile ad hoc networking and consequently QoS support in these networks a challenging task. Indeed, the first drawback of MANETs is that they are often considered as a network, so its characteristics is not considered at the application layer. On the one hand, the question is what is the QoS model that can optimize the network resource utilization while satisfying application requirements? On the other hand, what are the mechanisms that can be offered to the applications to adapt to available resources?

In this report, we discuss the issues to deploy the quality of service models that have been proposed for wired networks in MANETs. Moreover, we investigate the QoS models dedicated for MANETs. Then, we present some features that have been introduced in order to improve the performance of MANET routing protocols and medium access mechanisms.

Keywords: mobile ad hoc networks, quality of service model, signaling, cross-layer design, routing, MAC.

1 Introduction

The goal of MANET architecture is to provide communication facilities between end-users without any centralized infrastructure. It is also possible to have an access to some hosts in a fixed infrastructure depending on the kind of mobile ad hoc network available. Some scenarios where an ad hoc network could be used are business associates sharing information during a meeting, military personnel relaying tactical and other types of information in a battlefield, and emergency disaster relief personnel coordinating efforts after a natural disaster such as a hurricane, earthquake or flooding.

* *Eurecom Research is particularly supported by its industrial partners: ASCOM, Swisscom, France Telecom, la fondation Cegetel, BouyguesTelecom, Thales, Stmiscroelectronics, Hitachi Europe and Texas Instruments.*

QoS support is critical for wireless home networking, video on demand, audio on demand and real-time Voice over IP applications. Time-bounded applications such as audio and video conference typically require a minimal bandwidth, bounded delay and low jitter, but can tolerate some losses. Providing QoS support in MANETs is a challenging task due to the frequent changes in network topology as well as the lack of wireless resources caused by the high collision rate, high error rate, low capacity, etc. In this report, we present the difficulties to apply directly wired QoS models in MANET. Moreover, we discuss some QoS enhancement mechanisms introduced for MANET including resource reservation [10, 14], QoS routing [13, 9], and service differentiation based-MAC layer [18, 19, 25]. In fact, fixed wired networks have relatively stable QoS characteristics for a given connection . However, radio-based connections may be affected by link failure more often. Therefore, it is essential to capture the aforesaid characteristics to identify the quality of links. Furthermore, the routing protocols that support QoS must be adaptive to cope with the considerable variations in link quality experienced during a connection.

The report is organized as follows. Section 2 presents the difficulties to deploy the current QoS model described for wired network, namely IntServ and DiffServ. In Section 3, we describe a variant of these two models that have been proposed for MANETs . An overview of some routing features in MANET will be presented in Sections 4. In Section 5, we investigate the wireless medium access mechanism characteristics and we outline the most QoS schemes that have been proposed to introduce service differentiation-based MAC layer. Finally, Section 6 provides concluding remarks.

2 Wired QoS models and MANETs

The presence of mobility implies that links make and break often and in an indeterministic fashion. Consequently, QoS models described for wired networks are insufficient for such networks [5, 8]. Integrated Services (IntServ) [17] and Differentiated Services (DiffServ) [20] are the two basic architectures proposed to deliver QoS guarantees in the Internet. Below, we discuss the different constraints to introduce these two models in MANET.

2.1 Integrated Services

IntServ [17] model is implemented with four main components: the signaling protocol, the admission control routine, the classifier, and the packet scheduler [3]. This architecture allows sources to reserve resources, using a signaling protocol, in order to provide special QoS for specific flows. However, the resource requirements (computational processing and memory consumption) for routing and running per-flow resource reservations on routers increase in direct proportion to the number of separate reservations that need to be accommodated. Moreover, in MANET every host should implement in addition to its traditional functionalities, the different router components. These characteristics are very difficult to support in wireless networks, since they cost a huge storage and processing overhead for the mobile node whose storage and computing resources are scarce. In addition, IntServ was introduced to provide guaranteed service which implies assured level of bandwidth, bounded end to end delay and no queuing losses for conforming packets. Hence, these metrics are guaranteed since the data path, which is selected for a specific QoS flow, is so maintained. However, the possible mobility of nodes that participate at route forward data and so they become unreachable, will decrease significantly the performance of QoS services. Thus, if we introduce another mechanism to repair the route, when this later is no more available, we will surely increase the algorithm complexity.

In addition, IntServ guaranteed service traffic is subject of traffic policing and traffic shaping. Traffic policing controls the incoming traffic which must conform to Traffic specification (Tspec). The non conforming traffic is treated as best effort. In the Internet, this policing is executed at network borders. Traffic shaping implies that bursty traffic is shaped in order to fit it into Tspec. So, traffic policing and traffic shaping are required, but how these concepts can be defined in MANET with its highly dynamic topology?

As a result, the introduction of an IntServ-based QoS in the MANET is a somewhat complicated challenge. In the next subsection, we investigate the different roles of signaling protocols and we describe the most features proposed to adapt signaling to MANET characteristics.

2.1.1 QoS signaling protocols

QoS signaling protocols enable hosts to reserve resources aimed to provide a specified QoS to a data flow. The QoS signaling system can be divided into two signaling approach, in-band signaling (INSIGNIA) and out-of-band signaling (RSVP, DRSVP). The in-band signaling refers to the fact that control information is carried along with data packets; the out-of-band signaling refers to the approach that uses explicit control packets.

- **ReSerVation Protocol (RSVP)**

RSVP [4] is an out-of-band signaling system. It is used by a host to request specific qualities of service from the network for particular application data streams or flows.

