Review

QoS in IEEE 802.11-based wireless networks: A contemporary review

Aqsa Malik [a], Junaid Qadir [b,*], Basharat Ahmad [a], Kok-Lim Alvin Yau [c], Ubaid Ullah [a]

[a] School of Electrical Engineering and Computer Science (SEECS), National University of Sciences and Technology (NUST), Pakistan
[b] Electrical Engineering Department, SEECS, NUST, Pakistan
[c] Faculty of Science and Technology, Sunway University, Malaysia

ABSTRACT

Apart from mobile cellular networks, IEEE 802.11-based wireless local area networks (WLANs) represent the most widely deployed wireless networking technology. With the migration of critical applications onto data networks, and the emergence of multimedia applications such as digital audio/video and multimedia games, the success of IEEE 802.11 depends critically on its ability to provide Quality of Service (QoS). A lot of research has focused on equipping IEEE 802.11 WLANs with features to support QoS. In this survey, we provide an overview of these techniques. We discuss the QoS features incorporated by the IEEE 802.11 standard at both physical (PHY) and Media Access Control (MAC) layers, as well as other higher-layer proposals. We also focus on how the new architectural developments of software-defined networking (SDN) and cloud networking can be used to facilitate QoS provisioning in IEEE 802.11-based networks. We conclude this paper by identifying some open research issues for future consideration.

© 2015 Elsevier Ltd. All rights reserved.

Contents

* Corresponding author.

1. Introduction

The IEEE 802.11 wireless local area networking (WLAN) standard defines one of the most widely deployed wireless technologies in the world. The popularity of wireless networking is driven by the ubiquity of portable mobile hand-held devices, and the convenience of untethered communications. With the increasing deployment of multimedia content on the Internet—such as digital video, Voice over IP (VoIP), videoconferencing, and multi-player networked games—along with the deployment of time-sensitive critical applications, there is a strong motivation to develop QoS features to meet the more stringent performance requirements (Baghaei and Hunt, 2004).

While the Internet and data networking models of the IEEE 802.11 WLAN technology, which are based on the *datagram delivery* model of IP, provide simple, adaptive and fault resilient network, they are ill-suited to QoS provisioning. The underlying datagram model of IP is a *best-effort* service—i.e., while the network tries to deliver packet to the destination correctly without any packet losses, it makes no guarantees. Multimedia applications, in particular, need stronger guarantees about the minimum throughput and maximum latency to work satisfactorily. An expensive solution for ensuring QoS is to *overprovision*. Most of the Internet QoS effort has focused on how to get a network with less capacity meet application requirements at a lower cost. In IEEE 802.11 WLANs, the error and interference prone nature of wireless medium—due to fading and multipath effects (Gast, 2005)—makes QoS

provisioning even more challenging. The combination of best-effort routing, datagram routing, and an unreliable wireless medium, makes the task of QoS provisioning in IEEE 802.11 WLANs very challenging.

In this survey, we provide a focused overview of work done to ensure QoS in the IEEE 802.11 standard. We have the following three goals: (i) to provide a self-contained introduction to the QoS features embedded in the IEEE 802.11 standard; (ii) to provide a layer-wise description and survey of techniques adopted for ensuring QoS in the IEEE 802.11 networks; and (iii) to survey the applications of new networking architectures—such as software defined networks (SDN) and cloud computing—for QoS provisioning in the IEEE 802.11-based WLANs.

Contributions of this paper: A lot of research has been conducted on the topic of QoS (Wang, 2001), including numerous surveys that have focused on the QoS problems for specific wireless networks such as Wireless Sensor Networks (WSNs) (Chen and Varshney, 2004), wireless mesh networks (WMNs) (Mogre et al., 2007), and IEEE 802.11-based WLANs (Zhu et al., 2004; Lindgren et al., 2001; Ni et al., 2004; Aboul-Magd, 2009). Our work is different from the previous work in that we provide an updated account of QoS literature in IEEE 802.11-based wireless networks including a discussion of recent architectural developments, such as cloud computing and SDN, that facilitate finer network management control, so we have reinvigorated the interest of the research community in providing high QoS in IEEE 802.11-based wireless

networks. In addition to highlighting the QoS features incorporated into the IEEE 802.11 networking standard, we also highlight different QoS approaches pertaining to different layers of the TCP/IP model.

Organization of this paper: This survey is organized in the following way. In Section 2, we provide a broad-based introduction on the general area of Internet QoS. In Section 3, we present the QoS features which are intrinsic to IEEE 802.11[1] for the physical (PHY) and the medium access control (MAC) layers. We provide an overview of the IEEE 802.11 a/b/g and higher-throughput IEEE 802.11 standards (802.11 n/ac/ad) in Sections 3.1 and 3.2. We follow this by a discussion on MAC layer QoS features proposed in IEEE 802.11 in Section 3.3. Apart from the QoS features that are part of the IEEE 802.11 standard, various work has focused on QoS improvement including work at the network layer (discussed in Section 4), the transport layer (discussed in Section 5), the application layer (discussed in Section 6) as well as cross-layer work (discussed in Section 7). The promise of recent architectural developments, such as cloud computing and SDN, in enabling QoS, along with a survey of proposed work, is provided in Section 8. Thereafter, we discuss some open research issues in Section 9. Finally, we provide concluding remarks in Section 10.

To facilitate the reader, acronyms used in this paper are collected in Table 1 as a convenient reference.

2. Internet QoS—a broad introduction

There has been a lot of work on Internet QoS, the bulk of which has focused on wired networks (Wang, 2001; Jha and Hassan, 2002). While many of the ideas developed for Internet QoS are also applicable more broadly to wireless QoS, wireless networks do provide some unique challenges motivating the development of new methods (Setton et al., 2005). In this section, we provide a broad overview of the abundant literature on Internet QoS.

The original applications of the Internet—such as file transfer and email—are *elastic* applications which are not bound by stringent performance requirements, and therefore match well with the Internet's datagram delivery model. The modern Internet world, which is full of multimedia applications, requires QoS guarantees that users have come to expect from the telecommunications networking world. To support multimedia and other interactive/high performance applications, there is a need to support QoS features through *QoS provisioning* that provides resource assurance along with service differentiation. Various techniques have been developed to facilitate QoS provisioning, including (i) congestion control, (ii) admission control, and (iii) traffic shaping and engineering.

In the remainder of this section, we study the problems of resource allocation and service differentiation, and will introduce the techniques of admission control, congestion control, scheduling, as well as traffic shaping and engineering that can be used to facilitate QoS provisioning.

2.1. Resource allocation

Fundamentally, many QoS issues stem from the problem of *resource allocation*. A computer network is composed of various resources—such as links of varying bandwidths, routers with varying buffer sizes—that are shared by the different network applications and users. Packet delays and losses occur if the network resources cannot meet all the traffic demands. A network that supports QoS must actively manage resource allocation to satisfy various users' and applications' demands. Without appropriate resource allocation, network performance and service quality deteriorate rapidly under heavy load due to dropped packets and congestion. There are two main architectural approaches to resource allocation in the Internet: Integrated Services (IntServ) and Differentiated Services (DiffServ). Apart from IntServ and DiffServ, other QoS frameworks have also been proposed (Aurrecoechea et al., 1998). We, however, focus only on the more important QoS framework proposals, namely IntServ and DiffServ.

2.1.1. IntServ

IntServ performs *per-flow resource reservation* for service differentiation. IntServ provides services on a per-flow basis where a flow is a packet stream with common source address, destination address and port number. In IntServ, a packet scheduler is used to enforce resource allocation to individual flows while supporting prioritization. The IntServ scheduler can be used to provide delay bounds. The delay bounds can be deterministic or statistical—for deterministic bounds, isolation or dedication of resources is required, while statistical bounds can be provided when statistical multiplexing (Mahadevan and Sivalingam, 1999) is used. There are two key IntServ abstractions, namely Reserved Resources and Standard Resources. In the Reserved Resource abstraction, the router must know the amount of resources currently reserved for on-going sessions. The Standard Resource abstraction includes the capacities of the links and the router buffers, respectively. An example is Call Setup in which buffers are kept at the routers. These buffers ensure a specific amount of bandwidth is allocated to the flows at each router (Zhang et al., 1993; Clark et al., 1992; Shenker et al., 1997; Braden et al., 1994).

2.1.2. DiffServ

DiffServ, on the other hand, performs *per-class resource reservation* for service differentiation, and makes use of prioritization, multiple forwarding classes, and edge policing to categorize traffic into different classes; and the traffic is treated according to its respective classes. The edge routers are responsible for the complex operations in the network; while the core routers perform simple and easy computations. The packet-handling rule in DiffServ is termed as Per-Hop Behavior (PHB). In other words, each network device along a path behaves in a certain way in which a specific group of packets have the same priority value. The PHB rule decides whether a packet needs to be forwarded or dropped depending on the QoS-based precedence value of the packet. However, the framework is very complex and cannot be applied to heterogeneous networks (Blake et al., 1998; Jacobson et al., 1999; Nichols and Jacobson, 1999). DiffServ has been used for implementing QoS in various IEEE 802.11-based wireless networks such as Chaouchi and Munaretto (2004) and García-Macías et al. (2003).

2.2. Service differentiation

Service differentiation is used to support multiple services with diverse requirements—such as interactive delay-sensitive services along with elastic delay-tolerant file transfer services (Christin and Liebeherr, 2003). The overprovisioning of network resources is not always possible in radio networks, thus making service differentiation an integral component of most QoS-based solutions. In service differentiation, several parameters (e.g., packet deadline) can be modified to define how a flow should access the wireless medium (Chen and Varshney, 2004). A variety of services can be provided by the use of simple network parameters deployed in network nodes, and these services can be classified according to a

---

[1] We note here that the IEEE 802.11 standard directly addresses the PHY and MAC layers only.

Table 1
Acronyms used in this paper.

Acronym	Expanded form
AMC	Adaptive Modulating Scheme
APSD	Automatic Power Save Delivery
ARQ	Automatic Repeat reQuest
ATM	Asynchronous Transfer Mode
BER	Bit Error Rate
BPSK	Binary Phase Shift Keying
CAC	Call Admission Control
CAP	Control Access Period
CbWN	Cloud-based Wireless Network
CW	Contention Window
DCF	Distributed Coordination Function
DFS	Distributed Fair Scheduling
DIFS	DCF Interframe Space
DiffServ	Differentiated Services
DSSS	Direct-Sequence Spread Spectrum
EDCA	Enhanced Distributed Channel Access
EDCF	Enhanced DCF
EDD	Earliest Due Date
FEC	Forward Error Correction
FHSS	Frequency Hopping Spread Spectrum
FCFS	First-Come First-Served
FIFO	First In First Out
HCF	Hybrid Coordination Function
HCCA	HCF Controlled Channel Access
IntServ	Integrated Services
LTE	Long-Term Evolution
MAC	Media Access Control
MDP	Markov Decision Process
MPDU	MAC Protocol Data Unit
MPLS	Multiprotocol Label Switching
MIMO	Multiple Input Multiple Output
NUC	Network Utilization Characteristic
OMAR	Opportunistic Medium Access and Adaptive Rates
OSAR	Opportunistic Scheduling and Auto Rate
PCF	Point Coordination Function
PIFS	PCF Interframe Spacing
PHB	Per-Hop Behavior
PSTN	Public Switched Telephone Network
QAM	Quadrature Amplitude Modulation
QoE	Quality of Experience
QoS	Quality of Service
RTS	Request To Send
SDN	Software Defined Networking
SIFS	Short Interface Spacing
SISO	Single Input Single Output
STA	(Wireless) Station
SWN	Software Defined Wireless Network
TDM	Time Division Multiplexing
TXOP	Transmission Opportunity
VoIP	Voice over IP
VM	Virtual Memory
WFQ	Weighted Fair Queuing
WLAN	Wireless Local Area Networking
WRR	Weighted Round Robin
WSN	Wireless Sensor Network

Fig. 1. Taxonomy of QoS metrics (adapted from Hanzo and Tafazolli (2007) and Chalmers and Sloman (1999)).

audio, require less bandwidth, while video and file transfers require high bandwidth. Similarly, the delay requirements also differ with the type of application. Some applications, such as email, are not delay-sensitive. However, interactive applications, such as web browsing, videoconferencing and live streaming, have more stringent delay requirements. The variation in the packet arrival time within a stream of packets is called jitter. Email, file sharing and remote login are not affected by jitters in the network traffic; while real-time applications, audio and video do. The loss of packets in the case of audio and video is not of significance. The other applications, as discussed previously, cannot tolerate an increased packet loss rate in transmissions (Tanenbaum, 2003). In this work, we will focus mostly on the QoS parameters of delay, throughput, and guaranteed bandwidth.

2.3. Admission control

One way of supporting QoS is through *admission control*—in which new sessions are allowed onto the network only if sufficient resources are available to provide service to the new and existing sessions (Knightly and Shroff, 1999). The interest in the field of admission control, has been driven by the idea that regulation of incoming traffic flows prevents network congestion, and helps in ensuring QoS. Call Admission Control (CAC) is a traffic management system employed in Public Switched Telephone Networks (PSTNs) (Perros and Elsayed, 1996). The CAC scheme is easy to implement because of the homogeneous environment. However, the present networking environment of homogeneous network is not preserved. Hence, the admission control function is more challenging in heterogeneous networks (e.g., joint WLANs and IP networks) (Gao et al., 2005). A new flow request is admitted only if the ongoing flows are not negatively affected. The common parameters used for admission control are peak bandwidth requirement and the average rate (Aboul-Magd, 2009). Therefore, the task of admission control is to maximize resource utilization in the network, and to control the amount of traffic to achieve the predefined performance objectives of the current flows.

Hou et al. (2009) have presented a formal theory of QoS provisioning in unreliable wireless networks, such as the IEEE 802.11-based wireless networks, which subsumes a framework for jointly addressing three important QoS criteria, namely delay, delivery ratio, and channel reliability. They also propose algorithms and policies for admission control and scheduling that can

large number of characteristics (Aad and Castelluccia, 2001). The QoS of the system is enhanced by differentiating the priority of each host and offering them different levels of QoS parameters.

Service requirements are often application-specific. For example, certain applications are delay-sensitive (e.g., voice conferencing which is sensitive to round-trip delay), while others are concerned more with average transmission rate (e.g., bulk file transfer). Service requirements are often expressed using metrics (i) bandwidth, (ii) delay, (iii) jitter, and (iv) loss rate. A more comprehensive, but still non-exhaustive, listing of QoS metrics is displayed in Fig. 1. To accommodate the impact of these metrics, the network must support multiple QoS strategies to support different applications (Ksentini et al., 2004; Jha and Hassan, 2002).

The bandwidth requirements of different applications are different. Some applications, such as email, remote login and
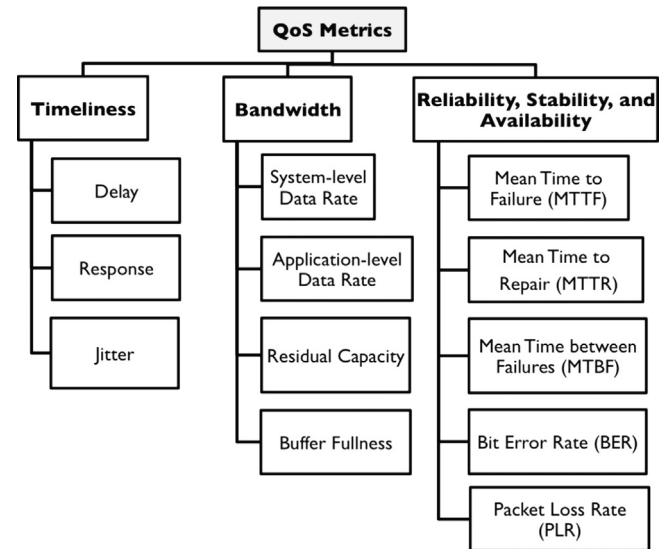
be implemented in IEEE 802.11-based networks. The authors analytically develop necessary and sufficient conditions to satisfy these three criteria. More details of admission control techniques in the context of IEEE 802.11 standard are described later in Section 4.

2.4. Congestion control

Congestion control in the modern Internet is typically performed by using the TCP protocol (Vicisano et al., 1998). Congestion in a network may occur if the number of packets sent to the network is greater than the number of packets a network can handle. Congestion control refers to the techniques to control the congestion level and keep the load below the capacity. In the QoS-Integrated Services, the congestion control mechanism should be different for different kinds of sources: e.g., file transfer/email is different from real-time voice/video applications (Xiao and Ni, 1999). The QoS enabled routers provide services to certain flows based on their requirements. Congestion control helps to provide priority differentiation of flows by servicing queues in different manners (e.g., the order in which the flows are serviced).

2.5. Scheduling

Scheduling is the key to share network resources fairly among users in a network, and it provides service guarantees to time-critical applications. The scheduler first decides the order of requests to be served, and then it manages the queues of these awaiting requests. The scheduling scheme is important for the networks because there are two types of applications. One is insensitive to the performance that users receive from the network, and the other has a strict bound on the performance. The scheduling can provide different services to the flows using parameters such as different *bandwidths*—by serving only a single flow at a particular interval; different *mean delays*—according to the level of priority defined for the flow; and different *loss rates*—by assigning more or fewer buffers to the flows (Keshav, 1997). The scheduling mechanism adopted in the IEEE 802.11 standard is explained in detail later in Section 3.3.

2.6. Traffic shaping and engineering

Traffic in data networks is bursty in nature. Traffic shaping is a technique for handling the bursty nature of the traffic entering a network through controlling and allocating appropriate levels of network bandwidth (Tanenbaum, 2003). The goal is to regulate average traffic rate and reduce congestion. The traffic shaping is performed at the boundary nodes. These nodes have classifiers that mark the flows according to their service requirements. The mechanisms of traffic management can be classified in a number of ways (Wang, 2001). One possible criterion is time scale (Aboul-Magd, 2009). In order to achieve QoS guarantees, decisions on buffering and forwarding must be performed quickly. Traffic engineering is the process that maximizes network utilization through careful distribution of network resources (Wang, 2001). Most of the Internet backbones currently rely on label switching by adopting 'multi protocol label switching' (MPLS) technology. The purpose of label switching is to enhance the scope of traffic engineering, QoS provisioning and overlay networks (Peterson and Davie, 2007). The traffic shaping mechanism for the IEEE 802.11 standard is defined in Section 3.3.

3. QoS support in IEEE 802.11

Standards in the IEEE 802 project target the PHY layer and the MAC layer. While IEEE 802.3 defines the PHY and MAC layers for wired LANs; the prominent IEEE 802.11 standard, which is the focus. The first IEEE 802.11 specification was published in 1997, and it has undergone numerous subsequent amendments. The IEEE 802.11 working group has various task groups focusing on a myriad of niche concerns with an elaborated description of the IEEE 802.11 universe provided in (Hiertz et al., 2010). The focus of the various task groups can be seen in Zhu et al. (2004) and Ni et al. (2004).

Providing QoS services in applications using traditional IEEE 802.11 standards is difficult since they provide no explicit mechanisms for service differentiation. Various parameters extracted from the general traffic layout are used to ensure acceptable QoS in these networks, such as *goodput*—which is the measure of packet arrival rate during a fixed period of time; *load level*—which indicates the usage of a medium on per time basis; and *available bandwidth*—which measures the rate at which new flows can send traffic without affecting the existing flows in the network (Dujovne et al., 2010; Ng et al., 2005). The legacy standards only have the basic Distributed Coordination Function (DCF) and the optional Point Coordination Function (PCF) enhancements, such as collision avoidance and a First In First Out (FIFO) scheduler (Ni, 2005). Even with the contention-free PCF, the QoS problem could not be solved. Consequently, the services provided to the users do not have optimal performances for various applications including audio and video applications during heavy network loads (Wu et al., 2001). In Section 3.4, we present the IEEE 802.11e standard, which provides MAC layer enhancements for QoS, that incorporates traffic priority and queueing to enable service differentiation among the flows (Mangold et al., 2002).

3.1. QoS support in IEEE 802.11 a/b/g

The IEEE 802.11 has evolved in different eras to satisfy differing requirements of applications (LaMaire et al., 1996; Raniwala and Chiueh, 2005). The IEEE 802.11a standard is part of the original IEEE 802.11 standard, and it operates in the 5 GHz range with a data rate of 54 Mbps while supporting the Frequency Hopping Spread Spectrum (FHSS) and direct sequence spread spectrum (DSSS). Unfortunately, due to the use of high frequency spectrum, the technique of DSSS faced the problems of short transmission range and interference (Zhou et al., 2006). To address this problem, IEEE 802.11b, which is also based on DSSS technology, operates in the 2.4 GHz spectrum with a data rate of 11 Mbps. The 802.11b standard is not backward compatible with the IEEE 802.11a standard. The IEEE 802.11g standard—operating at 2.4 GHz with a data rate upto 54 Mbps—is introduced for backward compatibility with the IEEE 802.11a standard.

QoS limitations of IEEE 802.11a/b/g: In the DCF-based schemes, the access to the medium is given on first come first served (FCFS) basis. This creates some fairness problems, which can potentially result in flows being deprived of their fair bandwidth share. There is no proper mechanism to distinguish between the flows on priority basis in the PCF-based environments. The legacy standards of IEEE 802.11 a/b/g have no standard mechanisms to ensure QoS (Mohapatra et al., 2003; Guo and Chaskar, 2002). Because these standards do not incorporate admission control, performance degradation occurs during heavy traffic load.

3.2. QoS support in high throughput IEEE 802.11 (802.11 n/ac/ad)

To support the need of high throughput wireless networking, various IEEE 802.11 standards have been proposed in recent time

such as the 802.11n, 802.11ac, and 802.11ad standards. The IEEE 802.11n standard is based on the Multiple Input Multiple Output (MIMO) technology, and it offers a high data rate of upto 600 Mbps. The IEEE 802.11ac standard aims to support an even higher data rate of 1 Gbps, while IEEE 802.11ad standard aims at achieving a rate upto 7 Gbps exploiting the wideband channels available in the 60 GHz band. These new standards incorporate scheduling mechanisms, Call Admission Control algorithms, and PHY and MAC layer enhancements for supporting multimedia applications with QoS. The interested reader is referred to a comprehensive study of QoS support in very high throughput IEEE 802.11 architectures (IEEE 802.11 n/ac/ad) presented in (Charfi et al., 2013).

3.3. MAC layer QoS features for IEEE 802.11

In general, the major techniques used for ensuring QoS at the MAC layer include admission control and scheduling. In the IEEE 802.11 standard, the MAC layer provides the functionality of addressing, framing, reliability check, and access coordination to the wireless medium (Li et al., 2007). The MAC layer with QoS enhancements aims to provide the network with a much reduced overhead, segregating frames on the priority basis, and keeping the collisions to the least possible level. This section describes the techniques implemented in the legacy IEEE 802.11 standard. The rest of this subsection presents a description of the QoS-focused IEEE 802.11e standard.

The wireless systems can be configured in two different modes in the IEEE 802.11 architecture: (i) the ad hoc mode, and (ii) the infrastructure mode. The infrastructure mode has multiple stations that can communicate with each other through an access point (AP), where the APs have connectivity with a wired network at the backend. While in the ad hoc mode, the stations can communicate directly with each other without any intervening access points or a backend wired network. The IEEE 802.11 has two medium access coordination functions, namely the DCF and the PCF.

Various kinds of QoS enhancement techniques have been proposed for the IEEE 802.11 standard (Zhao et al., 2002; Li and Prabhakaran, 2005; Achary et al., 2012), and they are explained in the rest of this subsection.

3.3.1. Priority queueing

This method is used to provide priority queues at the MAC layer where data packets are segregated on the basis of priorities. Whenever a particular station has access to the channel, it transmits the one which has the highest priority among the queued packets. All the stations must contend with each other for access to the medium.

Table 2
Priority levels corresponding to various applications types for supporting priority queueing in IEEE 802.11.

Priority	802.1 D User priority	802.11e access category (AC)	Description
Lowest	1	AC_BK	Background traffic
	2	AC_BK	Background traffic
	0	AC_BE	Best effort
	3	AC_BE	Best effort
	4	AC_VI	Video
	5	AC_VI	Video
	6	AC_VO	Voice
Highest	7	AC_VO	Voice, network management

Priority queueing is done in a way that there are eight different levels of priority, and therefore eight different queues must be maintained. Table 2 shows the classification of these priorities queues. The highest level or the seventh level has the highest priority and it is assigned to the most critical applications. The next two levels, i.e. levels 5 and 6, correspond to delay-sensitive video and audio applications. Levels 4 and below are used for regular data traffic, as well as streaming video. Level 0 is left for the traffic that can tolerate all the deficiencies of the best-effort service (Sundareswaran et al., 2007).

3.3.2. Differentiated Services

The QoS enhancements can also be classified in the terms of the DCF-based or the PCF-based enhancements. Figure 2 provides a taxonomy of DCF- and PCF-based enhancements (Ni and Turletti, 2004) for both priority queueing and Differentiated Services.

We initially discuss four main techniques for deploying *Differentiated Services using DCF*:

- *Distributed Fair Scheduling*: For good performance of a system, it is not a fair practice to restrict the services of low-priority traffic and to provide better services to high priority traffic. One way is to assign more bandwidth to the high priority traffic in comparison to the low priority traffic. Distributed Fair Scheduling (DFS) is a technique used in this respect. In this technique, each flow is assigned some weight depending on its priority and the bandwidth it gets is then proportional to this weight. This is a centralized technique in the sense that it uses a central AP, which has the information regarding all the traffic flows from different stations, and can therefore assign different weights to each of these flows. This technique thus differentiates among all the traffic flows going through the AP (Lindgren et al., 2003).
  The DFS scheme uses the backoff mechanism of IEEE 802.11 to decide the transmission order of each station. When the transmission starts, each station chooses a random backoff time. This backoff interval is a function of packet length and the priority of the flow. The stations with low priority flows have longer backoff intervals than the stations with high priority flows. Using packet size in the backoff calculation ensures fairness among the stations, resulting in smaller packets being sent more often. In the case of a station experiencing a collision, the new backoff interval is generated using the same algorithm.
- *Varying DIFS*: Another solution is to vary the distributed inter-frame spacing (DIFS) duration for differentiation among flows (Aad and Castelluccia, 2001). For example, we know that the ACK packet in the IEEE 802.11 standard gets higher priority
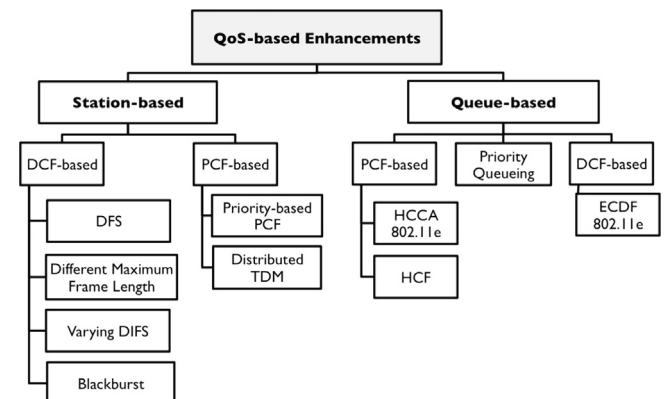

Fig. 2. MAC layer QoS enhancement schemes for IEEE 802.11-based wireless networks (described in Section 3.3).

than RTS packets, due to the fact that ACK packet waits short inter-frame spacing (SIFS) amount of time, while RTS packet waits DIFS amount of time, which is much longer. The same idea can be taken to the data frames; in which each flow's priority is set with a different DIFS duration. To avoid collisions, a backoff time is maintained similarly in these packets as well. Such technique is much beneficial in real-time applications, where delays have a greater significance compared to packet loss (Drabu, 1999).

- *Differentiated Maximum Frame Length*: In this approach, service differentiation is achieved in a way that different stations can transmit frames with different maximum frame sizes. The stations with high priority flows can transmit a larger frame than the one with the lower priority flows. To ensure this, there are two mechanisms: either the packets that exceed the maximum frame size are discarded or an upper bound on the size of packets is maintained in each station (Aad and Castelluccia, 2000). In some cases, when the packet size is greater than the maximum limit, the packets are fragmented. These fragments are sent without any RTS in between, waiting just for the reception of corresponding ACKs. These mechanisms provide us with the same data rates as those without fragmentation (Drabu, 1999).

- *Blackburst*: The blackburst scheme imposes certain constraints on high priority flows rather than the low priority flows which has been considered until now (Sharma et al., 2013). In this technique, every station gets access to the medium for a fixed interval of time (Wang et al., 2011). Once the station gets access to the medium, it jams the medium for a certain duration. Consider a station that has higher priority than others, and it has data packets to transmit, so it senses the channel. Once it detects the channel that has been idle for PIFS amount of time, it has the potential to transmit its frames. Hence, after waiting for a PIFS amount of time, it enters a blackburst contention period. A jamming signal, which is called blackburst, is then sent by this station to jam the channel. The length of this blackburst signal is proportional to the amount of time a particular station must wait before getting access to the medium. After the station has transmitted its blackburst signal, it again listens to check if any other stations are also sending a blackburst signal. The length of this blackburst signal is

compared to check whether it is longer or shorter than its own. Subsequently, the station with the longest blackburst shows that it has been waiting for a longer amount of time to access the channel, hence it is the next station to access the channel. This technique is similar to how TDM shares the same medium among the different flows, and it is used in real-time traffic and synchronization (Ni et al., 2004).

We next discuss two techniques for offering *Differentiated Services using PCF*:

- *Distributed TDM*: This mechanism uses a polling method as in the regular PCF mechanism, but time slots are also defined as in the TDM approach, and each of these time slots is assigned to a specific station. Once these time slots are assigned, each station knows when to transmit, and thus transmission of packets can be done with a very little involvement of the AP (Drabu, 1999).

- *Hybrid Coordination Function*: Hybrid Coordination Function (HCF) is a new coordination function proposed in IEEE 802.11e to enhance both DCF and PCF. HCF uses two methods: the first method is contention-based and it is known as Enhanced Distributed Channel Access (EDCA), and the second method is contention-free and it is known as HCF-controlled channel access (HCCA). HCF uses the AP as a traffic manager which is termed as the hybrid coordinator (HC) (Chen et al., 2011), which is a centralized coordinator. The HC negotiates the exchange of frames and the frame handling rules given in HCF. The HC is located within the range of AP and works both in the contention-based and contention-free periods. The traffic is composed of wireless station (STA) "streams" or pipes, with each STA stream associated with a set of QoS parameters (Kowalski, 2013) negotiated with the AP. The AP uses a polling method to control the traffic. It sends polling packets to the stations. When a station is polled, it replies to the poll in a frame that contains the response and the data to be transmitted. In this method, the polling is based upon the priority on which QoS has to be ensured (Garg et al., 2003).

The various techniques for service differentiation covered in this section are summarized in Table 3 along with their main features and advantages.

Table 3
Comparison of the MAC layer service differentiation schemes using DCF in the IEEE 802.11 standard.

MAC scheme	Main features	Advantages	Drawbacks	References
Distributed Fair Scheduling (DFS)	The DFS algorithm uses the backoff mechanism as a function of packet length and the priority of the flow	Provides fairness to all the flows; performance of high priority flows is increased	It has a high implementation complexity	Lindgren et al. (2003)
Varying DIFS	Flow priority is given by setting different DIFS durations. To avoid collisions, a similar backoff time is maintained	Provides benefits to real-time applications where higher delay is more significant than lower packet loss	Low priority traffic suffers more, when high priority frames are queued	Aad and Castelluccia (2001)
Differentiated maximum frame length	Services are differentiated by defining maximum frame size proportionately to a flow's priority	Reduces contention overhead and achieves good differentiation	In noisy environments, longer packets are more likely to be corrupted than shorter ones, decreasing the service differentiation efficiency	Aad and Castelluccia (2000)
Blackburst	The blackburst scheme jams the channel according to the amount of time it has waited	Minimizes delay of real-time flows; high priority flows get maximum benefit in the absence of low priority flows	It requires constant access intervals for high-priority traffic, otherwise the performance degrades considerably	Sharma et al. (2013) and Ni et al. (2004)
Enhanced Distributed Coordinated Function (EDCF)	EDCF is a contention-based channel access function of IEEE 802.11e which can provide differentiated service	Provides better service differentiation using priority queues	When load increases, the network saturates and throughput decreases	Romdhani et al. (2003) and Qashi et al. (2011)
HCF Controlled Channel Access (HCCA)	The hybrid controller provides transmission opportunities to stations with higher priority packets	Combines the advantages of PCF and DCF and provides Guaranteed Services with a much higher probability than EDCF	Due to no admission control policy, it works well only until a certain traffic load and fails when load exceeds a QoS-load upper bound	Garg et al. (2003) and Viegas et al. (2013)

3.3.3. QoS scheduling

A priority scheduler always selects packets from a queue with the highest priority. Such an approach is simple to understand, but can unfortunately lead to starvation of lower priority packets, particularly when there is a steady flow of high priority packets. There are also deadline-based and rate-based scheduling schemes.

The process of QoS scheduling in the IEEE 802.11 standard chooses packets amongst the various flows and distributes them on to specific links depending upon the requirements of each flow. This distribution of flows on each link has to be done within a small time interval and should be hardware-friendly. Scheduling is designed to provide a better throughput while reducing transmission times—throughput and delay being the key metrics quantifying better QoS. Resource reservation for different traffic flows requires synchronization among nodes to effectively monitor the changes in resource adaptation (Yu et al., 2013). For such kind of insurances, we need to have a real-time monitoring mechanism for the changing network environment. The scheduler is generally operating at the MAC layer of the TCP/IP model.

In Zhang et al. (2009), a cross-layer design algorithm for QoS packet scheduling has been defined which considers delay and information shared at the PHY, MAC and network layers. It helps in high-speed data transmission through careful monitoring of the constant changes in the network while providing fairness to all flows. Note that, for best-effort services, the scheduling scheme treats all flows with the same priority. Several QoS scheduling techniques have been proposed (Lu et al., 1999; Tsao, 2000) and many enhancements of current techniques have been discussed (Grilo et al., 2003; Lim et al., 2004; Xiao, 2004).

In Ansel et al. (2004), a HCF-based packet scheduler is designed and implemented with special reference to the IEEE 802.11e standard. This design caters both the constant bit rate and variable bit rate of the QoS-sensitive traffic and provides bandwidth support and smaller delays to all network flows. The QoS scheduling provides guaranteed services taking into consideration the bit rate, delay, and throughput. The general algorithms that ensure QoS in scheduling are: (i) strict priority; (ii) weighted fair queueing (WFQ); (iii) Weighted Round Robin (WRR); and (iv) Earliest Due Date (EDD).

3.3.3.1. Strict priority.
In this algorithm, the buffer is partitioned into a number of different queues, which is equal to the number of different priority flows. The packets are then stored in these queues by the scheduler according to their own priority levels. The flows in the same queue are then sent using the FIFO scheme. The strict priority algorithm is easy to implement but it does not guarantee any bit rate and losses. Moreover, the lower priority flows may have a zero-valued throughput. In Georges et al. (2004), Jasperneite et al. (2002), and Georges et al. (2005), a network calculus method is used to evaluate the performance of a switch as it provides a good model of packet exchanges, and it determines end-to-end delay. Note that, the strict priority scheduling is implemented in Ethernet switches. A slight modification to the strict priority algorithm is proposed in Jiang et al. (2002), where the different flows are assigned with different parameters. The technique is important in the Per-Hop Behavior of Differentiated Services network.

3.3.3.2. Weighted fair queueing.
The same idea of assigning each flow with a certain priority is used, however the queues are not served on FIFO. Each flow is assigned a specific weight according to the QoS requirements (Parekh and Gallager, 1993). Hence, the bit rate varies with each flow. A certain upper bound on the buffer size is implemented to give all the flows a share of the bandwidth, which is unlike to what we have seen above. An interleaved WFQ scheme is implemented in Chen et al. (2005), where a table specifies the queue sequence. The table is interleaved, so higher priority flows are visited more frequently. The scheme improves on latency and jitter which are associated with the traffic queues. In Banchs and Perez (2002), the WFQ scheme that is backward compatible with the IEEE 802.11 standard is discussed. The simulation results show that the scheme can provide appropriate bandwidth distribution even in the presence of flows that need to be transmitted at all times.

3.3.3.3. Weighted Round Robin.
Weighted Round Robin is a frame-based implementation of WFQ. The flows are segregated similarly in separate queues with a specific weight assigned to each queue. The management can get difficult at times with different packet sizes. A new scheduling algorithm, called the dynamic WRR is proposed in Kwon et al. (1998). This algorithm is suitable for all traffic forms having variable and constant bit rates. The queues of traffic are assigned a dynamic weight. It helps the network in providing multimedia services even in the presence of bursty traffic. In Kwak et al. (2002), a modified dynamic WRR scheme is proposed. This scheme guarantees the delays in real-time traffic and provides efficient transmission of other forms of traffic.