RSVP is not suitable for MANET since the signaling overhead of RSVP is heavy for the mobile hosts that causes a scalability problem. Moreover, the signaling control message will contend with data packets for the channel and cost a large amount of bandwidth. Furthermore, it is not adaptive for the time-varying topology because it has no mechanism to rapidly respond to the topology change in MANETs.

- **Dynamic RSVP (DRSVP)**

DRSVP is a protocol for the "Dynamic QoS" concept by extending the RSVP protocol to support QoS in MANET [10]. The following extensions and modifications have been made to the RSVP protocol to obtain this new protocol:

- An additional flow specification (flowspec) in Reservation (Resv) messages, and an additional traffic specification (tspec) in Path messages, so that they describe ranges of traffic flows.
- A "measurement specification" (mspec) to the Resv messages, which is used to allow nodes to learn about "downstream" resource bottlenecks.
- A new reservation notification (ResvNotify) message, which carries a "sender measurement specification" (smspec) that is used to allow nodes to learn about "upstream" resource bottlenecks.
- The admission control processing is modified to deal with bandwidth ranges.
- A bandwidth allocation algorithm is added that divides up available bandwidth among admitted flows, taking into account the desired range for each flow as well as any upstream or downstream bottlenecks for each flow.
- An application programming interface is described to handle bandwidth ranges.

The DRSVP is a flexible QoS reservation protocol. It uses separate queue for each flow. Despite of the new features introduced in DRSVP, several problems are not yet resolved as scalability.

- **INSIGNIA**

INSIGNIA uses an in-band signaling system that supports QoS in a mobile environment [14]. So, the INSIGNIA informations are carried in-band with the data and encoded using the IP option field in datagrams. As shown in Figure 1, the INSIGNIA IP option supports the establishment of adaptive reservation-based services. It includes five fields that are named as follow: service mode, payload

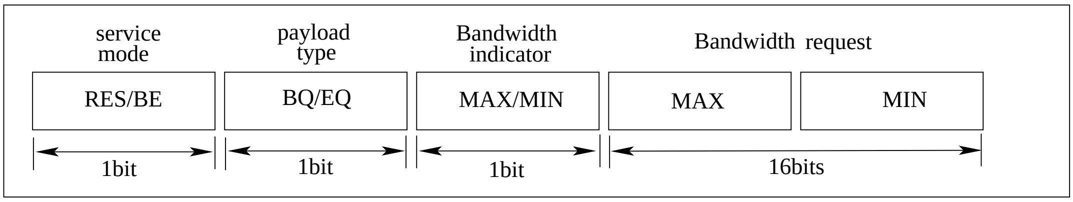

Figure 1: the INSIGNIA IP option

type, bandwidth indicator, and bandwidth request, for more detail see the IETF Internet Draft [14]. Although in-band signaling costs some bandwidth more or less, it will not contend for the transmission channel with data packets since it is included in every data packet. This feature is very important in wireless networks, where the transmission channel is shared by all neighboring hosts.

The flow state information is managed in soft-state method, that is, the flow state information is periodically refreshed by the received signaling information. Coordinating with the admission control module, INSIGNIA allocates bandwidth to the flow if the resource requirement can be satisfied. Otherwise, the flow will be degraded to best-effort service. To keep the processing simple and lightweight, INSIGNIA does not send rejection and error messages if the resource request is not satisfied. For fast responding to the changes in network topology and end-to-end quality of service conditions, the destination node periodically sends QoS statistical reports (such as loss rate, delay, and throughput) to inform the source node of the status of the real-time flows. Through this kind of feedback information, the source node can take corresponding actions to adapt the flows to observed network conditions. As a whole, INSIGNIA is an effective signaling protocol for MANETs. Coordinating with other network components (viz. routing protocol, scheduling, and admission control), INSIGNIA can efficiently deliver adaptive real-time flows in MANETs [15]. However, since the flow state information should be kept in the mobile hosts, it can cost a huge information storage when the number of flows increases. So, the scalability problem may hinder its deployment in the future.

Due to bandwidth and power constraints, keeping the signaling lightweight and simple is more important than designing a powerful but complex signaling system. At least at present, this should be one of the main principles of designing signaling system in MANETs.

2.2 Differentiated Services

DiffServ model [20] uses a per-class service differentiation for traffic transmission. This architecture avoids the scalability problem by defining a small number of per hop behaviors (PHBs) at the network edge routers and associating a different DiffServ Code Point (DSCP) in the IP header of packets belonging to each class of PHBs. Core routers use DSCP to differentiate between different QoS classes on per-hop basis. Thus, DiffServ is scalable but it does not guarantee services on end-to-end basis. This remains to be a drawback for MANET. This architecture can provide an efficient service differentiation in the Internet.

Since DiffServ is designed for fixed wired networks, we still face some challenges to implement this model in MANETs. First, it is ambiguous as to what the boundary routers in MANETs are. Intuitively, every node should have the functionality as both boundary router and interior router because the source nodes can not be predefined. This arouses a heavy storage cost in every host. Second, the concept of Service Level Agreement (SLA) in the Internet does not exist in MANETs. SLA is a service contract between a customer and its Internet Service Provider (ISP) that specifies the forwarding service that a customer should receive [20]. A SLA may include traffic conditioning rules which constitute a Traffic Conditioning Agreement (TCA) that used to re-mark traffic streams, discard or shape packets according to the traffic characteristics such as rate and burst size. How to make a SLA in MANETs is difficult because the mobility of nodes complicate the definition of a traffic rules negotiation scheme for the mobile nodes.