3.3.3.4. Earliest Due Date.
In the normal EDD scheme for wired networks, packets of several different flows are assigned deadlines according to which packets are served first by the packet scheduler with the smaller deadline indicating higher priority. Since wireless networks show varying characteristics, the deployment of EDD is not an easy task. Therefore, in Elsayed and Khattab (2006), a channel-dependent EDD (CD-EDD) is described. It depends on the channel state, and the packets are queued by the scheduler on the basis of earliest expiry time and other channel parameters. The prioritized flow consequently gets the highest transmission rate among all the flows.

3.3.4. Traffic shaping

Traffic shaping is used to control the flows of traffic in a channel. The basic idea is to limit the amount of packets per station. A traffic controller is used to comply the QoS requirements of each flow. Traffic shaping can split the resources according to different requirements of different flows. The traffic shaper must adapt to the variations in a channel. The traffic shaping mechanism has a strong impact on the performance of a system (Morris et al., 2008). Several traffic shaping parameters are used in the QoS model of IEEE 802.11 standard: e.g., the aggregation level and the bursting level. *Aggregation level* refers to the amount of packets that are aggregated into a single IEEE 802.11 packet. *Bursting level* refers to the amount of packets transmitted at each Transmission Opportunity (Zhang and Ionescu, 2007).

3.4. QoS support in IEEE 802.11e

The IEEE 802.11e standard is an important extension of the IEEE 802.11 standard focusing on QoS (Mangold et al., 2003) that works with any PHY implementation. Wireless nodes equipped with IEEE 802.11e features are now known as QoS stations (QSTAs) and they are associated with a QoS access point (QAP) to form a QoS basic service set (QBSS). The main feature of the IEEE 802.11e standard is that it improves the MAC layer for QoS provisioning by providing support for: (i) segregation of data packets based on priority requirements; (ii) negotiation of QoS parameters through a central coordinator or AP; and (iii) admission control.

The IEEE 802.11e standard introduces a contention-based MAC layer scheme called Enhanced DCF (EDCF) and a polling-based scheme called HCF Controlled Channel Access (HCCA). Both these

schemes are useful for QoS provisioning to support delay-sensitive voice and video applications (Choi et al., 2003), and they are described next.

3.4.1. Enhanced DCF (EDCF)

In the DCF configuration, a Contention Window is set after a frame is transmitted. This is done to avoid any collisions. The window defines the contention time of various stations who contend with each other for access to channel. However, each of the stations cannot seize the channel immediately, rather the MAC protocol uses a randomly chosen time period for each station after that channel has undergone transmission (Yang and Vaidya, 2002).

EDCF uses this Contention Window to differentiate between high priority and low priority services (Romdhani et al., 2003). The central coordinator assigns a Contention Window of shorter length to the stations with higher priority that helps them to transmit before the lower priority ones (Krithika and Pushpavalli, 2012; Qashi et al., 2011). To differentiate further, inter-frame spacing (IFS) can be varied according to different traffic categories. Instead of using a DIFS as for the DCF traffic, a new inter-frame spacing called arbitration inter-frame spacing (AIFS) is used. The AIFS used for traffic has a duration of a few time slots longer than the DIFS duration. Therefore, a traffic category having smaller AIFS gets higher priority (Villalón et al., 2005).

3.4.2. HCF Controlled Channel Access

The HCF Controlled Channel Access (HCCA) is IEEE 802.11e specific, and it makes use of a Hybrid Coordinator (HC) to manage the bandwidth allocation of wireless medium (Khan et al., 2013). The HC can obtain a Transmission Opportunity (TXOP) and initiate data deliveries to provide transmission opportunities to a station with a higher priority without any backoff; that is to say, the HC can access the channels after a PIFS amount of time rather than a DIFS amount of time as for the other stations (Mangold et al., 2003). As PIFS is smaller than DIFS and AIFS, the HC has a priority over the DCF traffic, and also over the ECF traffic that uses AIFS.

3.4.3. Control Access Period (CAP)

The CAP in HCCA is a period when access to the wireless medium is controlled (Ni, 2005). During this time, the HC, or the AP, gives the right of using the medium to a device. The AP can gain access to the medium before any other stations, and can then provide the Transmission Opportunity to any station. This guarantees data transfer from a station irrespective of the congestion level in the channel (Rashid et al., 2008). The AP can schedule such transmission opportunities for each of the stations, and can provide the parameters needed for QoS provisioning (Dujovne et al., 2010).

3.4.3.1. CAP in the contention period.
A CAP in the contention period is used to regulate access to the medium to guarantee various QoS parameters (Reddy et al., 2006). However, in this method, the AP is not the controller (or the sole decision maker) (Gu et al., 2004). This is because any stations having the DCF traffic, or any EDCA traffic, can interfere with the scheduling performed at the AP, so this can delay the already scheduled data transfer at a particular station. Moreover, CAPs may use a RTS to prevent other devices while contending for the medium (Ni, 2005). This causes a marginal overhead (Rashid et al., 2007).

3.4.3.2. CAP in the contention free period.
The contention free period (CFP) is the most efficient way to use the channel, and it allows the AP to have a fine control of the medium (Cervello et al., 2006). During this period, the AP has full command of the medium, so the stations do not contend for the access to the

medium (Yeh and Chen, 2002). The scheduling of the traffic, and the provision of QoS guarantee to the stations, is handled by the AP itself. The AP can set multiple CAPs following each other and uses the smallest possible time intervals to separate every CAP (Rashid et al., 2007).

3.4.4. Other IEEE 802.11e QoS features

We now outline some other important features of the IEEE 802.11e standard. The TXOP parameter defines a time limit for the utilization of radio resources at the stations (Mangold et al., 2003). The Automatic Power Save Delivery (APSD) mechanism is used by the AP to deliver multiple frames within a service period. Thereby, APs can enter sleep period until the next service period to conserve energy (Pérez-Costa and Camps-Mur, 2010). The APSD mechanism has scheduled APSD and unscheduled APSD, both of which lead to power saving as compared to the legacy IEEE 802.11. The IEEE 802.11e standard also supports block ACKs for the acknowledgment of multiple MAC Protocol Data Units (MPDUs) in a single block acknowledgment frame resulting in reduced overhead (Tinnirello and Choi, 2005).

The NoAck is another enhancement that indicates the loss of a packet, so retransmission can be ensured quickly to reduce delay (Politis et al., 2011). Direct link setup is another supported feature that allows direct station-to-station transfer within a service set.

4. Network layer QoS solutions for IEEE 802.11

The bulk of research investigating QoS solutions for the network layer of IEEE 802.11 networks has focused on admission control and QoS routing. These two important facets of network layer QoS solutions are discussed next in separate subsections.

4.1. Admission control

Although the enhancements explained at the MAC layer provides service differentiation among different traffic flows, it can ensure QoS only when network load is reasonable. If the load increases beyond a certain limit, the QoS guarantees are not ensured even to high priority traffic (Mangold et al., 2002; Xiao and Li, 2004). This is where the admission control mechanism helps in preventing the network from becoming congested, by allowing or disallowing flows depending on whether the conditions are favorable to meet QoS requirements. More specifically, the purpose of admission control is to limit the amount of newly admitted traffic such that the QoS performance of existing flows is not degraded (Gao et al., 2005). Admission control is a key component to adapt to the traffic variations according to the changing environment of IEEE 802.11-based wireless networks (Andreadis and Zambon, 2012). In Hanzo and Tafazolli (2009), they have presented a very comprehensive survey on different admission control schemes available in the literature. Admission control can be categorized into three different methodologies (Brewer and Ayyagari, 2010).

4.1.1. Measurement-based admission control

In this scheme, the decisions are made through continuous monitoring of network status, such as throughput and delay. A certain threshold is maintained according to the network status for admission of new traffic flows. Nor et al. (2006) proposed a metric called Network Utilization Characteristic (NUC) as a means for admission of traffic flows into network. NUC defines the amount of channel utilized to transmit the flow over the network. This scheme guarantees QoS to high priority flows under loaded channel environments. Another scheme presented by Wu et al. (2010) is that each traffic class is assigned a certain portion of

available resources, and these resources are then remaining reserved for that particular class. In this regard, only the traffic with higher priority compared to the existing traffic is admitted.

4.1.2. Model-based admission control

In model-based schemes, the network status is measured based on some models. The Markov chain models are quite popular in attempts at modeling IEEE 802.11 although other approaches are also being explored due to some limitations of Markovian models (Chen et al., 2006). In Cano et al. (2007), an analytical model is used to estimate the minimum bandwidth requirement of all flows. When a newly admitted flow need to be activated, the algorithm checks if it is going to result in preservation of QoS requirements of existing flows.

4.1.3. Measurement-aided, model-based admission control

It is a hybrid of measurement-based and model-based schemes. The algorithm in Ksentini et al. (2007) takes network measurements in a loaded environment and also the data rate requirements of the flow that is requesting for admission. Furthermore, a channel model is applied to predict the network conditions and provides QoS enhancements accordingly. Another solution is the threshold-based approach proposed in Bensaou et al. (2009) in which the channel conditions are continuously monitored and the contention probability is measured. When any new flows request for admission, the admission control checks for the competing flows. The absolute bandwidth and the expected delay of the new flow are measured. If this satisfies the threshold conditions, then this flow is admitted.

4.2. QoS routing schemes

QoS routing is an essential part of the overall QoS architecture in the IEEE 802.11 standard. QoS routing allows the network to compute a path that supports the QoS objectives of various flows under the constraints of wireless medium. The chosen path may or may not be the shortest path, but it meets a particular service category objectives (Sivakumar et al., 1999; Yin et al., 2006).

There are various metrics that can be used for measuring QoS routing performance. We describe an example work for each metric. The metrics proposed for QoS-based routing in the literature are: (i) minimum throughput, or capacity, required in bits per second (Lin and Liu, 1999); (ii) maximum tolerable delay in seconds (Chen and Nahrstedt, 1999); (iii) maximum tolerable packet loss ratio (PLR) (Abdrabou and Zhuang, 2006); and (iv) maximum tolerable jitter (Bashandy et al., 2005; Wang and Kuo, 2005). In addition to these generic metrics, there are also other metrics specific to various layers of TCP/IP model. For example, at the *network layer*, achievable throughput or residual capacity (Lin and Liu, 1999), end-to-end delay (Chou et al., 2006; Chen and Nahrstedt, 1999), node buffer space (Sheng et al., 2003), and route lifetime (Rubin and Liu, 2003) are important metrics; at the link (or MAC) layer, link reliability (Barolli et al., 2003), and link stability (Rubin and Liu, 2003) are important performance metrics; finally, signal-to-interference ratio (SIR) (Kim et al., 2004), Bit Error Rate (BER) (Wisitpongphan et al., 2005), and node's residual energy (Toh, 2001) are important performance metrics at the PHY layer. A comprehensive summary of QoS-based routing metrics is provided in a survey paper (Hanzo and Tafazolli, 2007).

Many QoS routing works have been proposed in the literature. For instance, Matos et al. proposed that routing decisions of voice, video, and data can be computed in a decentralized fashion at intermediate nodes in wireless multi-service networks such that the overall network performance is optimized per the desired QoS

(Matos et al., 2012). There are a number of frameworks that can be used for QoS-based routing. We describe two such frameworks.

4.2.1. Measurement-based QoS routing

In Liu et al. (2004), a framework is provided to achieve fairness among different priority flows. The basic concept is to allow the high priority traffic to help the low priority traffic by sharing their unused bandwidth. This scheme can considerably improve the system performance and it can shorten the delays when the traffic load is very high. The QoS-supporting algorithm presented in Chen and Heinzelman (2005) helps applications to find routes that satisfy their service needs, or a feedback is provided in case of non-availability of these resources. Thus, a protocol that is QoS-aware and also has admission control and feedback mechanism is proposed. Another resource reservation algorithm is proposed by Xue and Ganz (2003) in which bandwidth and delay are measured very accurately using a MAC protocol with collision detection. These calculations are then used by the algorithm to make decision in admission and reservation of resources.

4.2.2. Ticket-based probing algorithm

This algorithm uses tickets to limit the number of paths observed. When any source wants to get a QoS satisfying path to any destination, it sends probe messages along with a ticket. The number of tickets is equivalent to the number of paths searched. When the destination receives this probe message, the path from source to destination is set (Chen and Nahrstedt, 1999).

5. Transport layer QoS solutions for IEEE 802.11

The classical version of TCP protocol performs rate control based on its assumption that packet losses occur solely due to network congestion. This assumption does not hold true for wireless networks where channel noise and interference can be another significant cause of packet loss. With such an assumption, TCP performs poorly in terms of end-to-end QoS since it may reduce the sending rate even in uncongested networks under the mistaken assumption that packet losses are only caused by network congestion.

The problem of suboptimal performance of TCP in wireless networks has been known for long and much effort has focused on improving TCP's performance (Balakrishnan et al., 1997; Chen et al., 2002). Most of the existing work in the IEEE 802.11 standard is focused on the QoS requirements of multimedia applications, such as VoIP (Jelassi et al., 2012), and data traffic, such as web, email, and media downloads. But with growing demand of wireless networks, time-critical applications with voice and video do place significant QoS requirements on wireless medium. To upgrade QoS at the transport layer, we can adopt several techniques taking into account the delay and loss as the basic parameters. With the main focus on congestion control and error control, several techniques are discussed below.

5.1. Congestion control

The bursty nature of the wireless media and path loss causes the degradation of services to applications that require high video quality. This degradation is caused by network congestion. This is why TCP congestion control is essential to minimize packet loss and reduce delay. Rate control is a congestion mechanism which reduces network congestion by comparing the required bandwidth for video with the available bandwidth (Bolot and Turletti, 1998). Multiple standards of rate-adaptive video encodings (Wu et al., 2000) exist for different applications, such as H.261 and H.263 for video conferencing (Martins et al., 1996; Wiegand et al.,

1996), as well as MPEG-1 and MPEG-2 for real-time transmission (Ding, 1997; Hsu et al., 1997). The main purpose of a rate-adaptive encoding scheme is to enhance the video quality under a certain encoding rate. Rate control and rate shaping (Ding, 1997) are the algorithms for congestion control in the IEEE 802.11 standard.

5.1.1. Rate control

It is very important for an end-to-end protocol to accurately estimate the appropriate sending rate for network transfer since an infeasibly high sending rate can result in packet losses and retransmissions. TCP retransmissions that result from packet losses may lead to unacceptably long delay for QoS-aware multimedia delivery over the wireless channels. Rate control subsumes flow control and congestion control which adjust sending rates to ensure that the sender's rate does not overwhelm the receiver and the network, respectively.

Two types of congestion control are in wide practice: window-based (Jacobson, 1988) and rate-based (Turletti and Huitema, 1996). The window-based approach analyzes the available network bandwidth by gradually increasing the size of congestion window. When congestion is detected (through the detection of packet loss), the protocol decreases the window size by a large amount. The abrupt decrease in the window size in response to congestion is necessary to prevent network failure. Window-based control performs retransmissions which result in extensive delay, which is intolerable in case of real-time video transmission. The rate-based control approach sends at a rate based on an estimated available network bandwidth. If the estimated bandwidth is accurate, then network congestion can be avoided. The rate-based control approach is usually used for transportation of real-time video. Existing rate control mechanism for real-time video are source-based, receiver-based or hybrid. The interested reader is referred to the paper Zhu (2011), and the references therein, for a description of rate-control work focusing on supporting real-time traffic in WLANs.

5.1.2. Rate shaping

Rate shaping is the practice in which the compressed video bit stream adjusts itself to the rate of a target rate. We can think of a rate shaper as an interaction medium between an encoder and the network, which matches the encoder's output to the available network bandwidth. Since rate shaper does not need any interaction with the encoder, it can be used for any video coding scheme for both live and stored videos.

5.2. Error control

The QoS guarantees can also be made through error control. The main role of congestion control is to avoid packet loss. However, we are unable to avoid packet loss completely in the Internet, and as a result the quality of video or other bandwidth hungry applications are affected. The error control schemes at the transport layer are application-aware. The error recovery schemes can be divided into two basic types:

5.2.1. Automatic Repeat reQuest (ARQ)

The ARQ scheme uses an acknowledgement packet to indicate that a packet has been received successfully. It is very efficient for high-speed wireless links because the round trip delay of the link is very small (Name et al., 2002). The ARQ scheme can be implemented at both transport and link layers of the OSI model (Crow et al., 1997). The traffic is segmented into queues such that QoS guarantees are ensured. The ARQ scheme can adapt to channel errors and is more efficient in terms of bandwidth utilization.

5.2.2. Forward Error Correction (FEC)

The FEC scheme adds redundant bits to the flow which helps in recovering the erroneous bits. The FEC is used for the transmission of real-time applications which have a strict delay requirements (Liu et al., 1997; Aikawa et al., 1996). However, a drawback of FEC is the increased overhead even in the absence of errors in the link (Choi et al., 2006). FEC helps in maintaining a uniform throughput and time delay in the networks. However, the overhead increases with channel errors because long FEC codes must be used.

5.3. Prioritization

The TCP ACK prioritization method uses both the AIFS and the minimum Contention Window (CWmin) parameters. The stations having smaller CWmin gets more transmission opportunities than stations having larger value of CWmin as their backoff counter is smaller. The AIFS parameters can be used to allow the AP to have quicker access to the wireless medium. Since the TCP ACKs can go freely through the bottleneck links, the performance of the system is upgraded (Leith et al., 2005).

6. Application layer QoS solutions for IEEE 802.11

Apart from work at the lower layers, it is also possible to implement QoS in IEEE 802.11 networks at the application layer. Traditionally, the work done on application layer QoS has focused on various aspects of multimedia delivery using techniques such as scalable video coding (van der Schaar et al., 2006), error correction coding and rate-distortion optimization (Chakareski and Girod, 2003), source coding (Luo et al., 2006), transcoding (Luo et al., 2008), adaptive transmission (Luo et al., 2008), as well as rate control (Luo et al., 2006). We note here that most of the QoS work done at the application layer is cross-layered in nature, specifically drawing upon information from, and interacting with, other layers. We provide two brief examples as illustration.

Chakareski and Girod (2003) proposed an optimization-based *error-correction coding* scheme, which works at the application layer, for rate-distortion optimized multimedia streaming to wireless clients. The scheme allows the sender to compute the prioritization levels of packets to satisfy an average transmission rate constraint while minimizing the average end-to-end distortion.

van der Schaar et al. (2006) proposed a framework for delay-constrained video streaming over IEEE 802.11 a/e WLANs. The authors considered the problem of video transmission over HCCA and developed a cross-layered optimization framework working at the PHY, MAC, and application layer of the TCP/IP layered model. In another cross-layered video streaming paper, Li and VanderSchaar (2004) proposed an error protection scheme for the transmission of layered coded video to provide adaptive QoS through prioritized queuing at the network layer and limitation of retries (or retransmissions) at the link layer. The basic insight of this work is that different video layers of varying importance may not receive uniform processing and protection, but may receive unequal priority depending on the channel conditions.