In table 1, we present a global cross-layer view of difficulties that are faced in ad hoc networks compared to wired networks. Moreover, we outline the most proposed solutions designed to deal with these problems. Furthermore, we give the different open issue objectives.

	Wired mechanisms	Deployment problems in ad hoc networks	Proposed solutions For ad hoc networks	Open issues
Application layer	QoS support by IntServ (RSVP) DiffServ	Frequent disconnection and reconnection Packet loss Signaling overhead Traffic rules negotiation	FEC INSIGNIA, DRSVP FQMM SWAN	Handle time-varying delay and packet loss statistics
Network layer	Table-based routing	Mobility of nodes Change of Power Route discovery Route maintenance	Proactive, reactive and hybrid protocols Hierarchical and flat structure	Minimize communication overhead Stability and scalability Fast convergence rate
MAC layer	CSMA/CA	Distributed control Hidden terminals Exposed node problem	IEEE 802.11 (CSMA/CD) Hiperlan Hybrid ARQ	Minimize collision Maximize throughput Fairness QoS
Physical layer	Reliable links High throughput	wireless link Varying received SNR	Software radio Power control Directional antenna Multiuser detection	Adaptive modulation Interaction with upper layers

Table 1: From wired networks to ad hoc networks

3 MANET QoS Models

IntServ and DiffServ are designed in the context of a static environments (fixed hosts and networks). Indeed, they require accurate link state (e.g. available bandwidth, packet loss rate , delay, and etc.) and topology information. As a result, these models cannot be applied directly to mobile environments. A variant of IntServ and DiffServ architectures have been proposed for ad hoc networks. Below, we provide a summary of the main features of these models.

3.1 Flexible QoS Model for Manet (FQMM).

FQMM is a quality of service model designed for MANETs [5]. FQMM consists of three key features: dynamic roles of nodes, hybrid provisioning and adaptive conditioning. FQMM defines three types of nodes: an ingress node which sends date, an interior node which forwards data to other nodes, and an egress node which is a destination. Obviously, each node may have multiple roles. This model selectively uses the per-flow state property of IntServ, and the service differentiation of DiffServ. On one hand, for applications with high priority, per-flow QoS guarantees of IntServ are provided. On the other hand, applications with low priorities are given per-class differentiation of DiffServ. Therefore, FQMM applies an hybrid provisioning using both IntServ and DiffServ schemes. Furthermore, the adaptive traffic conditioner used includes several components as traffic profile, marker and dropper.

In contrast to an absolute traffic profile, the traffic profile proposed in FQMM is defined as the relative percentage of the effective link capacity.

FQMM is designed for small and medium size MANET, with less than 50 nodes. Hence, the QoS performance may decrease under high load traffics. Therefore, the drawbacks related to IntServ and DiffServ remain to be a drawback in FQMM.

3.2 Two-Layered Quality of Service model (2LQoS)

The 2LQoS proposes an architecture that considers a cross layer design where Network Layer Metrics (NLMs) and application layer metrics cooperate together to provide service differentiation and QoS support[7]. The mapping between these two layer metrics is desirable because the QoS that an application requires depends on the network resources characteristics.

In this approach, the NLMs determine the quality of links in order to generate the paths with good quality. They try to evenly distribute the traffic in the network and avoid paths with a low quality regardless of the application. So, three NLMs are defined to provide a trade-off between load balancing and resource conservation: hop count, buffer level, and stability level. As all routing protocols, there are three phases that are considered: path generation, path selection, and data forwarding. During the path generation phase, buffer level and stability level are the NLMs that are considered in 2LQoS model to characterize both mobility rate and available resources reside at each node. Concave functions are used to represent the NLMs corresponding to a path given the values of these metrics for individual nodes on that path. Having known the ability of paths, an application selects exactly one path which is more likely to meet its requirements. To do this, application requirements are classified into three QoS classes: I, II, and III, with different priorities. Then, each class is mapped onto the appropriate metrics at the application layer, that is delay, throughput, and best effort respectively. This implies that applications may need to adapt to the available resources offered by the network.

The 2LQoS model address a mobile ad hoc network with a semi symmetric environment where all nodes have similar capabilities. Capabilities means transmission range, battery life, processing capacity, buffer capacity, and speed movement. However, the MANET reality is likely to be very heterogeneous. So how the network layer metrics can be redefined in order to validate the 2LQoS model in a fully asymmetric environment?

3.3 SWAN model

SWAN is a service differentiation model that is designed for stateless wireless ad hoc networks [6]. Each node in the mobile ad hoc network independently regulates best effort traffic using a rate control algorithm that is based on feedback from MAC layer. Moreover, each node uses sender-based admission control for UDP real-time traffic. Indeed, the admission control measures the local resource availability of real time traffic using the shared wireless channel that allows to listen to packets sent within radio transmission range. In addition, SWAN uses Explicit Congestion Notification (ECN) to dynamically regulate admitted real-time traffic in order to deal with network dynamics brought on by mobility or traffic overload conditions. In fact, each mobile node in the network can detect congestion or overload conditions using a periodic local bandwidth measurement of the traffic. When a node detects such a state, it starts marking the ECN bits in the IP header of the real-time packets. The destination node monitors the ECN bits and notifies the source using a regulate message. When the source node receives a regulate message, it initiates reestablishment of its real-time session based on its original bandwidth needs. To reestablish a real-time session a source node follows the same process as setting up a new session by sending a probing request toward the destination. A source node terminates the session if the estimated end-to-end bandwidth indicated in the probing response packet cannot meet its existing session requirements. Moreover, real-time sessions could be regulated or dropped due to mobility or excessive traffic overloading at wireless intermediate nodes. A novel aspect of SWAN is that it does not require the support of a QoS-capable/real-time MAC. Rather, soft real-time services are built using existing best effort wireless MAC technology.

The performance of SWAN can give a good results when a fixed route for sending packets from the source to the destination is used. However, this cannot properly reflect the performance of this model under the mobile ad hoc network characteristics as mobility and so dynamic routes. Indeed, the probing response or regulate message sent to the source by the intermediate nodes can be affected by delays. And so, when it reaches the source, the network topology and thus the bandwidth availability may have already changed significantly. Moreover, mobile ad hoc routing may result in the loss of data packets, regulate messages and probing requests/responses. This may lead to a significant degradation in the performance of the admission control mechanism proposed in SWAN. Therefore, it is extremely important to incorporate real implementation of ad hoc routing into the existing test bed in order to truly investigate and demonstrate the effectiveness of SWAN.

4 QoS routing in MANET

The goal of QoS routing is to optimize the network resource utilization while satisfying application requirements. Indeed, it is not enough to find a shortest path but also with available resources as battery and buffer. However, the time-varying topology, limited resources, and distributed aspects complicate the QoS routing protocol concept in MANETs. Routing process in MANET is based on four functionalities [7, 9]. First, path generation which is the algorithm used to generate routes between sources and destinations based on the available distributed informations in the network. Second, path selection is the process of choosing a path that meet between application requirements and available network resources. Third, data forwarding which is functionality of transporting packets along the selected route. Fourth, path maintenance which is the most important mechanism in MANET. It concerns the maintaining of the data forwarding path during the session life. Indeed, delivering services in wireless and mobile environment is mainly based on the routing protocol maintenance capabilities because new or alternative routes between sources and destinations can be recomputed during the lifetime of the ongoing sessions.

The routing protocols used in MANET can be classified into three classes:

4.1 Flooding protocols

In flooding approach, the sender broadcasts the data packets to all its neighbors. Then, each node receiving these packets forwards them to its neighbors. To avoid the forwarding of the same packet more than once, nodes use sequence numbers. The flooding for data delivery may be more efficient than other protocols when the information transmission rate is low enough that the overhead of explicit route discovery and maintenance generated by other protocols is relatively higher. This is the case where nodes exchange a small data packet no frequently but the mobility rate of the nodes is high. However, the flooding uses broadcasting and so it is hard to implement reliable broadcast delivery without significantly increasing overhead.

4.2 Proactive protocols

Proactive routing protocols require that each node in the network maintain global topology information in a routing table [9]. Thus, a route can be provided immediately when it is requested. This kind of protocols work better when the number of communications increases and the mobility of nodes decreases. However, when the number of nodes increases, storage requirement and communication overhead increase.

4.3 Reactive protocols

Reactive routing protocols have the feature on-demand. Each host computes route for a specific destination only when necessary [9]. Moreover, topology changes that do not influence active routes that do not trigger any route maintenance function. Thus, communication overhead is lower compared to proactive routing protocol. However, when the network size increases, the delay of route determination increases. Moreover, if the number of communications increases, the network may saturates.

4.4 Hybrid protocols

Hybrid routing protocols maintain partial topology information in some hosts. Routing decisions are made either proactively or reactively [11].

One important observation on these protocols is that none of them can avoid the involvement of flooding. Indeed, proactive protocols rely on flooding for the dissemination of topology update packets. Reactive protocols rely on flooding for route discovery.

Choosing between proactive, reactive or hybrid algorithm to select routes between nodes in MANET is not easy to decide. In fact, many factors should be considered as network size, mobility speed, global link stability and reliability, QoS level requested, etc. These factors affect the route establishment delay, the route maintenance capability and so the support of QoS. In Table 2, we summarize the most characteristics of the routing design issues as a function of different criterion.

	Flooding	**Proactive**	**Reactive**	**Hybrid**
The motivation behind this strategy	Simple	Always the routing info. is available	No need to periodic route updates	Meet between the advantages of proactive and reactive algorithms
Time complexity	o(D)	o(D)	o(2D)	Intra-zone:o(M) Inter-zone:Depends on the used protocol
Impact of a big network size	Huge comunication overhead	Huge storage and communication of routing info.	less effect	Depends on the zone size
Impact of the traffic load	—	+++	—	++
implementation Complexity	low	high	low	medium (it depends on the routing protocol concept)
Complexity of the algorithm	o(N^2)	o(N)	o(2N)	
Stored information in each node	Any stored information	Entire topology	A few information	Zone topology

D: the network diameter
N: the number of nodes in the network
M: the average number of nodes per zone

Table 2: Routing architecture characteristics

Typically, routing protocols in ad hoc networks use a broadcast approach to determine routes by using flooding-based algorithm. However, the flooding of route request has been shown to be very unreliable because of the hidden and exposed terminal problems [13]. To reduce the effects of broadcast problems and enhance the medium utilization, some works routing-based use unicast mechanism instead of local broadcast to find route in a domain or in zone[13]. They use the broadcast mechanism to search for the route from the domain of the source to the domain of the destination, which is the case in hybrid protocols as described in [2, 13, 11]. Note that each node within the domain has a table based-routing enable each host to select a route to reach any local destination. In other cases, only one node in the zone, called the dominator, has a view of the local topology. The knowledge of the local topology enables the dominator to decide the route to any node in the domain.