The techniques for enhancements of various layers (e.g., MAC, network, transport and application layers) are summarized in Table 4.

7. Cross-layer QoS solutions for IEEE 802.11

While most QoS enhancement techniques are implemented at the MAC layer, wireless QoS can benefit from cross-layered interaction and implementation (Toumpis and Goldsmith, 2003).

Table 4
Layered overview of QoS techniques in the IEEE 802.11-based wireless networks.

QoS enhancement technique(s)	Description	QoS metri cused	Reference(s)
MAC layer			
Priority queueing	Data packets are segregated based on their priorities in queues. The packets with the highest priority are transmitted first and so on	Buffer fullness	Sundareswaran et al. (2007) and Ni and Turletti (2004)
Distributed Fair Scheduling	Each flow is assigned bandwidth according to its priority	Backoff interval	Lindgren et al. (2003)
Varying DIFS	DIFS is varied in order to differentiate among flows. Each flow's priority is set by giving it a different DIFS	Time to transmit	
Maximum frame length	High priority stations can transmit larger frames comparatively	Packet length	Drabu (1999)
Enhanced DCF	Shorter Contention Windows are assigned to higher priority stations helping them to transmit first	System-level data rate	Krithika and Pushpavalli (2012) and Villalón et al. (2005)
Blackburst	Contention period is used to indicate the waiting time for medium access	Normalized MAC load	Ni et al. (2004)
Network layer			
Admission control	The network is thoroughly examined and when congestion occurs, the nodes decrease their best-effort traffic in response	Loss, Jitter	Domingo and Remondo (2004) and Mangold et al. (2002)
QoS Routing	Some routing mechanisms are used under which QoS paths are determined. QoS path may or may not be similar to the shortest path	Delay, PLR	Zhang and Mouftah (2005), Sivakumar et al. (1999), Yin et al. (2006), and Matos et al. (2012)
Transport layer			
Congestion control	The congestion control mechanism orders the source to transmit traffic at a rate that is not greater than the available network bandwidth	Application-level data rate, Buffer fullness	Bolot and Turletti (1998), Jacobson (1988), and Turletti and Huitema (1996)
Error correction	Forward Error Correction and Automatic Repeat reQuest are used to ensure reliability	BER, PLR	Name et al. (2002) Aikawa et al. (1996), and Crow et al. (1997)
TCP ACK prioritization	The Contention Window and AIFS are used for flow prioritization	Residual capacity, PLR	Leith et al. (2005)
Application layer			
Scalable Video Coding	A cross-layered framework is proposed for delay-constrained video streaming over IEEE 802.11 a/e WLANs. The framework works at the PHY, MAC, and application layers of the TCP/IP layered model	End-to-end delay, PLR, Jitter	van der Schaar et al. (2006)
Transcoding	Video transcoding reencodes the stream to adapt the bit rate to the available resource	Average Service Response Ratio, Startup Latency, Cache hit ratio	Lee and Lee (2014) and Van Beek et al. (2004)
Application layer error control	An application layer optimization-based error correction coding scheme for rate-distortion optimized multimedia streaming to wireless clients	End-to-end distortion, Relative one-way trip time (ROTT)	Cen et al. (2003) and Chakareski and Chou (2004)
Real-time retry limit adaptation	A real-time retry-limit adaptation is proposed at the link layer for video with adaptive QoS	Buffer capacity, Delay, BER	Li and VanderSchaar (2004)
Hybrid ARQ/ FEC	Multicast and unicast real-time video streaming approaches over WLANs are implemented through a hybrid ARQ algorithm that combines FEC and ARQ	Latency, Reliability, Energy efficiency	Majumda et al. (2002) and Rao et al. (2014)

Since QoS provisioning entails various issues that span the range of the TCP/IP layered stack, cross-layer solutions are finding increasing deployment.

This section discusses a few cross-layered solutions for implementing QoS in wireless networks in general, and in IEEE 802.11 networks in particular. The various techniques for cross-layer enhancements discussed in this section are summarized in Table 5.

7.1. Cross-layer features for wireless multimedia

With increasing multimedia traffic on the Internet and wireless access being anticipated to become the future predominant Internet technology (Zander and Mahonen, 2013), delivering multimedia applications with enhanced QoS has become extremely important. Wireless channel information cannot be predicted easily due to deep fades and multipath effects, but getting information of source motion using video sequence is not hard. Due to the error-prone nature of wireless medium, and the undifferentiated nature of Internet's best-effort service model, multimedia delivery over the wireless networks is technically challenging. To improve user experience over the wireless Internet, QoS support can be introduced at different layers.

In Zhang et al. (2005), the authors propose a cross-layered architecture—combining application-level, transport-layer, as well as link-layer controls—for supporting multimedia delivery over wireless Internet. These controls incorporate issues such as dynamic estimation of network and channel; adaptive error control, congestion control, and ARQ mechanisms; and priority

based scheduling. In another work (Qu et al., 2006), a cross-layer framework is proposed in which the source motion is captured from a video sequence, and it consists of a packetization scheme, a cross-layer FEC-based unequal error protection scheme, and an intracoding rate selection scheme. This significantly improves transmission of bursty traffic and its losses over the wireless network without making the system complex.

In order to ensure QoS in real-time applications, we can bound delay instead of high spectral efficiency (Indumathi and Murugesan, 2010). For example, we can ensure QoS for multimedia applications by analyzing the impact of the PHY layer on the data link layer. The single input and single output (SISO) mechanism is used with Adaptive Modulating Schemes at the physical layer; and at the data link layer, we check the impact of the physical layer on the delay bound. Physical layer is modeled using a finite-state Markov chain. The use of appropriate scheduling schemes, and the resources allocated to the users can thus ensure high QoS for each user. This technique allocates resources to real-time users in time slots in a dynamic method using SISO along with adaptive modulating codes.

7.2. Adaptive Modulating Scheme (AMC)

There are various works in the literature that have proposed combining QoS reservation and scheduling at the MAC layer with adaptive modulation and coding (AMC) at the PHY layer. For instance, Liu et al. (2005) have proposed an hybrid architecture combining QoS reservation and scheduling at the MAC layer with

Table 5
Overview of cross-layer enhancement techniques in the IEEE 802.11-based wireless networks.

Feature(s)	Description	References
Wireless multimedia	SISO is used with adaptive modulation schemes at the PHY layer; and the impact on delay bound is measured at the data link layer	Indumathi and Murugesan (2010)
Adaptive Modulating Scheme	The node with the highest priority is assigned resources first, and each type of connection adopts AMC at the PHY layer	Liu et al. (2006)
Wireless scheduling	The information received from the PHY layer is utilized, and an efficient cross-layer packet scheduling approach is proposed which guarantees QoS parameters like delay, BER and received signal strength	Abd El-atty (2006)
Cooperative communications	QoS is ensured by using cross-layer design taking into account the PHY and network layers using cooperative communication	Sheng et al. (2011)
Channel coding and retransmissions	The use of hybrid schemes corrects the errors in an adaptive manner by using a combined scheme of channel coding and retransmissions, thus improving the TCP performance	Girod and Farber (1999) and Farber et al. (1999)
Dynamic priority functions	Defined for new nodes which are updated dynamically. AMC and MIMO are used accordingly for QoS provisioning	Liu et al. (2006)
Cross-layer perceptual ARQ	A cross-layer priority-based ARQ algorithm for H.264 video streaming in IEEE 802.11 wireless networks	Bucciol et al. (2004)

AMC at the PHY layer. With AMC, the physical transmission parameters can adapt to the changes in the link quality. As an example, the PHY layer can fallback to a modulation scheme more robust to noise, such as BPSK instead of QAM-16, in the case link quality degrades.

In Agarwal and Jagannatham (2013), the authors derive an optimal policy to reduce the average amount of dropped packets of a delay-controlled wireless node. The presented framework utilizes adaptive modulation for transmission of the optimal amount of packets to satisfy the QoS requirements. This framework, devised as a Markov Decision Process (MDP), works on reducing the long-term packet drop rate. In another work (Liu et al., 2006), a cross-layered approach is presented for mobile wireless networks which studies the impact of the PHY layer infrastructure over the link layer QoS performance. This work considers MIMO diversity schemes along with AMC in its PHY layer analysis, while also studying the impact of the PHY layer infrastructure on real-time multimedia QoS provisioning performance at the link layer.

7.3. Wireless scheduling

The interaction between packet scheduling and the PHY layer is studied in Abd El-atty (2006) where the network is used efficiently by predicting the future state of the wireless channel, as well as controlling the transmission power in multipath fading wireless CDMA networks. It is based on cross-layered model in which the information received from PHY layer is utilized by the scheduler, and an efficient cross-layer packet scheduling is proposed which guarantees users guaranteed QoS performance in terms of delay, BER, and received signal strength.

7.4. Cooperative communications

QoS can also be ensured by using cross-layered design taking into account the PHY and networking layers using cooperative communication (Sheng et al., 2011). Cooperative communication is first studied at PHY layer, followed by routing to ensure QoS in the network: i.e., we use an optimized link cost for the decision making of our routing leading to better path selection. The power consumption is kept to a minimum possible value, and end-to-end reliability is achieved by reducing the error rate. The selection of the best possible path leads to end-to-end reliability and thus the ensuring of QoS. In Xianyang and Feng (2014), the authors describe a novel concept named cooperative QoS routing, which sets up a routing path that helps to satisfy user's bandwidth requirement. The authors propose an optimization problem, called 'widest cooperative routing path' (WCRP) problem, which finds a cooperative routing path with the maximum uninterrupted bandwidth and a scheduling scheme to evade interference.

7.5. Combining channel coding and retransmissions

The transport layer ensures a reliable transmission by re-sending corrupted packets due to congestion. However, packets may get corrupted in wireless networks due to other reasons such as fading and multipath effects. The requests for repeated transmission for the packet in such cases would negatively impact the performance of the network. To cater to this problem, joint cross-layer techniques are used. FEC and ARQ are used as error correction codes; where FEC is used in delay-sensitive applications (Boutremans and Le Boudec, 2003), while ARQ is used for delay-tolerant applications (e.g., audio/video streaming). In most of the applications, FEC does not negotiate with the receiver for error correction because of the delay-sensitive nature although feedback from the receiver can be effective (Girod and Farber, 1999; Farber et al., 1999). Using hybrid schemes corrects the errors in an adaptive manner by using combined channel coding and retransmissions thus improving the TCP performance. Recently, network coding has been extensively used in wireless networks to upgrade the limited wireless capacity. In WLANs, network coding can be applied to packet retransmission. More than one packet can be evenly transmitted by a single retransmission at base station. In Tanigawa et al. (2011), the retransmission is based on network coding, and it cooperates with IEEE 802.11e EDCA. Consequently, QoS of high priority group is upgraded from the aspect of efficient loss recovery.

7.6. Joint cross-layer techniques

This subsection presents four main joint cross-layer techniques.

7.6.1. Joint rate control, admission control, and scheduling

The main goal of an Internet designer is to share the resources efficiently. To use the scarce bandwidth in an efficient way, it is usually the case that real-time traffic co-exists with the rest of the traffic. We can jointly solve rate-control, admission-control, and scheduling problems for optimized solutions. In infrastructure-based networks, the EDCF is providing QoS to the nodes; but in distributed multihop networks, it is not possible due hidden terminal and interference problems. Thus, admission control puts a limit on real-time traffic that can overwhelm the system. By collaboration with nearby nodes, the interfering best-effort traffic is cut down to reduce the contention of real-time traffic. A scheme called SoftMAC is proposed in Wu et al. (2007) to use a control mechanism in order to regulate real-time and best-effort traffic in a distributed manner by coordinating with neighboring nodes. SoftMAC uses admission control to make sure that real-time traffic

has sufficient bandwidth along its path. It also caters for rate control to avoid collisions between the real-time traffic and the best-effort traffic. Moreover, it has a priority queueing module to prioritize the real-time traffic.

QoS for wireless networks can be ensured by joint solutions working at the PHY and network layers. Such solutions can include joint routing and rate allocation to ensure QoS for different applications. Zhang et al. have proposed a framework for cross-layer design for QoS support in multihop wireless networks, and have reviewed in detail the interplay between joint routing at the network layer and rate allocation at the transport layer. In another paper, Zhu et al. (2007) discuss an optimization of joint allocation rate and multipath routing which allocates rates depending upon the distortion rate and congestion level. This paper considered the problem of rate allocation for multi-user video streaming sharing multiple heterogeneous access networks. The problem was formulated as a convex optimization problem and distributed approximation of the optimization was proposed.

7.6.2. Joint power control, scheduling, and routing

Various works in the literature have demonstrated the deficiencies of the traditional siloed approach of independently performed power control, scheduling, and routing in wireless networking (Zhang and Zhang, 2008; Toumpis and Goldsmith, 2003). It is worth highlighting the strongly coupled nature of the these problems by noting that a change in power allocation, or the schedules on a given link, can impact flows that do not utilize the modified link. Due to the strong coupling between the network, MAC, and PHY layers, the power control, scheduling, and routing problems are best addressed jointly.

With joint scheduling and power control, a network can generally achieve higher throughput and lower delay in a network (Zhang and Zhang, 2008), although for some unbalanced topologies, scheduling alone cannot satisfy bandwidth requirements and rerouting is also needed to send some packets through alternative routes and thereby release congestion. The routes are then selected according to the joint metric of energy consumed and traffic accumulated, with priority given to nodes with longer queue length. A similar approach is presented in Cruz and Santhanam (2003) which aims to minimize the total average transmission power in a wireless multi-hop network through optimal link scheduling and power control. This work, however, requires tight time synchronization between transmitters, and quasi-static channel conditions that remain constant over several time slots, thus limiting the applicability of this work to interconnecting stationary nodes.

7.6.3. Joint scheduling and rate optimization for opportunistic transmission

In order to utilize the scarce resources of the wireless networks, opportunistic transmission takes advantage of the varying nature of the channel and improves throughput of the network. Two approaches are used in this regard: the first one uses the time diversity of an individual link by changing the transmission rate according to the channel conditions (Moh, 2009; Liu and Knightly, 2003), while the second one considers multi-user diversity and jointly considers the time and spatial heterogeneity of a channel. Wang et al. (2004) have proposed a MAC solution named 'Opportunistic Scheduling and Auto Rate' (OSAR) which jointly considers rate adaptation and multi-user diversity. OSAR protocols exploits the channel variations by automatically adjusting the sending rate to best match the channel conditions. In a followup work, Wang et al. (2006) have proposed another solution 'Opportunistic Medium Access and Adaptive Rates' (OMAR), which aims at efficient utilization of the shared medium in IEEE 802.11-based

ad-hoc networks through joint consideration of multi-user diversity, distributed scheduling, and adaptivity. OMAR uses a clustering framework in which a node with a predefined number of links can function as the clusterhead to locally coordinate multiuser communication. The clusterhead is responsible for initiating medium access, while the cluster embers make medium access decisions in a distributed manner.

7.6.4. Joint channel assignment and routing

In this section, we present techniques considering the data link layer and network layer jointly. Interference among the channels is one of the main hurdle in achieving QoS in wireless networks. Orthogonal channel assignment is a potential solution to this problem. In Kyasanur and Vaidya (2006), the authors present a joint algorithm for channel assignment and routing. The proposed solution allows the creation of multiple routes in the network to achieve higher throughput. The channel assignment algorithm take into account two factors: the first is network topology, while the second is information about the traffic.

8. Modern network architectural trends and wireless QoS

This section presents three types of modern network architectures. Various research areas of recent architectures discussed in this section are summarized in Table 6.

8.1. Software-defined wireless networks (SWNs)

With increasing deployment and diversification of wireless technology, managing wireless networks has become very challenging. Software-defined networking (SDN) is a promising architecture that can be used for conveniently operating, controlling, and managing wireless networks. The defining characteristic of SDN is generally understood to serve as the separation of the control and data planes. The presence of programmable controllers to adjust the operating parameters enables us to call these networks 'software defined'.

Traditionally, networking devices, such as firewalls and routers, require vendor-specific software for programming their operating parameters. This programming can be done manually by a network administrator through the command line interface (CLI). This limits the margin of innovation that can be incorporated into the modern networks such as the world wide web or the WLANs. SDN changes this notion of network programming by extracting the control intelligence from the data plane and managing all the data plane devices at centralized controller(s) (Nunes et al., 2014). Figure 3 shows a traditional network where the control and data planes are co-located in each networking device, and so the network control is decentralized. In comparison, SDN has a centralized architecture where the central SDN controller is controlling the multiple data planes; specifically, southbound APIs are used to communicate with data plane, and northbound APIs are used to communicate with SDN applications. SDN provides the flexibility of programming a network through the control plane. This can help in simplifying network management and operations. The rest of this subsection presents QoS efforts for IEEE 802.11-based SWNs.

8.1.1. QoS efforts for IEEE 802.11-based SWNs

Traditionally, SDN has been focused on campus and data center networks; but increasingly, researchers are also focusing on wireless access networks. Since it can be difficult to manually configure various QoS knobs, OpenFlow and SDN architectures can be very useful in automating scalable control of network QoS based on high-level descriptions of application/service requirements. For instance,

Table 6
Sample of research on QoS with recent architectural developments.

Project	Description of QoS enhancement technique(s)	References
Software defined wireless networks		
meSDN	Achieves real-time detection of QoS demands in a network and provides end-to-end QoS control	Lee et al. (2014)
QoSFlow	Provides packet scheduling algorithm to improve QoS mechanism in OpenFlow/ SDN-based networks	Ishimori et al. (2013)
OpenQoS	Provides a dynamic routing scheme that generates shortest path for data delivery in order to minimize packet loss and latency	Egilmez et al. (2012)
OpenFlow controller for multimedia delivery	Facilitates multimedia delivery with QoS using the best path with optimal service configuration	Kassler et al. (2012)
FlowVisor	Supports 'QoS-enabled network slicing' that provides a user or an application with a certain network capacity "slice", which is isolated from other coexisting slices servicing other users/networks	Sherwood et al. (2009)
Multimedia Streaming QoS Architectures for SDN	Provides QoS extensions for multimedia delivery using distributed control architecture in multi-operator SDNs	Egilmez and Tekalp (2014)
Interference mitigation in enterprise WLAN	Proposes an OpenFlow-based framework for interference mitigation in enterprise WLANs	Zhao et al. (2014)
OpenFlow-based QoS support for Ofelia	Proposes architectural extensions to make Ofelia a QoS-supporting federated experimental testbed	Sonkoly et al. (2012)
PANE	An SDN-based API for participatory networking was proposed that can be used by a user to reserve minimum guaranteed bandwidth between two hosts	Ferguson et al. (2013)
OpenQFlow	Proposes a flexible variant of OpenFlow supporting a two-tiered flow-based QoS framework	Nam-Seok et al. (2013)
SDN for future Internet	Proposed a SDN-based QoS framework that implemented adifferentiation framework for OpenFlow networks. The framework was tested in failure-conditions using single and multiple autonomous system scenarios	Sharma et al. (2014)
SDN control framework for QoS	Programming of network devices to provide QoS guarantees The SDN controller manages the traffic according to data collected	Tomovic et al. (2014)
AQSDN	Proposed an autonomic QoS management mechanism for SDNs (AQSDN). The various QoS features in AQSDN are configured autonomically using OpenFlow and OF-config protocols	Wang et al. (2014)
FlowQoS	FlowQoS is an SDN-based QoS framework comprising two components: firstly, a flow classifier mapping application traffic to different parts of flow spaces, secondly, a SDN-based rate shaper	Seddiki et al. (2014)
Cloud−based wireless networks		
EDCA model for cloud	Proposes a model for QoS-aware differentiated multimedia cloud service provisioning in WLAN networks	Tursunova and Kim (2012)
Resource allocation in clouds	Proposes energy-aware resource allocation mechanisms for data centers set up in cloud environments	Sharkh et al. (2013)
Dalvi et al	Proposes centralized cloud-based approaches for managing WLANs	Dalvi et al. (2011)
CloneCloud	Utilizes computation offloading through elastic execution between mobile devices and cloud	Chun et al. (2011)
LWAPP (RFC 5412)	Proposes lightweight access point protocol (LWAPP) for centralized cloud-based WLAN management	Calhoun (2010)
CloudMAC	Enables APs to redirect MAC frames only. Processing of MAC data is done via cloud computing infrastructure	Dely et al. (2012)
IEEE 802.11 on cloud-based radio over fibre	Conducts a study on the feasibility of the architecture of IEEE 802.11 on cloud-based radio over fibre	Zhang and Franklin (2014)
Cloud-based multimedia mobile streaming	The proposed multimedia cloud computing solution considers the overall network environment and adjusts the transmission frequency and multimedia transcoding parameters accordingly	Lai et al. (2013)
Cloud-enabled Wireless Body Area Networks (WBANs)	Wan et al. have proposed using the hybrid of MCC and WBAN architectures to provide more reliable QoS to healthcare applications	Wan et al. (2013)
Cognitive wireless networks		
Coexistence of 802.15.4 with IEEE 802.11	Proposes distributed adaptation strategies to ensure coexistence of IEEE 802.11 WLAN and IEEE 802.15.4 wireless sensor networks (WSNs) in the ISM band	Pollin et al. (2006)
Coexistence between IEEE 802.11b and IEEE 802.16a networks	Proposes algorithms—based on dynamic frequency selection (DFS), power control (PC) and time-agility (TA)—to allow IEEE 802.11b and IEEE 802.16a networks to coexist in the same unlicensed band	Jing et al. (2005)
QoS-aware MAC for IEEE 802.11p	Proposes an efficient multichannel QoS cognitive MAC (MQOG) for cognitive vehicular networks	Ajaltouni et al. (2012)
Integration of IEEE 802.11 and 3G	Proposes schemes for integrating IEEE 802.11 and 3G seamlessly while satisfying QoS guarantees and roaming agreements	Buddhikot et al. (2003)
Managing TCP in DSA-based WLANs	Proposes a framework known as DSASync for improved end-to-end TCP performance in dynamic spectrum access (DSA) wireless networks	Kumar and Shin (2010)
Handover between IEEE 802.11b and overlay networks	Proposes algorithms for intersystem handover between IEEE 802.11b and an overlay network while satisfying the QoS parameters of minimum data rate, maximum data block delay, and maximum BER	Garmonov et al. (2008)

Ferguson et al. (2013) have proposed PANE that implements an API for participatory networking using SDN principles. PANE can be used to allow a user to reserve a certain guaranteed minimum bandwidth between two hosts. The flexibility afforded by PANE can be used for enforcing QoS amongst other tasks (such as access control and path configuration).

Numerous QoS works exploiting SDN principles and/or the OpenFlow protocol have been proposed. OpenFlow is the most prominent SDN protocol used for implementing the architectural vision of separated control and data planes. OpenFlow implements a protocol used by the SDN controller to communicate with controlled devices. OpenFlow provides basic QoS primitives including support of only FIFO scheduling which may be insufficient for some applications like multimedia streaming. QoS support for OpenFlow is improving: queues are available in OpenFlow 1.0 which enables traffic shaping, while in the latest version,
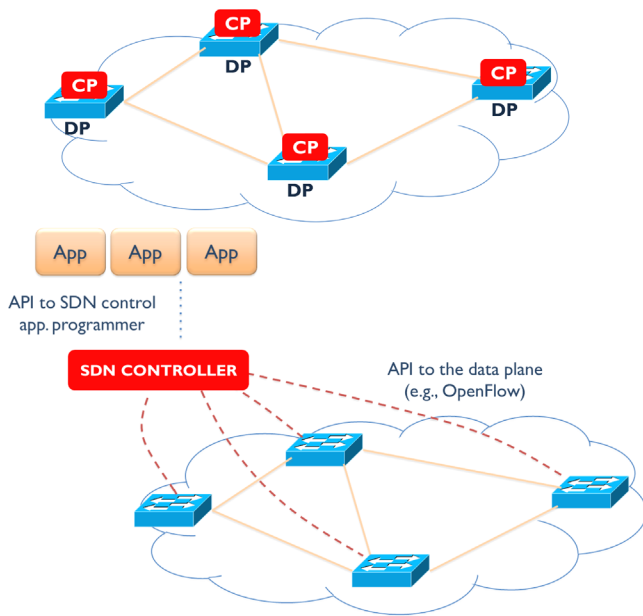
Fig. 3. Comparison of a traditional network and a SDN network architecture (Qadir et al., 2014).

namely OpenFlow 1.3, rate limiting can be supported through meter tables. The QoS routing significantly affects the overall throughput of a network. The major portion of Internet traffic is composed of best-effort traffic. Therefore, an efficient QoS routing algorithm must take into account the existence of best-effort traffic and its impact on the overall performance. The OpenFlow controllers can generate flow tables to manage the QoS within a network using different routing protocols (Civanlar et al., 2010). The controller performs additional functions of QoS contract management and QoS route management. The controller also polls switches to report on the congestion level in the network.

Kim et al. (2010) proposed an automated fine-grained QoS control approach for managing OpenFlow networks. In particular, the proposed OpenFlow-based network QoS control framework utilized per-flow rate limiters and dynamic priority assignment. Another QoS solution called *QoSFlow* for OpenFlow-based SDN networks was proposed by Ishimori et al. (2013). The proposed QoS solution relies on appropriate control of packet scheduling. In particular, QoSFlow is proposed as a QoS development strategy that relies on multiple packet schedulers for OpenFlow supported networks to overcome the limitations of FIFO packet scheduling. QoSFlow can provide control for the following packet schedulers: hierarchical token bucket (HTB), random early detection (RED) and stochastic fair queueing (SFQ). *OpenQoS* (Egilmez et al., 2012) is an OpenFlow controller designed for supporting multimedia flows with end-to-end QoS requirements. It enables QoS by placing multimedia traffic on QoS-guaranteed routes. OpenQoS presents a new dynamic QoS routing scheme that maintains the shortest path for the data delivery, which helps in minimizing packet loss and latency. The results show that the network turbulence has a minimal effect on video quality with QoS support. On the other hand, the videos without QoS support suffer significantly from quality degradation.

In another OpenFlow-based work, Sharma et al. (2014) have presented a QoS solution for the future Internet using SDN technologies. The SDN controller is used to configure the queues on the OpenFlow enabled routers. It uses the Type of Service (ToS) field to enqueue packets in the appropriate queues at the edge routers. The routers identify QoS requirement of the traffics and send them to appropriate queues accordingly. The model is tested

for both the single and multiple AS environments to test the resilience in face of network failures. In another work (Tomovic et al., 2014), an SDN-based controller that uses the OpenFlow protocol is proposed to provide the QoS guarantees to high priority flows. The solution minimizes the degradation of best-effort traffic by using a new algorithm, which is based on resource utilization. The paper also shows the improvements brought about by the proposed algorithm on the traditional best-effort and IntServ QoS schemes through experimental evaluation. In yet another OpenFlow-based QoS work, *OpenQFlow* (Nam-Seok et al., 2013)— a two tiered flow-based QoS framework was proposed. In OpenQ-Flow, the flow table is divided further into three tables: the first for the flow state, the second for the forwarding rules, and the final for defining the QoS rules in which multiple micro-flows could share a single QoS profile.

There exists work in leveraging SDN principles for improved QoS performance in wireless networks. SDN-enabled IEEE 802.11 networks provide the benefits of agile QoS provisioning. In an SDN network, bandwidth allocation, rate limiting, and traffic shaping can be efficiently implemented at the network level through automated QoS network APIs via the network controller. Real-time measurements can help to steer policies that can efficiently enforce the QoS mechanisms in real time (Amani et al., 2014). QoS policies can be deployed at the controller and the wastage of resources can be minimized. In an SDN-enabled network, several controllers in a single area can communicate with each other to exchange network information. This communication allows the users to connect with APs, regardless of their operators, thus improving the user's Quality of Experience (QoE). An AP that receives packets with a destination address of another network can forward the packets flexibly through rules defined at the controller (Chaudet and Haddad, 2013). Zhao et al. (2014) have proposed a framework for a single SDN controller controlling all the APs via OpenFlow interface. The proposed framework utilizes OpenFlow to mitigate interference in enterprise WLANs. The proposed framework adds specific rules in various APs for packet scheduling without modifying the conventional DCF mechanism. In another work, Lee et al. (2014) extend the SDN controller network to mobile devices. It helps in achieving real-time detection of QoS demands in a network and can provide end-to-end QoS control.

Kassler et al. (2012) have proposed an SDN-based QoE-driven multimedia service negotiation and path optimization architecture. In the proposed architecture, the QoS measurements are taken at the service level and the network level. At the network level, the data flows from a source to a destination along the same path. This helps the media to be delivered in the best possible path and service configuration. In return, the overall QoS for the users is improved. SDNs are proposed to combine the configurations of network elements and end hosts. This enables network operators to define their own set of rules to control the traffic routing and QoS. A dynamic framework for ensuring QoS in streaming videos at the control plane is presented in Egilmez et al. (2013). The scheme works in the OpenFlow-based networks by optimizing forwarding decisions at the SDN controller. The controller acts as the brain of a network where the forwarding decisions are made. The routing choices are associated with the priority of the data flows. The resources are reserved at the controller, based on the type of delivery the controller can provide (Sezer et al., 2013). This reservation scheme does not affect the other types of flows due to the dynamic routing mechanism in the OpenFlow architecture.

There also has been work on using SDN for QoS management of home networks. Seddiki et al. (2014) have proposed an SDN-based QoS framework called *FlowQoS* for managing the QoS in home broadband networks. Using an SDN based framework is helpful for managing home networks since home routers are typically feature

constrained due to limited computational capabilities. In addition, since users cannot be assumed to be networking experts, the QoS configuration is also often too complicated for end-users. FlowQoS manages this by outsourcing the QoS functions to a separate centralized control logic that allows a user to specify desired QoS policies more abstractly. More specifically, the FlowQoS framework comprises two components: firstly, a flow classifier mapping application traffic to different parts of flow spaces, secondly, a SDN-based rate shaper.

SDN has also made it possible to bring autonomic capabilities in the network nodes. Such an approach is presented in Wang et al. (2014). This approach proposes Packet Context-aware QoS model (PCaQoS) that configures various QoS features in OpenFlow switches by extending the OpenFlow configuration protocol. This model checks the QoS requirements of the packets and then forwards them into specific queues.

8.2. Cloud-based Wireless Networks (CbWNs)

The main idea of cloud computing is to offer computing services (e.g., networks, storage, applications)—provisioned through APIs via the web from a shared pool of resources—in a virtualized data center in utility computing style (Armbrust et al.,
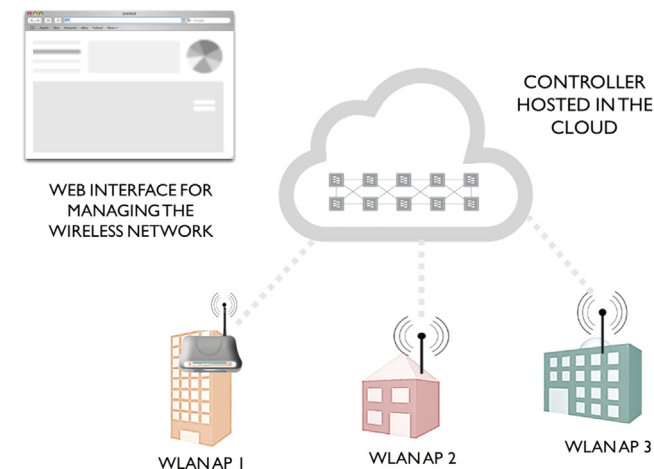

Fig. 4. Cloud-based Wireless Networks (CbWNs) (Qadir et al., 2014).

2010). The capability to program a network through the cloud/SDN architectures allow revolutionary dynamism in service provisioning, network management and control. The trend of CbWNs is to extend the cloud computing concept to wireless networks (Qadir et al., 2014). An illustration of the CbWN architecture is shown in Fig. 4. Some salient advantages of CbWNs are described next.

- *Centralized management*: A major application of CbWNs is *centralized remote management* of wireless networks. The control and provisioning of wireless access points (CAPWAP) protocol—defined by IETF in RFC 5415 (Calhoun, 2009)—is a control and management (C&M) protocol that aims at migrating functionalities from the hardware AP equipment to an external controller potentially managed via the cloud. There exists significant interest in the research community in proposing efficient approaches for central management of Wi-Fi networks (Dalvi et al., 2011). Various industrial solutions, such as Meraki Networks (2014) from Cisco, Aruba Networks (2014), and AeroHive, have also been proposed to perform cloud-based management of WLANs.
- *Zero-touch auto-configuration*: The centralized management paradigm of CbWNs can allow plug and play *zero-touch*

auto-configuration of wireless APs allowing the APs to function without any manual configuration by the network administrator. The centralized management of wireless networks also allow cloud-based performance management and the use of *advanced data analytics* for performance optimization including *real-time reconfiguration* of wireless parameters.

8.2.1. QoS efforts for IEEE 802.11-based CbWNs

Most of the devices used for cloud computing applications are portable and are connected through IEEE 802.11 WLAN. The WiFi access network may not be able to cope with the need of providing services to QoS-assured cloud multimedia applications. In particular, when the devices are in motion, QoS-aware handover should take into account the traffic load and currently accessible bandwidth at each location and at particular AP with high precision. Consequently, a precise modeling of the genuine surroundings of IEEE 802.11 WLAN is important for proficient QoS-aware cloud service provisioning. Tursunova and Kim (2012) proposed a pragmatic IEEE 802.11e EDCA model for QoS-aware differentiated multimedia cloud service provisioning in WLAN networks.

Most of the previous work concentrated on the analysis of EDCA in saturated and non-saturated states. In Engelstad and Osterbo (2005), Malone et al. (2007), Giustiniano et al. (2010), Wu et al. (2006), and Kosek-Szott et al. (2011), the authors examine only clean channel surroundings, with an assumption that the frame error is caused by packet collision only, while in real surroundings, frame errors may take place due to channel noise. In Tursunova and Kim (2012), the authors considered the time-varying frame error probability of independent stations. They strengthened the mathematical model of IEEE 802.11e Kosek-Szott et al. (2011) by using the calculated packet error probability, which may be caused by frame collisions and channel noise. Experimental outcomes show that the suggested model gives more precise assessment, compared to existing analytical models.

WLANs deployed by large firms or universities can compose of hundreds or even thousands of APs. Similarly the size of the operating system of an AP also increases with time due to the inclusion of software packages in each release. Therefore, it is getting harder for the network administrators to configure each AP individually. Reducing complexity of networking appliances and uncovering data flow management tasks via standardized interfaces and high-level programming primitives are some of the main concepts of SDN. In Sharkh et al. (2013), the authors present a resource allocation mechanism based on the cloud environments, as well as an energy-aware model for the data centers. To obtain similar advantages in WLANs, Dely et al. (2012) introduced CloudMAC, which is a novel management architecture in which access points redirect MAC frames only. The remaining functionalities, like the processing of MAC data or management frames, are executed in typical servers that are operated in data centers and can be allocated via cloud computing infrastructure. OpenFlow is used to organize the flow and transmission characteristics of MAC frames.

An emerging trend in cloud computing is the 'mobile cloud computing' (MCC) architecture that integrates mobile environments with cloud computing to enable mobile wireless clients to perform computations in the cloud. Various QoS solutions have been proposed for the MCC architecture. Chun et al. (2011) proposed CloneCloud which allows unmodified mobile applications running in an application-level Virtual Memory (VM) to seamlessly offload part of their execution from the mobile devices onto device clones operating in the cloud. The ability to offload computation can be exploited in a QoS framework to meet stringent deadlines. Wan et al. (2013) have proposed using cloud-enabled wireless body area
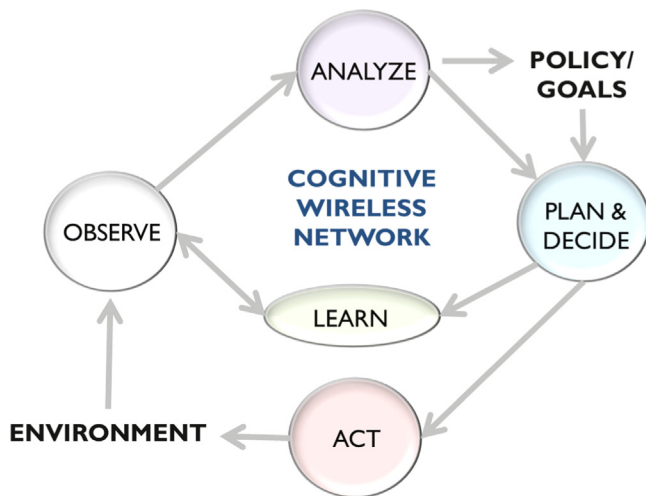
Fig. 5. Cognitive wireless networks (CWN) include an embedded cognitive engine which can observe network conditions, orient itself with the context, learn from experience, and decide to act (Qadir et al., 2014).

networks (WBANs) for healthcare applications. WBANs—which are characterized by limited energy, memory, computation, and communication capabilities—can utilize the mobile cloud computing (MCC) architecture to remotely manage its processing requirements and provision QoS more reliably. In addition, Wan et al. proposed that semantic models and ontology technology can be used to provide generic standard protocols through which QoS may be assured. In another work, Lai et al. (2013) have presented a network and device aware self-adaptive QoS solution for cloud-based multimedia streaming suitable for mobile devices. The proposed multimedia cloud computing solution considers the overall network environment and adjusts the transmission frequency and multimedia transcoding parameters accordingly.