To achieve good performance, QoS routing protocols consider some QoS metrics to select the path between source and destinations. In [1, 8, 12], the authors summarize the features used by several routing protocols to provide best quality of service. These issues are based on the computation of one or several parameters such as bandwidth, delay, available buffer , link stability, and link cost. These parameters are configured based on some threshold values. Indeed, the route is selected according to the current available resources.

Moreover, it is very important to give routing packets a specific treatment and more priority at MAC layer, in order to quickly establish routes while providing the optimal available resources to the applications. Furthermore, a good routing protocol should broadcast a minimum control packets during the process of route establishment to deal with the constraint-based MAC layer. In the next section, we describe the most widely medium access mechanism used in MANET. Then, we outline several works that have been proposed to introduce service differentiation based MAC layer protocol.

5 Service differentiation enhancement-based MAC layer

QoS researches in wireless networks are still mostly pre-mature. Due to the difference between wireless and wireline networks in the underlying physical and link layer characteristics, the wireless networks rely much more on the physical and link layer capabilities. Due to the high difference between transmitted and received power levels, traditional random channel access mechanisms used in wired networks as CSMA/CD are not applicable in wireless networks. To deal with this problem, the CSMA/CA (Carrier Sense Multiple Access with Collision Avoidance) protocol is used in WLAN. The IEEE 802.11 standard is the most widely deployed Wireless Local Area Network (WLAN) standard today [21]. Its MAC layer includes a set of protocols which are responsible for maintaining order in the use of a shared medium. Distributed Coordination Function (DCF) is the basic medium access mechanism of 802.11 for both ad-hoc and infrastructure modes, which uses the CSMA/CA protocol. It can only support best-effort services, without any QoS guarantees. Thus, several enhancement mechanisms have been described to introduce QoS based on the basic scheme.

Moreover, the service differentiation based MAC layer is investigated also in Hiperlan standard. We describe in this section some of the mechanisms that have been proposed to introduce QoS based MAC layer.

5.1 Overview of QoS enhancement schemes proposed for IEEE 802.11

- **Basic mechanism: DCF**

DCF is the basic medium access mechanism of the 802.11 [21]. In this mode, if the medium is found idle for longer than a DIFS (Distributed Inter Frame Space) then the station can transmit a packet. Otherwise, a backoff process is started. More specifically, the station computes a random value called backoff time, in the range of 0 and CW (Contention Window) size. The backoff timer is periodically decremented by one for every time slot the medium remains idle after the channel has been detected idle for a period greater than DIFS. As soon as the backoff timer expires, the station can access the medium. If no acknowledgment is received, the station assumes that collision has occurred. The CW is doubled until a predefined maximum value (CW_{max}) is reached. Then, the station schedules a retransmission by re-entering the backoff process. Otherwise, after a successful frame transmission, the CW is reset to its initial value, CW_{min}. Note that all data packets use this mechanism without any differentiation.

To introduce priorities in IEEE 802.11 using DCF, three techniques have been proposed in [24]. Each scheme uses different parameters to provide service differentiation:

(a) Backoff increase function: Each priority level has a different backoff increment function. Experiments show that this scheme performs well with UDP but not with TCP because ACKs affect the differentiation mechanism.

(b) Different DIFS: This scheme ensures that no high priority station has queued frames when station of low priority starts transmission. The main issue of this scheme is that low priority traffic suffers as long as high priority frames are queued.

(c) Different Maximum Frame Length: This mechanism is used to increase both transmission reliability and differentiation, and works well for TCP and UDP flows. However, in a noisy environment, long packets are more likely to be corrupted than short ones, which decreases the efficiency of this scheme.

In [27], an algorithm is proposed to provide service differentiation using two parameters of IEEE 802.11, the backoff interval and the IFS (used before each data transmission). This scheme proposes four levels of priority which ensure that high priority classes have a short waiting time when accessing the medium. Indeed, when a collision occurs, high priority stations are more likely to access the medium than a low priority ones. However, when there is not any high priority stations that want to transmit packets, the low priority ones still use a long backoff time.

[18] proposes a Distributed Fair Scheduling (DFS) scheme, which utilizes the ideas of fair queuing in the wireless domain. A distributed algorithm for rate-based service differentiation is described. This mechanism solves the problem of throughput fairness between different flows of traffic. However, the paper does not present an analysis of the delay differentiation.

Based on DCF, a fully distributed service quality estimation, radio monitoring, and admission control approach are proposed in [19] to support service differentiation. A Virtual MAC (VMAC) algorithm monitors the radio channel and estimates locally achievable service levels. The VMAC estimates MAC level statistics related to service quality such as delay, jitter, packet collision, and packet loss. A Virtual Source (VS) algorithm utilizes the VMAC to estimate application-level service quality. The VS allows application parameters to be tuned in response to dynamic channel conditions based on virtual delay curves. To provide service differentiation, they also introduce backoff timer differentiation, such as $CW_{min}^{high-pri} < CW_{min}^{low-pri}$, $CW_{max}^{high-pri} < CW_{max}^{low-pri}$. Results show that when these distributed virtual algorithms are applied to the admission control of the radio channel, then, a globally stable state can be maintained without the need for complex centralized radio resource management.