8.3. Cognitive wireless networks (CWN)

Cognitive wireless networks (CWNs) are next-generation wireless networks—that demonstrate network-wide intelligent behavior—in which network nodes are incorporated with cognitive engines (see Fig. 5) consist of substantial artificial intelligence (AI) approaches in the form of machine learning, knowledge reasoning, optimization, and natural language processing (Akyildiz et al., 2006; Haykin, 2005). Such networks are composed of network nodes equipped with cognitive radios (CR), which display device-level intelligent behavior.

Along with network-level reconfiguration capabilities afforded by technologies such as SDN which can be used to realize programmable data plane and programmable control plane, future wireless programmable networking will also implement some variant of a "*knowledge plane*" (Clark et al., 2003). Traditionally, a network of nodes equipped with CRs is called a cognitive radio network (CRN) with the dominant application of CR technology being dynamic spectrum access (DSA), which can resolve the 'artificial spectrum scarcity' problem resulting from the classic command-and-control licensing approach (Fette, 2009) adopted in various countries around the world. Since CRNs inherently embody AI techniques with wireless communications, it seems natural to explore using CRs to provide mechanisms for implementing the knowledge plane of future programmable wireless devices.

While the bulk of CRN work has focused on enabling device-level intelligent behavior, the concept of CWNs, initially proposed in Thomas et al. (2007), generalizes CRNs and emphasizes *network-level* intelligence and self-aware behavior. While DSA is

the most popularly cited application of CRNs, developing network-level intelligence in CRNs enables numerous other applications—including the ability to reprogram itself optimally according to network conditions.

In previous CRN research, it has been observed that PHY and MAC layers offer many "*knobs*" that can be tweaked to optimize performance which can be measured through some "*meters*". In Fette (2009), many examples of knobs and meters at the PHY and MAC layers have been provided. Since CRNs operate in dynamic, often unknown, conditions, configuring the knobs optimally is not a trivial problem. Various AI-based techniques have been proposed in the literature to assist CRNs in their quest of performing autonomous optimal adaptations in such settings. Apart from AI techniques, CRN also borrows techniques and tools from various other fields such as game theory, control theory, optimization theory, and metaheuristics (Haykin, 2005).

Game theory has been used in various work to model, analyze, and develop QoS solutions for CRNs. Berlemann et al. (2005) have proposed the use of radio resource sharing games to enable distributed QoS solutions in unlicensed bands shared by multiple users. Attar et al. proposed a game-theoretic resource allocation framework that guarantees QoS in a DSA environment (in which the primary network is assumed to be OFDM-based cellular network). The QoS is defined by the minimum rate available to the primary network and the target BER.

Optimization theory has also been used in the literature to address the problem of QoS-constrained dynamic spectrum access. For example, Xing et al. (2007) considered QoS differentiation for various unlicensed users while incorporating interference temperature constraints assuming a spectrum underlay access (alternatively, known as a shared-use model). In addition, various cross-layered solutions have been developed for ensuring QoS in CRNs. For example, Su and Zhang (2008) proposed a distributed cross-layered solution incorporating spectrum sensing at the PHY layer and packet scheduling at the MAC layer for QoS provisioning in CRNs.

There has also been work in using cognitive technologies to facilitate QoS-aware coexistence among multiple 802.11 WLANs, between 802.11 and 802.16 networks (Berlemann et al., 2006; Jing et al., 2005), and between 802.11 WLANs and overlay networks (Garmonov et al., 2008).

9. Open research issues and future work

QoS enhancement schemes in modern wireless-based networks still need further attention. This section highlights some of the important issues.

9.1. Convergence of different technologies

Wireless technologies are proliferating at a breakneck pace, and in such a dynamic ecosystem, technologies that facilitate multi-technology convergence is becoming increasingly important. In the future, IEEE 802.11-based networks will increasingly coexist with other wireless technologies such as 4G/5G, WiMAX, RFID, Internet of Things (IoT), and Wireless Sensor Networks. As an example, 5G—expected to materialize by 2020—will be highly integrative and will tie Wi-Fi with other wireless mobile standards such as 3G and LTE (Andrews et al., 2015). IEEE 802.11-based wireless networks is facing stiff competition from other technologies that coexist in the unlicensed spectrum space, such as the IEEE 802.16-based wireless metropolitan area networks. Management of handover, spectrum sharing, coexistence, and interworking of diverse technologies, become important for ensuring QoS. There will be a lot of interest in QoS-aware spectrum sharing and

coexistence between IEEE 802.11-based wireless networks and other technologies. This entails work at various layers including the potential use of multi-path TCP at the transport layer for improving QoS by exploiting multi-homing with a diverse range of wireless networks, such as Wi-Fi and 3G. More research needs to be done to ensure QoS in such environments where heterogeneous technologies exist.

9.2. Context-aware and cognitive/semantic computing

To improve the QoS and QoE, it is imperative for researchers to seamlessly incorporate user preferences, and awareness of context, which can be based on *identity*, *location*, *time*, or *activity*, into IEEE-based networks. Since the end user's traffic varies with time, traffic behavior should be analyzed to predict the future traffic patterns and subsequently to adopt appropriate strategies. This helps in fulfilling the requirement of end devices with higher efficiency. Also, since the wireless networks are mobile, predicting the future locations of nodes helps in data forwarding, and thus in reducing the overall delay. If the future location of a node can be predicted from its mobility pattern and its speed, this helps in successful delivery of packets. Capturing the mobility patterns and its behavior ensure enhanced QoS.

Similar to the concept of cognitive networking, there is a rich literature of semantic computing and context-aware computing that can be used for developing QoS-aware solutions (Wang et al., 2006). While traditionally focused on web services (Kritikos and Plexousakis, 2007), and ubiquitous and pervasive computing environs (Mokhtar et al., 2008), these techniques also are a good match for emerging architectures such as IoT, cloud computing, and global-scale machine-to-machine communication. QoS solutions in such environments can be enhanced by semantic support that leverages available contextual information. Recent work have proposed using context-aware semantic service discovery and dissemination methods to meet QoS requirements in such networks.

Although some work has been done in the design of QoS-aware cognitive routing (Qadir, 2013; How et al., 2011) and cognitive transport layer protocols (Kumar and Shin, 2010; Chowdhury et al., 2009) for CWNs, more work needs to be done to realize the considerable promise of using AI- and machine learning-based techniques for developing IEEE 802.11-based wireless networks that can provide high QoS and QoE.

9.3. Challenges due to virtualization

Virtualization has transformed both operational efficiency and the economics of the computing industry, and more recently, the data center environment. With the growing role of virtualization in networking, it is highly likely that IEEE 802.11-based networks will increasingly utilize virtualization technology in the future (Qadir et al., 2014). In particular, the combination of cloud computing and network virtualization (including network functions virtualization) allow programmability that leads to unprecedented flexibility in rapidly creating, deploying, and managing novel services in virtualized settings as per the demands of users. This can create a new service-oriented architecture for wireless networking where heterogeneous wireless access technologies including IEEE 802.11 may coexist and converge as *extended cloud infrastructure* (Wen et al., 2013). The QoS challenges associated with virtualization/cloud management of IEEE 802.11 WLANs need to be investigated thoroughly to resolve potential issues.

9.4. Exploiting online social networking information for QoS

An interesting direction of future work is to explore using information and properties of online social networks (OSN) for QoS provisioning. OSN information can be used in the context of QoS to predict and anticipate demand for resources as well as provision of network resources. OSN information can also be used to develop policies that translate OSN relationships into traffic engineering decisions—e.g., in community shared networks, OSN information can be used to provide different QoS levels to different users based on their social relationship (so that 'friends' get prioritized service over non-friends). In a recent work, Sathiaseelan have proposed using OSN credentials for resource provisioning using SDN principles. With the increasing role of OSNs—such as Facebook, Twitter—in modern networks, the information in OSN provide rich metadata that can be used profitably for efficient QoS implementations.

9.5. Fog computing and QoS

The propensity of network architects towards centralization and decentralization has seesawed over the past many years (cf. the change in focus from the mainframes to PCs to the cloud again). Fog computing is the latest architecture sustaining this trend. Fog computing is an emerging paradigm that extends the cloud computing concept by bringing resources and services closer to the network's edge (Bonomi et al., 2012). Fog computing advances cloud computing into the internet-of-things (IoT) era by descending the "cloud" and diffusing it to the end-user client devices. With the IoT architecture envisioning the connectivity of billions of devices, the proposition of using end devices for hosting services has become very attractive.

This migration of computation, control and storage from the cloud to the client devices has become possible due to the advances in client device technology which is much more powerful than ever before. This allows applications to run close to the data, which is the key characteristic of fog computing that distinguishes it from cloud computing. Such an architecture is driven by increasingly stringent requirements of modern multimedia applications. Satisfying these requirements with QoS will require augmentation of the traditional cloud computing model with the fog computing model as well as end user caching using concepts like information-centric networking (ICN) (Nef et al., 2012; Duan et al., 2011). With fog computing implemented at the edge of the network, it provides an excellent framework for enabling low-latency QoS-aware solutions (Stojmenovic and Wen, 2014).

10. Conclusion

It is anticipated that the bulk of access technologies in the near future will be wireless. With emerging standards such as 5G, and the convergence of the telecom and Internet industries on IP-based technologies, the ability to provide high QoS has become paramount. In this paper, we have surveyed various QoS enhancement techniques proposed for IEEE 802.11-based wireless networks with our discussion encompassing both classical techniques as well as proposals for new and emerging architectures such as SDN and cloud-networks. We have classified these techniques using different criteria. We have discussed QoS solutions that have been proposed for various TCP/IP layers along with a discussion on cross-layered protocols. We have also highlighted open research challenges and directions for future work.

References

Aad I, Castelluccia C. Introducing service differentiation into IEEE 802.11. In: Proceedings of fifth IEEE symposium on computers and communications, 2000. ISCC 2000. NY, USA: IEEE; 2000. p. 438–43.

Aad I, Castelluccia C. Differentiation mechanisms for IEEE 802.11. In: INFOCOM 2001. The 20th annual joint conference of the IEEE computer and communications societies. Proceedings of IEEE, vol. 1. IEEE; 2001. p. 209–18.

Abd El-atty S. Efficient packet scheduling with pre-defined QoS using cross-layer technique in wireless networks. In: Proceedings of the 11th IEEE symposium on computers and communications, 2006. ISCC'06. NY, USA: IEEE; 2006. p. 820–6.

Abdrabou A, Zhuang W. A position-based QoS routing scheme for uwb mobile ad hoc networks. IEEE J Sel Areas Commun 2006;24(4):850–6.

Aboul-Magd O. Wireless local area networks quality of service: an engineering perspective. NY, USA: IEEE; 2009.

Achary R, Chellaih PR, Vaityanathan V, Nagarajan S. Enhanced QoS by service differentiation in MAC-layer for WLAN. Int J Comput Appl 2012;55.

Agarwal A, Jagannatham AK. Optimal adaptive modulation for QoS constrained wireless networks with renewable energy sources. IEEE Wirel Commun Lett 2013;2(1):78–81.

Aikawa S, Motoyama Y, Umehira M. Forward error correction schemes for wireless ATM systems. In: 1996 IEEE international conference on communications, 1996. ICC'96, conference record, converging technologies for tomorrow's applications, vol. 1. NY, USA: IEEE; 1996. p. 454–8.

Ajaltouni H, Pazzi RW, Boukerche A. An efficient QoS MAC for IEEE 802.11p over cognitive multichannel vehicular networks. In: 2012 IEEE international conference on Communications (ICC). IEEE; 2012. p. 413–7.

Akyildiz I, Lee W, Vuran M, Mohanty S. Next generation/dynamic spectrum access/cognitive radio wireless networks: a survey. Comput Netw 2006;50(13):2127–2159.

Amani M, Mahmoodi T, Tatipamula M, Aghvami H. SDN-based data offloading for 5G mobile networks. ZTE Commun Issue 2, 2014. http://wwwen.zte.com.cn/endata/magazine/ztecommunications/2014/2/articles/201406/t20140618_425075.html.

Andreadis A, Zambon R. Techniques for preserving QoS performance in contention-based IEEE 802.11e networks. In: Advanced wireless LAN. Rijeka, Croatia: Intech Open; 2012.

Andrews JG, Buzzi S, Choi W, Hanly S, Lozano A, Soong AC, et al. What will 5G be? IEEE JSAC, in press; arXiv preprint arXiv:1405.2957; 2015.

Ansel P, Ni Q, Turletti T. An efficient scheduling scheme for IEEE 802.11e. In: Proceedings of modeling and optimization in mobile, ad hoc and wireless networks; 2004. p. 24–6.

Armbrust M, Fox A, Griffith R, Joseph AD, Katz R, Konwinski A, et al. A view of cloud computing. Commun ACM 2010;53(April (4)):50–8 [Online]. Available: http://dx.doi.acm.org/10.1145/1721654.1721672.

Aruba Networks [Online]. ⟨http://cloud.arubanetworks.com/⟩; 2014 [accessed: 2014-06-1].

Aurrecoechea C, Campbell AT, Hauw L. A survey of QoS architectures. Multimed Syst 1998;6(3):138–51.

Baghaei N, Hunt R. Review of quality of service performance in wireless LANs and 3 multimedia application services. Comput Commun 2004;27(17):1684–92.

Balakrishnan H, Padmanabhan VN, Seshan S, Katz RH. A comparison of mechanisms for improving TCP performance over wireless links. IEEE/ACM Trans Netw 1997;5(6):756–69.

Banchs A, Perez X. Distributed weighted fair queuing in 802.11 wireless LAN. In: IEEE international conference on communications, 2002. ICC 2002, vol. 5. NY, USA: IEEE; 2002. p. 3121–7.

Barolli L, Koyama A, Shiratori N. A QoS routing method for ad hoc networks based on genetic algorithm. In: Proceedings of the 14th international workshop on database and expert systems applications, 2003. IEEE; 2003. p. 175–9.

Bashandy AR, Chong EK, Ghafoor A. Generalized quality-of-service routing with resource allocation. IEEE J Sel Areas Commun 2005;23(2):450–63.

Bensaou B, Kong Z-N, Tsang DH. A measurement-assisted, model-based admission control algorithm for IEEE 802.11e. J Interconnect Netw 2009;10(04):303–20.

Berlemann L, Hiertz GR, Walke BH, Mangold S. Radio resource sharing games: enabling QoS support in unlicensed bands. IEEE Netw 2005;19(4):59–65.

Berlemann L, Hoymann C, Hiertz GR, Mangold S. Coexistence and interworking of IEEE 802.16 and IEEE 802.11(e). In: IEEE 63rd vehicular technology conference, 2006. VTC 2006-Spring, vol. 1. IEEE; 2006. p. 27–31.

Blake S, Black D, Carlson M, Davies E, Wang Z, Weiss W. An architecture for differentiated services, IETF RFC 2475; 1998. https://tools.ietf.org/html/rfc2475.

Bolt J-C, Turletti T. Experience with control mechanisms for packet video in the internet. ACM SIGCOMM Comput Commun Rev 1998;28(1):4–15.

Bonomi F, Milito R, Zhu J, Addepalli S. Fog computing and its role in the internet of things. In: Proceedings of the first edition of the MCC workshop on mobile cloud computing. NY, USA: ACM; 2012. p. 13–6.

Boutremans C, Le Boudec J-Y. Adaptive joint playout buffer and FEC adjustment for internet telephony. In: INFOCOM 2003. The IEEE 22nd annual joint conference of the IEEE computer and communications, vol. 1. NY, USA: ACM; 2003. p. 652–62.

Braden R, Clark D, Shenker S. Integrated services in the internet architecture: an overview. RFC 1633; 1994.

Brewer OT, Ayyagari A. Comparison and analysis of measurement and parameter based admission control methods for quality of service (QoS) provisioning. In: Military communications conference, 2010-Milcom 2010. NY, USA: IEEE; 2010. p. 184–8.

Bucciol P, Davini G, Masala E, Filippi E, De Martin JC. Cross-layer perceptual ARQ for h.264 video streaming over 802.11 wireless networks. In: IEEE global telecommunications conference, 2004. GLOBECOM'04, vol. 5. NY, USA: IEEE; 2004. p. 3027–31.

Buddhikot M, Chandranmenon G, Han S, Lee Y-W, Miller S, Salgarelli L, 2003. Integration of 802.11 and third-generation wireless data networks. In: INFOCOM 2003. The IEEE 22nd annual joint conference of the IEEE computer and communications, vol. 1. NY, USA: IEEE; 2003. p. 503–12.

Calhoun P. Rfc 5415. Control and provisioning of wireless access points (CAPWAP) protocol specifications, https://tools.ietf.org/rfc/rfc5415.txt; 2009.

Calhoun P. Lightweight access point protocol. ⟨http://tools.ietf.org/html/rfc5412⟩; 2010 [accessed: 2013-09-12].

Cano C, Bellalta B, Oliver M. Adaptive admission control mechanism for IEEE 802.11e WLANs. In: IEEE 18th international symposium on personal, indoor and mobile radio communications, 2007. PIMRC 2007. NY, USA: IEEE; 2007. p. 1–5.

Cen S, Cosman PC, Voelker GM. End-to-end differentiation of congestion and wireless losses. IEEE/ACM Trans Netw 2003;11(5):703–17.

Cervello G, Choi S. Collision avoidance in IEEE 802.11 contention free period (cfp) with overlapping basic service sets (BSSs). US Patent 7,054,329; May 30, 2006.

Chakareski J, Chou PA. Application layer error-correction coding for rate-distortion optimized streaming to wireless clients. IEEE Trans Commun 2004;52(10):1675–87.

Chakareski J, Girod B. Rate-distortion optimized packet scheduling and routing for media streaming with path diversity. In: Proceedings of data compression conference, 2003. DCC 2003. NY, USA: IEEE; 2003. p. 203–12.

Chalmers D, Sloman M. A survey of quality of service in mobile computing environments. IEEE Commun Surv Tutor 1999;2(2):2–10.