A distributed solution for the support of real-time sources over IEEE 802.11, called *Blackburst*, is discussed in [28]. This scheme modifies the MAC to send short transmissions in order to gain priority for real-time service. It is shown that this approach is able to support bounded delays. The main drawback of this scheme is that it requires constant channel access intervals for high priority traffic, otherwise the performance degrades very much. Moreover, this scheme is optimized to meet the service requirements of isochronous traffic sources, which is a significant limitation for variable data rate applications.

Despite the obtained improvements, the mechanisms presented above do not provide bounded delays and efficient medium utilization at high load due to the high collision rate.

5.2 New features of IEEE 802.11e

The IEEE 802.11 working group is currently working on the support of QoS in a new standard, called IEEE 802.11e [23]. A new access method called Hybrid Coordination Function (HCF) is introduced.

It is a queue-based service differentiations. HCF describes some enhanced QoS-specific functions, called contention-based HCF channel access and polling-based HCF access channel. Enhanced DCF (EDCF) is the contention-based HCF channel access. The goal of this scheme is to enhance DCF access mechanism of IEEE 802.11 and to provide a distributed access approach that can support service differentiation. More specifically, the CW_{min} parameter is set differently for different priority classes, yielding high priority classes with small CW_{min}.

For further differentiation, 802.11e proposes the use of different IFS set according to traffic classes. Instead of DIFS, an Arbitration IFS (AIFS) is used. The AIFS for a given class should be a DIFS plus some (possibly zero) time slots. Classes with the smallest AIFS will have the highest priority as shown in Figure 2.

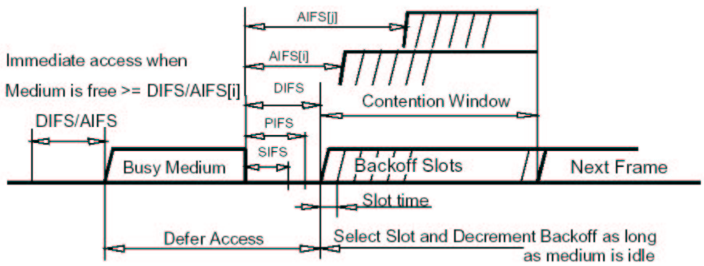

Figure 2: IFS relationships of IEEE 802.11e

Each Traffic Category (TC) within the station behaves like a virtual station: it contends independently for access to the medium. Figure 3 compares the 802.11e architecture that supports queue-based differentiation with the original one queue based DCF access mechanism.

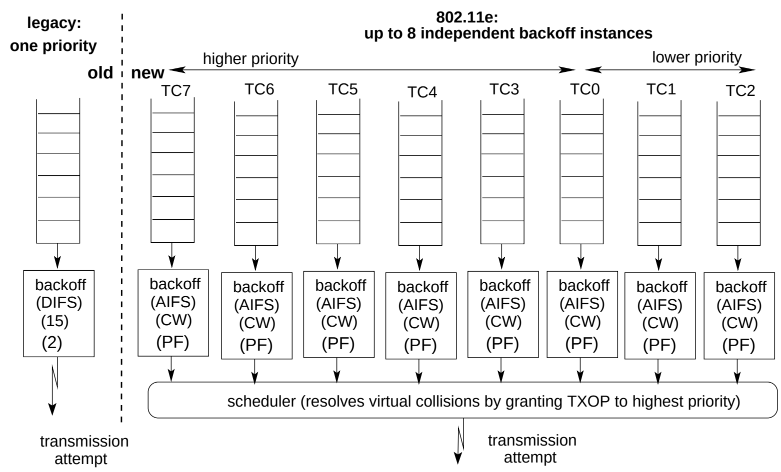

Figure 3: Queue-based EDCF vs. basic DCF

To decrease delay, jitter, and achieve higher medium utilization, packet bursting is proposed in this standard. So, once a station has gained access to the medium, it can be allowed to send more than one frame without contending for the medium again, as long as the total access time does not exceed a certain limit (TxOpLimit) and no collision occurs.

Per priority differentiation used by EDCF ensures better services to high priority class while offering a minimum service for low priority traffic. Although this mechanism improves the quality of service of real-time traffic, the performance obtained are not optimal since EDCF parameters cannot be adapted to the network conditions. In fact, since each TC is implemented as a virtual station, the collision rate increases very fast when the contentions to access the medium are high, which significantly affects the goodput, the latency and thus, decreases the performance of delay-bounded applications [25].

5.3 Adaptive EDCF (AEDCF)

When two or more TCP senders share the same receiver, they all receive TCP-ACKs with the same priority (limited to the same receiver priority). This tends to reduce the service differentiation. Furthermore, if the shared receiver is slow, the observed relative priority will also be reduced [26]. This motivates the use of queue-based differentiation where a shared node handles simultaneously several flows with different priorities. To improve the performance under different load rates and to increase the service differentiation in EDCF-based networks, a new scheme called Adaptive EDCF (AEDCF) has been proposed in [22]. This scheme extends the basic EDCF by making it more adaptive taking into account network conditions. Indeed, AEDCF uses a dynamic procedure to change the contention window value of each priority class differently. In fact, each class updates its contention window based on the estimated collision rate computed during a constant period. For further differentiation, each traffic category multiplies this collision rate by a priority factor [22].This mechanism offers to high priority traffic a higher probability to generate smaller CW value than low priority traffic and so they can access the medium first. Moreover, this scheme achieves a high medium utilization and it is much more efficient at high load. Furthermore, it improves total goodput, delay and delay-jitter.