Chaouchi H, Munaretto A. Adaptive QoS management for IEEE 802.11 future wireless ISPs. Wirel Netw 2004;10(4):413–21.

Charfi E, Chaari L, Kamoun L. PHY/MAC enhancements and QoS mechanisms for very high throughput WLANs: a survey. IEEE Commun Surv Tutor 2013;15(4):1714–35.

Chaudet C, Haddad Y. Wireless software defined networks: challenges and opportunities. In: 2013 IEEE international conference on microwaves, communications, antennas and electronics systems (COMCAS). NY, USA: IEEE; 2013. p. 1–5.

Chen D, Varshney PK. QoS support in wireless sensor networks: a survey. In: International conference on wireless networks, vol. 233; 2004. p. 1–7.

Chen L, Heinzelman WB. QoS—aware routing based on bandwidth estimation for mobile ad hoc networks. IEEE J Sel Areas Commun 2005;23(3):561–72.

Chen S, Nahrstedt K. Distributed quality-of-service routing in ad hoc networks. IEEE J Sel Areas Commun 1999;17(8):1488–505.

Chen W-P, Hsiao Y-C, Hou JC, Ge Y, Fitz MP. Syndrome: a light-weight approach to improving TCP performance in mobile wireless networks. Wirel Commun Mobile Comput 2002;2(1):37–57.

Chen Y-M, Chung H, Lee E, Tong Z. Interleaved weighted fair queuing mechanism and system. US Patent 6,975,638; December 13, 2005.

Chen X, Zhai H, Tian X, Fang Y. Supporting QoS in IEEE 802.11e wireless LANs. IEEE Trans Wirel Commun 2006;5(8):2217–27.

Chen Y-S, Chuang M-Y, Tseng F-C, Ke C-H. High performance distributed coordination function with QoS support in IEEE 802.11e networks. In: Australasian telecommunication networks and applications conference (ATNAC), 2011. NY, USA: IEEE; 2011. p. 1–6.

Choi S, Del Prado J, Sai Shankar N, Mangold S. IEEE 802.11e contention-based channel access (edcf) performance evaluation. In: IEEE international conference on communications, 2003. ICC'03, vol. 2. NY, USA: IEEE; 2003. p. 1151–6.

Choi S, Choi Y, Lee I. IEEE 802.11 MAC-level FEC scheme with retransmission combining. IEEE Trans Wirel Commun 2006;5(1):203–11.

Chou CT, Misra A, Qadir J. Low-latency broadcast in multirate wireless mesh networks. IEEE J Sel Areas Commun 2006;24(11):2081–91.

Chowdhury KR, Di Felice M, Akyildiz IF. Tp-crahn: a transport protocol for cognitive radio ad-hoc networks. In: IEEE INFOCOM 2009. NY, USA: IEEE; 2009. p. 2482–90.

Christin N, Liebeherr J. A QoS architecture for quantitative service differentiation. IEEE Commun Mag 2003;41(6):38–45.

Chun B-G, Ihm S, Maniatis P, Naik M, Patti A. Clonecloud: elastic execution between mobile device and cloud. In: Proceedings of the sixth conference on computer systems. NY, USA: ACM; 2011. p. 301–14.

Civanlar S, Parlakisik M, Tekalp AM, Gorkemli B, Kaytaz B, Onem E. A QoS -enabled openflow environment for scalable video streaming. In: 2010 IEEE GLOBECOM workshops (GC Wkshps). NY, USA: IEEE; 2010. p. 351–6.

Clark DD, Shenker S, Zhang, L. Supporting real-time applications in an integrated services packet network: architecture and mechanism, vol. 22, no. 4. NY, USA: ACM; 1992.

Clark DD, Partridge C, Ramming JC, Wroclawski JT. A knowledge plane for the internet. In: Proceedings of the 2003 conference on applications, technologies, architectures, and protocols for computer communications. NY, USA: ACM; 2003. p. 3–10.

Crow BP, Widjaja I, Kim JG, Sakai PT. Ieee 802.11 wireless local area networks. IEEE Commun Mag 1997;35(9):116–26.

Cruz RL, Santhanam AV. Optimal routing, link scheduling and power control in multihop wireless networks. In: INFOCOM 2003. IEEE the 22nd annual joint conference of the IEEE computer and communications, vol. 1. NY, USA: IEEE; 2003. p. 702–11.

Dalvi A, Swamy P, Meshram B. Centralized management approach for WLAN. In: Computer networks and information technologies. NY, USA: Springer; 2011. p. 578–80.

Dely P, Vestin J, Kassler A, Bayer N, Einsiedler H, Peylo C. CloudMAC: an OpenFlow based architecture for 802.11 MAC layer processing in the cloud. In: 2012 IEEE Globecom workshops (GC Wkshps). NY, USA: IEEE; 2012. p. 186–91.

Ding W. Joint encoder and channel rate control of VBR video over ATM networks. IEEE Trans Circuits Syst Video Technol 1997;7(2):266–78.

Domingo MC, Remondo D. An interaction model between ad hoc networks and fixed IP networks for QoS support. In: Proceedings of the seventh ACM international symposium on modeling, analysis and simulation of wireless and mobile systems. NY, USA: ACM; 2004. p. 188–94.

Drabu Y. A survey of QoS techniques in 802.11. Internet'Online; 1999. p. 0001–03.

Duan R, Chen X, Xing T. A QoS architecture for IOT. In: 2011 international conference on and fourth international conference on cyber, physical and social computing internet of things (iThings/CPSCom). NY, USA: IEEE; 2011. p. 717–20.

Dujovne D, Turletti T, Filali F. A taxonomy of IEEE 802.11 wireless parameters and open source measurement tools. IEEE Commun Surv Tutor 2010;12(2):249–62.

Egilmez H, Tekalp M. Distributed QoS architectures for multimedia streaming over software defined networks. IEEE Trans Multimed 2014;16(October (6)): 1597–1609.

Egilmez HE, Dane ST, Bagci KT, Tekalp AM. In: 2012 Asia-Pacific signal and information processing association annual summit and conference (APSIPA ASC). IEEE; 2012. p. 1–8.

Egilmez HE, Civanlar S, Tekalp AM. An optimization framework for QoS-enabled adaptive video streaming over openflow networks. IEEE Trans Multimed 2013;15(3):710–5.

Elsayed KM, Khattab AK. Channel-aware earliest deadline due fair scheduling for wireless multimedia networks. Wirel Pers Commun 2006;38(2):233–52.

Engelstad PE, Osterbo ON. Non-saturation and saturation analysis of IEEE 802.11e EDCA with starvation prediction. In: Proceedings of the eighth ACM international symposium on modeling, analysis and simulation of wireless and mobile systems. NY, USA: ACM; 2005. p. 224–33.

Farber N, Stuhlmuller K, Girod B. Analysis of error propagation in hybrid video coding with application to error resilience. In: 1999 Proceedings of international conference on image processing, 1999. ICIP 99, vol. 2. NY, USA: IEEE; 1999. p. 550–4.

Ferguson AD, Guha A, Liang C, Fonseca R, Krishnamurthi S. Participatory networking: an API for application control of sdns. In: ACM SIGCOMM computer communication review, vol. 43, no. 4. ACM; 2013. p. 327–38.

Fette BA. Cognitive radio technology. Academic Press; 2 edition; ISBN-10: 0123745357; ISBN-13: 978-0123745354; 2009.

Gao D, Cai J, Ngan KN. Admission control in IEEE 802.11 e wireless LANS. IEEE Netw 2005;19(4):6–13.

García-Macías JA, Rousseau F, Berger-Sabbatel G, Toumi L, Duda A. Quality of service and mobility for the wireless internet. Wirel Netw 2003;9(4):341–52.

Garg P, Doshi R, Greene R, Baker M, Malek M, Cheng X. Using IEEE 802.11e MAC for QoS over wireless. In: Proceedings of the 2003 IEEE international performance, computing, and communications conference, 2003. IEEE; 2003. p. 537–42.

Garmonov AV, Cheon SH, Han KL, Park YS, Savinkov A, Filin S, et al. QoS -oriented intersystem handover between IEEE 802.11b and overlay networks. IEEE Trans Veh Technol 2008;57(2):1142–54.

Gast M. 802.11 wireless networks: the definitive guide. CA, USA: O'Reilly Media, Inc.; 2005.

Georges J-P, Divoux T, Rondeau E. A formal method to guarantee a deterministic behaviour of switched ethernet networks for time-critical applications. In: 2004 IEEE international symposium on computer aided control systems design. NY, USA: IEEE; 2004. p. 255–60.

Georges J-P, Divoux T, Rondeau E. Strict priority versus weighted fair queueing in switched ethernet networks for time critical applications. In: Proceedings of the 19th IEEE international parallel and distributed processing symposium, 2005. NY, USA: IEEE; 2005. p. 141.

Girod B, Farber N. Feedback-based error control for mobile video transmission. Proc IEEE 1999;87(10):1707–23.

Giustiniano D, Malone D, Leith DJ, Papagiannaki K. Measuring transmission opportunities in 802.11 links. IEEE/ACM Trans Netw (TON) 2010;18(5):1516–29.

Grilo A, Macedo M, Nunes M. A scheduling algorithm for QoS support in IEEE 802.11 networks. IEEE Wirel Commun 2003;10(3):36–43.

Gu D, Yuan Y, Zhang J. Sequential coordinated channel access in wireless networks. US Patent App. 10/888,398; July 10, 2004.

Guo Y, Chaskar H. Class-based quality of service over air interfaces in 4G mobile networks. IEEE Commun Mag 2002;40(3):132–7.

Hanzo L, Tafazolli R. A survey of QoS routing solutions for mobile ad hoc networks. IEEE Commun Surv Tutor 2007;9(2):50–70 [2nd quarter].

Hanzo I, Tafazolli R. Admission control schemes for 802.11-based multi-hop mobile ad hoc networks: a survey. IEEE Commun Surv Tutor 2009;11(4):78–108.

Haykin S. Cognitive radio: brain-empowered wireless communications. IEEE J Sel Areas Commun 2005;23(2):201–20.

Hiertz GR, Denteneer D, Stibor L, Zang Y, Costa XP, Walke B. The IEEE 802.11 universe. IEEE Commun Mag 2010;48(1):62–70.

Hou I-H, Borkar V, Kumar P. A theory of QoS for wireless. In: INFOCOM 2009, NY, USA: IEEE; April 2009. p. 486–94.

How KC, Ma M, Qin Y. Routing and QoS provisioning in cognitive radio networks. Comput Netw 2011;55(1):330–42.

Hsu C-Y, Ortega A, Reibman AR. Joint selection of source and channel rate for VBR video transmission under ATM policing constraints. IEEE J Sel Areas Commun 1997;15(6):1016–28.

Indumathi G, Murugesan K. An adaptive time slot allocation for statistical QoS guarantees in wireless networks using crosslayer approach. Int J Commun Netw Inf Secur 2010;2(1).

Ishimori A, Farias F, Cerqueira E, Abelém A. Control of multiple packet schedulers for improving QoS on OpenFlow/SDN networking. In: 2013 second european workshop on software defined networks (EWSDN), NY, USA: IEEE; 2013. p. 81–6.

Jacobson V, Nichols K, Poduri K. An expedited forwarding PHB. RFC 2598. RFC 2598; 1999.

Jacobson V. Congestion avoidance and control. In: ACM SIGCOMM computer communication review, vol. 18 no. 4. NY, USA: ACM; 1988. p. 314–29.

Jasperneite J, Neumann P, Theis M, Watson K. Deterministic real-time communication with switched Ethernet. In: Proceedings of the fourth IEEE international workshop on factory communication systems. Citeseer. NY, USA: IEEE; 2002. p. 11–8.

Jelassi S, Rubino G, Melvin H, Youssef H, Pujolle G. Quality of experience of VoIP service: a survey of assessment approaches and open issues. IEEE Commun Surv Tutor 2012;14(2):491–513.

Jha S, Hassan M. Engineering internet QoS. Norwood, Massachusetts, USA: Artech House; 2002.

Jiang Y, Tham C-K, Ko C-C. A probabilistic priority scheduling discipline for multi-service networks. Comput Commun 2002;25(13):1243–54.

Jing X, Mau S-C, Raychaudhuri D, Matyas R. Reactive cognitive radio algorithms for co-existence between IEEE 802.11b and 802.16a networks. In: IEEE global telecommunications conference, 2005. GLOBECOM'05, vol. 5. NY, USA: IEEE; 2005. 5p.

Kassler A, Skorin-Kapov L, Dobrijevic O, Matijasevic M, Dely P. Towards QoE-driven multimedia service negotiation and path optimization with software defined networking. In: 2012 20th international conference on software, telecommunications and computer networks (SoftCOM). NY, USA: IEEE; 2012. p. 1–5.

Keshav S. An engineering approach to computer networking: ATM networks, the internet, and the telephone network, vol. 11997, Reading, MA: Addison-Wesley Professional; 1997.

Khan M, Khan TA, Beg M. Optimization of wireless network MAC layer parameters. Int J Innov Technol Explor Eng 2013;2(6) (ISSN: 2278-3075), http://www.ijitee. org/attachments/File/v2i6/F0831052613.pdf.

Kim D, Min C-H, Kim S. On-demand SIR and bandwidth-guaranteed routing with transmit power assignment in ad hoc mobile networks. IEEE Trans Veh Technol 2004;53(4):1215–23.

Kim, W, Sharma, P, Lee, J, Banerjee, S, Tourrilhes, J, Lee, et al. Automated and scalable QoS control for network convergence. In: Proceedings of INM/WREN, vol. 10; 2010. p. 1-1.

Knightly EW, Shroff NB. Admission control for statistical QoS: theory and practice. IEEE Netw 1999;13(2):20–9.

Kosek-Szott K, Natkaniec M, Pach AR. A simple but accurate throughput model for IEEE 802.11 EDCA in saturation and non-saturation conditions. Comput Netw 2011;55(3):622–35.

Kowalski J. Hybrid coordination in an IEEE 802.11 network. US Patent 8,472,416; June 25, 2013.

Krithika P, Pushpavalli M. Quality of service optimization in IEEE 802.11e networks using enhanced distributed channel access techniques. In: International journal of computer networks and wireless communications (IJCNWC); 2012.

Kritikos K, Plexousakis D. Semantic qos-based web service discovery algorithms. In: The fifth European conference on web services, 2007. ECOWS'07. NY, USA: IEEE; 2007. p. 181–90.

Ksentini A, Naimi M, Nafaa A, Gueroui M. Adaptive service differentiation for QoS provisioning in IEEE 802.11 wireless ad hoc networks. In: Proceedings of the first ACM international workshop on performance evaluation of wireless ad hoc, sensor, and ubiquitous networks. NY, USA: ACM; 2004. p. 39–45.

Ksentini A, Nafaa A, Gueroui A, Naimi M. ETXOP: a resource allocation protocol for QoS-sensitive services provisioning in 802.11 networks. Perform Eval 2007;64 (5):419–43.

Kumar A, Shin KG. Managing TCP connections in dynamic spectrum access based wireless LANs. In: 2010 the seventh annual IEEE communications society conference on sensor mesh and ad hoc communications and networks (SECON). NY, USA: IEEE; 2010. p. 1–9.

Kwak J-Y, Nam J-S, Kim D-H. A modified dynamic weighted round robin cell scheduling algorithm. ETRI J 2002;24(5):360–72.

Kwon T-G, Lee S-H, Rho J-K. Scheduling algorithm for real-time burst traffic using dynamic weighted round robin. In: Proceedings of the 1998 IEEE international symposium on circuits and systems, 1998. ISCAS'98, vol. 6. NY, USA: IEEE; 1998. p. 506–9.

Kyasanur P, Vaidya NH. Routing and link-layer protocols for multi-channel multi-interface ad hoc wireless networks. ACM SIGMOBILE Mob Comput Commun Rev 2006;10(1):31–43.

Lai C-F, Honggang W, Chao H-C, Guofang N. A network and device aware QoS approach for cloud-based mobile streaming. IEEE Trans Multimed 2013;15 (4):747–57.

LaMaire RO, Krishna A, Bhagwat P, Panian J. Wireless LANs and mobile networking: standards and future directions. IEEE Commun Mag 1996;34(8):86–94.

Lee C, Lee G. A version adaptive transcoding (VAT) mechanism for wireless mobile coverage networking. In: World automation congress (WAC), 2014. NY, USA: IEEE; 2014. p. 900–5.

Lee J, Uddin M, Tourrilhes J, Sen S, Banerjee S, Arndt M, et al. meSDN: mobile extension of SDN. In: Proceedings of the fifth ACM workshop on mobile cloud computing and services (MCS); 2014.

Leith DJ, Clifford P, Malone D, Ng A. TCP fairness in 802.11e WLANs. IEEE Commun Lett 2005;9(11):964–6.

Li M, Prabhakaran B. MAC layer admission control and priority re-allocation for handling QoS guarantees in non-cooperative wireless LANS. Mob Netw Appl 2005;10(6):947–59.

Li Q, VanderSchaar M. Providing adaptive QoS to layered video over wireless local area networks through real-time retry limit adaptation. IEEE Trans Multimed 2004;6(2):278–90.

Li Y, Chen CS, Song Y-Q, Wang Z. Real-time QoS support in wireless sensor networks: a survey. In: The seventh IFAC international conference on fieldbuses and networks in industrial and embedded systems-FeT'2007; 2007.

Lim L, Malik R, Tan P, Apichaichalermwongse C, Ando K, Harada Y. A QoS scheduler for IEEE 802.11e WLANs. In: The first IEEE consumer communications and networking conference, 2004. CCNC 2004. IEEE; 2004. p. 199–204.

Lin CR, Liu J-S. QoS routing in ad hoc wireless networks. IEEE J Sel Areas Commun 1999;17(8):1426–38.

Lindgren A, Almquist A, Schelén O. Evaluation of quality of service schemes for IEEE 802.11 wireless lans. In: Proceedings of the 26th Annual IEEE conference on local computer networks, 2001. LCN 2001. NY, USA: IEEE; 2001. p. 348–51.

Lindgren A, Almquist A, Schelén O. Quality of service schemes for IEEE 802.11 wireless LANs: an evaluation. Mob Netw Appl 2003;8(3):223–35.

Liu Y, Knightly E. Opportunistic fair scheduling over multiple wireless channels. In: IEEE INFOCOM 2003. twenty-second annual joint conference of the IEEE computer and communications, vol. 2. NY, USA: IEEE; 2003. p. 1106–15.

Liu H, Ma H, el Zarki M, Gupta S. Error control schemes for networks: an overview. Mob Netw Appl 1997;2(2):167–82.