5.4 Service differentiation in Hiperlan standard

HIPERLAN (HIgh PErformance Radio Local Area Network) is a Wireless LAN standard[29]. The HIPERLAN MAC protocol explicitly supports a quality of service (QoS) for packet delivery. There are two mechanisms that are provided to introduce service differentiation: the user priority of a packet (high or low) and the packet lifetime. Thus, the residual lifetime of a packet and its user priority are used to determine its Channel Access Mechanism (CAM) priority which can fall into one of five priority levels. Thus, the priority of each packet increases while its lifetime expires. Since multi-hop routing is supported within the standard the lifetime of a packet and the residual lifetime are transmitted along with the packet. Moreover, packets that cannot be delivered within the allocated lifetime are discarded.

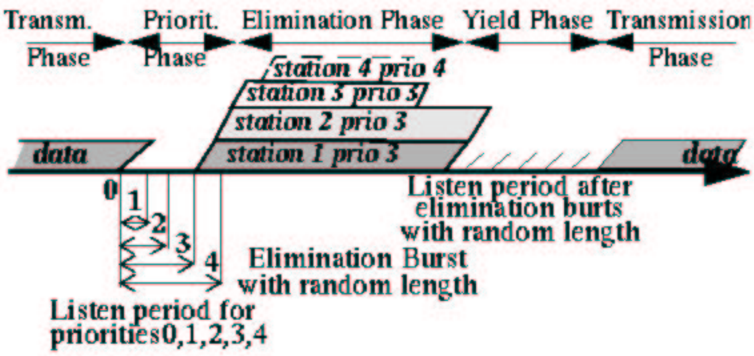

Figure 4: hiperlan medium access mechanisms

5.4.1 The Channel Access Mechanism

Each node that has packets to be transmitted sens the channel status. If the channel is sensed idle for at least 1700 bit-periods, then the node is allowed to immediately start transmission of the data frame. However, if the channel is considered not free, the access mechanism is split into three phases as shown in Figure 4: Priority resolution, Elimination and Yield phase. In the first phase, a station seeking access to the media, listens to the medium during the priority slots of the higher priorities that should have passed idle. If the channel was idle for p-1 priority slots only the stations with the same highest priority are admitted to the contention phase. Otherwise it stops contending and waits for the next channel access cycle.

In the second phase every remaining station transmits a burst with a random length. This burst length is calculated with a certain discrete probability distribution. Then, the station listens to the channel for an Elimination Survival Verification Period (ESVP). If the channel is sensed idle, the node is admitted to the yield phase. Else, it drops itself from contention and waits for the next channel access cycle. After this phase, at least one station survives.

In the The yield phase, the station has to listen the medium for a random length period. If it hears another station starting its transmission before its own yield phase is expired, it stands back from transmission. If not, it transmits immediately the data frame after the end of the yield period.

Although Hiperlan claims to support time bounded services it doesn't provide any services that guarantee QoS requirements. In fact, it does not support the allocation of a fixed portion of bandwidth nor any other QoS parameters. Thus, Hiperlan is still just a best effort network, not suitable to extend QoS-guaranteeing networks. Furthermore, the Hiperlan performance evaluation that has been presented in [30], shows that the performance of this standard seems similar to the IEEE 802.11 one. But, the IEEE 802.11 networks are simple setup and reasonable performance in typical network conditions.

6 Conclusion

Meeting QoS guarantees in mobile network systems is fundamentally an end-to-end issue, that is, from application to application. Obviously, QoS routing and QoS medium access mechanisms must cooperate together in order to achieve good performance by meeting between application requirements and available network resources. In fact, the routing protocols that support QoS must be adaptive to cope with the time-varying topology and time-varying network resources. For instance, it is possible that a route that was earlier found to meet certain QoS requirements no longer does so due to the dynamic nature of the topology. In such a case, it is important that the network intelligently adapts the session to its new and changed conditions. Moreover, QoS signaling will work better if it coordinates with QoS routing. Indeed, without QoS routing, signaling can work but the selected path may not have enough resources.

On the other hand, it is much more difficult to control the wireless radio channel than the fixed wireline networks. This becomes even more challenging during resource contention and under heavy usage. Thus, the QoS MAC protocol is an essential component in QoS support in MANETs. All upper-layer QoS components (QoS application, QoS routing and QoS signaling) are dependent on the QoS MAC protocol. If we don't consider a specific QoS MAC layer protocol that cooperates with the up- layer protocols, QoS provisioned by these later will be much disturb. Other QoS components in MANETs, such as scheduling and admission control, can be borrowed from other network architectures without or with few modifications.

The works that have been proposed in the literature to introduce QoS in MANET, seldom consider the cooperation between different layers to provide good performance for the applications. However, it is very important to consider a cross-layer QoS model for MANET. This QoS model should introduce different QoS mechanisms based on application, network, and MAC layer, in order to provide interaction between their different functionalities. The goal is to enable applications to select the routes that satisfy, as soon as possible, their requirements.

References

[1] D.D. Perkins and herman D. Hughes,"A survey on quality-of-service support for mobile ad-hoc", *Wireless Communications and Mobile Computing*, 2002, pp. 2:503-513.

[2] N. Nikaein, H. Labiod, and C. Bonnet, "DDR- Distributed Dynamic Routing algorithm for Mobile Ad Hoc Networks", *MOBI-HOC2002*, Lausanne, June 2002.

[3] X. Xiao and L. M. Ni, "Internet QoS: A big picture", *IEEE Network*, April 1999, pp. 8.18.

[4] R. Braden, L. Zhang, S. Berson, S. Herzog, and S. Jamin, "Resource reservation protocol (RSVP) . version 1 functional specification", *IETF RFC 2205*, Sept. 1997.

[5] H. Xiao, W. K. G. Seahand A. Lo, and K. C. Chua, "A flexible quality of service model for mobile ad-hoc networks", *IEEE VTC2000-spring*, Japon/Tokyo, 2000.

[6] G.Ahn, A.T.Campbell, A.Veres and L.Sun, "SWAN: Service Differentiation in Stateless Wireless Ad Hoc Networks ", *Proc. IEEE INFOCOM*, *2002*.

[7] N. Nikaein, C. Bonnet, Y. Moret, and I.A. Rai, "2lqos- twolayered quality of service model for reactive routing protocols for mobile ad hoc networks", *in Proceeding of SCI - 6th World MSCI*, Orlando/ Florida , 2002.

[8] K. Wu and J. Harms, "QoS support in mobile ad hoc networks", *Crossing Boundaries - an interdisciplinary journal VOL 1*, No 1 - Fall 2001.

[9] *Elizabeth M. Royer,* "A Review of Current Routing Protocols for Ad Hoc Mobile Wireless Networks", *IEEE Personal Communications*, April 1999.

[10] M. Mirhakkak, N. Schult, and D. Thomson, "Dynamic quality-of-service for ad hoc networks", *The MITRE corporation 2000.*

[11] N. Nikaein, C. Bonnet and N. Nikaein, "HARP - Hybrid Ad Hoc Routing Protocol", *IST: International Symposium on Telecommunications*, Iran/Tehran, 2001.

[12] Zeinalipour-Yazti Demertrios, "A Glance at Quality of Services in Mobile Ad-Hoc Networks", *Research Report for CS260 - Seminar in Mobile Ad Hoc Networks*, California, Nov. 2001.

[13] P. Sinha, R. Sivakumar, and V. Bharghavan, "CEDAR: A Core-Extraction Distributed ad hoc Routing Algorithm", *In Proc. IEEE INFOCOM*, March. 1999.

[14] G-S. Ahn, A.T. Campbell, S-B. Lee, and X. Zhang, "INSIGNIA Internet Draft", *draft-ietf-manet-insignia-01.txt*, October 1999.

[15] S-B. Lee, G-S. Ahn, and A.T Campbell, "Improving UDP and TCP Performance in Mobile Ad Hoc Networks with INSIGNIA", *IEEE Communications Magazine*, June 2001.

[16] P. Paul and S.V.Raghavan, "Survey of QoS Routing", *ICCC 2002*, Mumbai, Aug 2002.

[17] R. Braden, D. Clark, and S. Shenker, "Integrated Services in the Internet Architecture: an overview", *IETF RFC1633*, June 1994.

[18] N. H. Vaiday, P. Bahl, and S. Gupa, "Distributed Fair Scheduling in a Wireless LAN", *In the* 6^{th} *Annual International Conference on Mobile Computing and Networking*, Boston, Aug 2000.

[19] A. Veres, A. T. Campbell, M. Barry, and L-H. Sun, "Supporting service differentiation in wireless packet networks using distributed control", *IEEE JSAC*, Special Issue on Mobility and Resource Manegement in Next-Generation Wireless Systems, October 2001, Vol. 19, No. 10, pp. 2094-2104.

[20] K. Nichols, V. Jacobson, and L. Zhang, "A two-bit Differentiated Services Architecture for the internet", *IETF RFC2638*, August 1998.

[21] IEEE WG, Part 11, "Wireless LAN Medium Access Control (MAC) and Physical Layer(PHY) Specifications", *IEEE 802.11 Standard*, 1999.

[22] L. Romdhani, Q. Ni, and T. Turletti, "AEDCF: Enhanced Service Differentiation for IEEE 802.11 Wireless Ad-hoc Networks", *In Proc, of IEEE WCNC 2003*, march 2003.

[23] IEEE WG, " Wireless LAN Medium Access Control (MAC) and Physical Layer(PHY) Specifications:MAC *Enhancements for Quality of Service (QoS)", 802.11e Draft 3.1,* May, 2002.

[24] I. Aad and C. Castelluccia, "Differentiation mechanisms for IEEE 802.11", *Infocom*, April 2001.

[25] A. Lindgren, A. Almquist, and O. Schelen, "Evaluation of Quality of Service Schemes for IEEE 802.11", *In Proc. of IEEE LCN 2001*, November 2001.

[26] I. Aad and C. Castelluccia, "Remarks on per-flow differentiation in IEEE 802.11", *European Wireless 2002*, February 2002.

[27] J. Deng and R. S. Chang, "A priority Scheme for IEEE 802.11 DCF Access Method", *IEICE Transactions in Communications*, vol. 82-B, no. 1, January 1999.

[28] J. L. Sobrinho and A. S. Krishnakumar, "Real-time traffic over the IEEE 802.11 medium access control layer", *Bell Labs Technical Journal*, Autumn 1996, pages 172-187.

[29] ETSI. Radio Equipment and Systems (RES), "HIgh PErformance Radio Local Area Network (HIPERLAN), Type 1 Functional Specification", *Draft prETS 300 652*, France, Aug 1996.

[30] J. Weinmiller, M. Schlager, A. Festag, and A. Wolisz, "Performance Study of Access Control in Wireless LANs - IEEE 802.11 DFWMAC and ETSI RES 10 HIPERLAN", *ACM/Baltzer Mobile Networks and Applications*, 1997, vol. 2, no. 1, pp. 55-67.