Liu W, Chen X, Fang Y, Shea JM. Courtesy piggybacking: supporting differentiated services in multihop mobile ad hoc networks. IEEE Trans Mob Comput 2004;3(4):380–93.

Liu Q, Zhou S, Giannakis GB. Cross-layer scheduling with prescribed QoS guarantees in adaptive wireless networks. IEEE J Sel Areas Commun 2005;23(5):1056–66.

Liu Q, Wang X, Giannakis GB. A cross-layer scheduling algorithm with QoS support in wireless networks. IEEE Trans Veh Technol 2006;55(3):839–47.

Lu S, Bharghavan V, Srikant R. Fair scheduling in wireless packet networks. IEEE/ACM Trans Netw (TON) 1999;7(4):473–89.

Luo H, Shyu M-L, Chen S-C. An optimal resource utilization scheme with end-to-end congestion control for continuous media stream transmission. Comput Netw 2006;50(7):921–37.

Luo H, Shyu M-L, Chen S-C. Video streaming over the internet with optimal bandwidth resource allocation. Multimed Tools Appl 2008;40(1):111–34.

Mahadevan I, Sivalingam KM. Quality of service architectures for wireless networks: intserv and diffserv models. In: Proceedings of fourth international symposium on parallel architectures, algorithms, and networks, 1999. (I-SPAN'99). NY, USA: IEEE; 1999. p. 420–5.

Majumda A, Sachs DG, Kozintsev IV, Ramchandran K, Yeung MM. Multicast and unicast real-time video streaming over wireless lans. IEEE Trans Circuits Syst Video Technol 2002;12(6):524–34.

Malone D, Clifford P, Leith DJ. MAC layer channel quality measurement in 802.11. IEEE Commun Lett 2007;11(2):143–5.

Mangold S, Choi S, May P, Klein O, Hiertz G, Stibor L. IEEE 802.11e wireless LAN for quality of service. In: Proceedings of European wireless, vol. 2; 2002. p. 32–9.

Mangold S, Choi S, Hiertz GR, Klein O, Walke B. Analysis of IEEE 802.11e for QoS support in wireless LANs. IEEE Wirel Commun 2003;10(6):40–50.

Martins FC, Ding W, Feig E. Joint control of spatial quantization and temporal sampling for very low bit rate video. In: 1996 IEEE international conference on acoustics, speech, and signal processing, 1996. ICASSP-96. Conference proceedings, vol. 4. NY, USA: IEEE; 1996. p. 2072–5.

Matos R, Coutinho N, Marques C, Sargento S, Chakareski J, Kassler A. Quality of experience-based routing in multi-service wireless mesh networks. In: 2012 IEEE international conference on communications (ICC). NY, USA: IEEE; 2012. p. 7060–5.

Meraki Networks [Online], ⟨http://meraki.cisco.com⟩; 2014 [accessed: 2014-06-1].

Mogre PS, Hollick M, Steinmetz R. QoS in wireless mesh networks: challenges, pitfalls, and roadmap to its realization. In: International workshop on network and operating systems support for digital audio and video (NOSSDAV); 2007.

Moh S. Link quality aware route discovery for robust routing and high performance in mobile ad hoc networks. In: The 11th IEEE international conference on high performance computing and communications, 2009. HPCC'09. NY, USA: IEEE; 2009. p. 281–8.

Mohapatra P, Li J, Gui C. QoS in mobile ad hoc networks. IEEE Wirel Commun 2003;10(3):44–53.

Mokhtar SB, Preuveneers D, Georgantas N, Issarny V, Berbers Y. Easy: efficient semantic service discovery in pervasive computing environments with QoS and context support. J Syst Softw 2008;81(5):785–808.

Morris KJ, Hudson DJ, Goyal A. Automatic adaptive network traffic prioritization and shaping. US Patent 7,460,476; December 2, 2008.

Nam CH, Liew SC, Fu CP. An experimental study of ARQ protocol in 802.11 b wireless LAN. In: Proceedings of wireless personal multimedia communications (WPMC 2002); 2002.

Nam-Seok K, Hwanjo H, Jong-Dae P, Hong-Shik P. OpenQFlow: scalable openflow with flow-based QoS. IEICE Trans Commun 2013;96(2):479–88.

Nef M-A, Perlepes L, Karagiorgou S, Stamoulis GI, Kikiras PK. Enabling QoS in the internet of things. In: CTRQ 2012, the fifth international conference on communication theory, reliability, and quality of service; 2012. p. 33–8.

Ng AC, Malone D, Leith DJ. Experimental evaluation of TCP performance and fairness in an 802.11 e test-bed. In: Proceedings of the 2005 ACM SIGCOMM workshop on experimental approaches to wireless network design and analysis. NY, USA: ACM; 2005. p. 17–22.

Ni Q, Turletti T. QoS support for IEEE 802.11 wireless LAN. In: INRIA, Sophia Antipolis, France; 2004.

Ni Q, Romdhani L, Turletti T. A survey of QoS enhancements for IEEE 802.11 wireless LAN. Wirel Commun Mob Comput 2004;4(5):547–66.

Ni Q. Performance analysis and enhancements for IEEE 802.11e wireless networks. IEEE Netw 2005;19(4):21–7.

Nichols K, Jacobson V. A two-bit differentiated services architecture for the internet. RFC 2638; 1999.

Nor S, Mohd A, Cheow C. An admission control method for IEEE 802.11e. In: Network theory and applications. NY, USA: Springer; 2006. p. 105–22.

Nunes B, Mendonca M, Nguyen X, Obraczka K, Turletti T. A survey of software-defined networking: past, present, and future of programmable networks. IEEE Commun Surv Tutor 2014(99):1–18.

Pérez-Costa X, Camps-Mur D. IEEE 802.11e QoS and power saving features overview and analysis of combined performance [accepted from open call]. IEEE Wirel Commun 2010;17(4):88–96.

Parekh AK, Gallager RG. A generalized processor sharing approach to flow control in integrated services networks: the single-node case. IEEE/ACM Trans Netw (ToN) 1993;1(3):344–57.

Perros HG, Elsayed KM. Call admission control schemes: a review. IEEE Commun Mag 1996;34(11):82–91.

Peterson LL, Davie BS. Computer networks: a systems approach. NY, USA: Elsevier; 2007.

Politis A, Mavridis I, Manitsaris A. Exploiting multimedia frame semantics and MAC-layer enhancements for QoS provisioning in IEEE 802.11e congested networks. Int J Adv Netw Serv 2011;4(1 and 2):176–85.

Pollin S, Ergen M, Timmers M, Dejonghe A, Van der Perre L, Catthoor F, et al. Distributed cognitive coexistence of 802.15.4 with 802.11. In: The first international conference on cognitive radio oriented wireless networks and communications, 2006. IEEE; 2006. p. 1–5.

Qadir J, Ahmed N, Ahad N. Building programmable wireless networks: an architectural survey. EURASIP J Wirel Commun Netw (EURASIP JWCN) 2014, 1–31.

Qadir J. Artificial intelligence based cognitive routing for cognitive radio networks, arXiv preprint arXiv:1309.0085; 2013.

Qashi R, Bogdan M, Hanssgen K. Evaluating the QoS of WLANs for the IEEE 802.11 EDCF in real-time applications. In: 2011 international conference on communications and information technology (ICCIT). NY, USA: IEEE; 2011. p. 32–5.

Qu Q, Pei Y, Modestino JW, Tian X. Source-adaptation-based wireless video transport: a cross-layer approach. EURASIP J Appl Signal Process 2006;2006: 260.

Raniwala A, Chiueh T-C. Architecture and algorithms for an IEEE 802.11-based multi-channel wireless mesh network. In: INFOCOM 2005. Proceedings of the 24th annual joint conference of the IEEE computer and communications societies, IEEE, vol. 3. NY, USA: IEEE; 2005. p. 2223–34.

Rao KR, Bojkovic ZS, Bakmaz BM. Wireless multimedia communication systems: design, analysis, and implementation. CRC Press; 2014.

Rashid MM, Hossain E, Bhargava VK. HCCA scheduler design for guaranteed QoS in IEEE 802.11e based WLANs. In: IEEE wireless communications and networking conference, 2007. WCNC 2007. NY, USA: IEEE; 2007. p. 1538–43.

Rashid MM, Hossain E, Bhargava VK. Controlled channel access scheduling for guaranteed QoS in 802.11 e-based WLANs. IEEE Trans Wirel Commun 2008;7(4):1287–97.

Reddy TB, Karthigeyan I, Manoj B, Murthy C. Quality of service provisioning in ad hoc wireless networks: a survey of issues and solutions. Ad Hoc Netw 2006;4(1):83–124.

Romdhani L, Ni Q, Turletti T. Adaptive EDCF: enhanced service differentiation for IEEE 802.11 wireless ad hoc networks. In: 2003 IEEE wireless communications and networking, 2003. WCNC 2003, vol. 2. NY, USA: IEEE; 2003. p. 1373–8.

Rubin I, Liu Y-C. Link stability models for QoS ad hoc routing algorithms. In: 2003 IEEE 58th vehicular technology conference, 2003. VTC 2003-Fall, vol. 5. NY, USA: IEEE; 2003. p. 3084–8.

Seddiki MS, Shahbaz M, Donovan S, Grover S, Park M, Feamster N, et al. FlowQoS: QoS for the rest of us. In: Proceedings of the third workshop on hot topics in software defined networking. NY, USA: ACM; 2014. p. 207–8.

Setton E, Yoo T, Zhu X, Goldsmith A, Girod B. Cross-layer design of ad hoc networks for real-time video streaming. IEEE Wirel Commun 2005;12(4):59–65.

Sezer S, Scott-Hayward S, Chouhan P-K, Fraser B, Lake D, Finnegan J, et al. Are we ready for sdn? implementation challenges for software-defined networks IEEE Commun Mag 2013;51(7).

Sharkh MA, Jammal M, Shami A, Ouda A. Resource allocation in a network-based cloud computing environment: design challenges. IEEE Commun Mag 2013;51(11):46–52.

Sharma V, Malhotra J, Singh H. Quality of service (QoS) evaluation of IEEE 802.11 WLAN using different phy-layer standards. Optik—International Journal for Light and Electron Optics 2013;124(4):357–60.

Sharma S, Staessens D, Colle D, Palma D, Goncalves J, Figueiredo R, et al. Implementing quality of service for the software defined networking enabled future internet. In: The European workshop on software defined networking (EWSDN 2014). IEEE; 2014. p. 49–54.

Sheng M, Li J, Shi Y. Routing protocol with QoS guarantees for ad-hoc network. Electron Lett 2003;39(1):143–5.

Sheng Z, Leung KK, Ding Z. Cooperative wireless networks: from radio to network protocol designs. IEEE Commun Mag 2011;49(5):64–9.

Shenker S, Partridge C, Guerin R. Specification of guaranteed quality of service. RFC 2212. Internet RFC 2212. Technical Report; September 1997.

Sherwood R, Gibb G, Yap K-K, Appenzeller G, Casado M, McKeown N, et al. Flowvisor: a network virtualization layer. In: OpenFlow switch consortium. Technical Report; 2009.

Sivakumar R, Sinha P, Bharghavan V. CEDAR: a core-extraction distributed ad hoc routing algorithm. IEEE J Sel Areas Commun 1999;17(8):1454–65.

Sonkoly B, Gulyás A, Németh F, Czentye J, Kurucz K, Novak B, et al. On QoS support to Ofelia and OpenFlow. In: 2012 European workshop on software defined networking (EWSDN). NY, USA: IEEE; 2012. p. 109–13.

Stojmenovic I, Wen S. The fog computing paradigm: scenarios and security issues. In: 2014 federated conference on computer science and information systems (FedCSIS). NY, USA: IEEE; 2014. p. 1–8.

Su H, Zhang X. Cross-layer based opportunistic MAC protocols for QoS provisionings over cognitive radio wireless networks. IEEE J Sel Areas Commun 2008;26(1):118–29.

Sundareswaran N, Riley GF, Boyd K, Nainani A. Improving quality of service in MAC 802.11 layer. In: The 15th international symposium on modeling, analysis, and simulation of computer and telecommunication systems, 2007. MASCOTS'07. NY, USA: IEEE; 2007. p. 38–45.

Tanenbaum AS. Computer networks. 4th edition2003.

Tanigawa Y, Kim J-O, Tode H. QoS-aware retransmission with network coding based on adaptive cooperation with IEEE 802.11e EDCA. In: 2011 IEEE global telecommunications conference (GLOBECOM 2011). NY, USA: IEEE; 2011. p. 1–5.

Thomas RW, Friend DH, DaSilva LA, MacKenzie AB. Cognitive networks. NY, USA: Springer; 2007.

Tinnirello I, Choi S. Efficiency analysis of burst transmissions with block ACK in contention-based 802.11e WLANs. In: 2005 IEEE international conference on communications, 2005. ICC 2005, vol. 5. NY, USA: IEEE; 2005. p. 3455–60.

Toh C-K. Maximum battery life routing to support ubiquitous mobile computing in wireless ad hoc networks. IEEE Commun Mag 2001;39(6):138–47.

Tomovic S, Prasad N, Radusinovic I. Sdn control framework for QoS provisioning. In: The 22nd telecommunications forum Telfor (TELFOR), 2014. NY, USA: IEEE; 2014. p. 111–4.

Toumpis S, Goldsmith AJ. Performance, optimization, and cross-layer design of media access protocols for wireless ad hoc networks. In: IEEE international conference on communications, 2003. ICC'03, vol. 3. NY, USA: IEEE; 2003. p. 2234–40.

Tsao S-L. Extending earliest-due-date scheduling algorithms for wireless networks with location-dependent errors. In: The 52nd vehicular technology conference, 2000. IEEE-VTS Fall VTC 2000, vol. 1. NY, USA: IEEE; 2000. p. 223–8.

Turletti T, Huitema C. Videoconferencing on the internet. IEEE/ACM Trans Netw 1996;4(3):340–51.

Tursunova S, Kim Y-T. Realistic IEEE 802.11e EDCA model for QoS -aware mobile cloud service provisioning. IEEE Trans Consum Electron 2012;58(1):60–8.

Van Beek P, Deshpande S, Pan H, Sezan I. Adaptive streaming of high-quality video over wireless lans. In: Electronic imaging 2004. Bellingham, Washington, USA: International Society for Optics and Photonics(SPIE); 2004. p. 647–60.

van der Schaar M, Andreopoulos Y, Hu Z. Optimized scalable video streaming over IEEE 802.11 a/e HCCA wireless networks under delay constraints. IEEE Trans Mob Comput 2006;5(6):755–68.

Vicisano L, Crowcroft J, Rizzo L. TCP-like congestion control for layered multicast data transfer. In: INFOCOM'98. IEEE proceedings of seventeenth annual joint conference of the IEEE computer and communications societies, vol. 3. NY, USA: IEEE; 1998. p. 996–1003.

Viegas R, Guedes LA, Vasques F, Portugal P, Moraes R. A new MAC scheme specifically suited for real-time industrial communication based on IEEE 802.11 e. Comput Electr Eng 2013;39(6):1684–704.

Villalón J, Cuenca P, Orozco-Barbosa L. QoS provisioning mechanisms for ieee 802.11 wlan: a performance evaluation. In: Proceedings of 10th IFIP international conference on personal wireless communications. Colmar, France. Laxenburg, Austria: IFIP; 2005.

Wan J, Ullah S, Lai C-F, Zhou M, Wang X. Cloud-enabled wireless body area networks for pervasive healthcare. IEEE Netw 2013;27(5):56–61.

Wang M, Kuo G-S. An application-aware QoS routing scheme with improved stability for multimedia applications in mobile ad hoc networks. In: 2005 62nd IEEE vehicular technology conference, 2005. VTC-2005-Fall, vol. 3. IEEE; 2005. p. 1901–5.

Wang J, Zhai H, Fang Y, Yuang MC. Opportunistic media access control and rate adaptation for wireless ad hoc networks. In: 2004 IEEE international conference on communications, vol. 1; 2004. p. 154–8.

Wang J, Zhai H, Fang Y, Shea JM, Wu D. OMAR: utilizing multiuser diversity in wireless ad hoc networks. IEEE Trans Mob Comput 2006;5(12):1764–79.

Wang X, Vitvar T, Kerrigan M, Toma I. A QoS-aware selection model for semantic web services. In: Service-oriented computing–ICSOC 2006. NY USA: IEEE; 2006. p. 390–401.

Wang Z, Jiang T, Zhang L, Liu Y. Mobility and QoS oriented 802.11p MAC scheme for vehicle-to-infrastructure communications. In: 2011 sixth international ICST conference on communications and networking in China (CHINACOM). NY, USA: IEEE; 2011. p. 669–74.

Wang W, Qi Q, Gong X, Hu Y, Que X. Autonomic QoS management mechanism in software defined network. Commun China 2014;11(7):13–23.

Wang Z. Internet QoS: architectures and mechanisms for quality of service. Morgan Kaufmann; 2001.

Wen H, Tiwary PK, Le-Ngoc T. Wireless virtualization. In: Springer briefs in computer science. NY, USA: Springer; 2013.

Wiegand T, Lightstone M, Mukherjee D, Campbell TG, Mitra SK. Rate-distortion optimized mode selection for very low bit rate video coding and the emerging H. 263 standard. IEEE Trans Circuits Syst Video Technol 1996;6(2):182–90.

Wisitpongphan N, Ferrari G, Panichpapiboon S, Parikh J, Tonguz O. QoS provisioning using ber-based routing in ad hoc wireless networks. In: 2005 IEEE 61st vehicular technology conference, 2005. VTC 2005-Spring, vol. 4. NY, USA: IEEE; 2005. p. 2483–7.

Wu D, Hou YT, Zhu W, Lee H-J, Chiang T, Zhang Y-Q, et al. On end-to-end architecture for transporting MPEG-4 video over the internet. IEEE Trans Circuits Syst Video Technol 2000;10(6):923–41.

Wu D, Hou YT, Zhu W, Zhang Y-Q, Peha JM. Streaming video over the internet: approaches and directions. IEEE Trans Circuits Syst Video Technol 2001;11(3):282–300.

Wu H, Wang A, Zhang Q, Shen X. IEEE 802.11e enhanced distributed channel access EDCA throughput analysis. In: IEEE international conference on communications, 2006. ICC'06, vol. 1. NY, USA: IEEE; 2006. p. 223–8.

Wu H, Liu Y, Zhang Q, Zhang Z-L. SoftMAC: layer 2.5 collaborative MAC for multimedia support in multihop wireless networks. IEEE Trans Mob Comput 2007;6(1):12–25.

Wu H-T, Yang M-H, Ke K-W. The design of QoS provisioning mechanisms for wireless networks. In: 2010 the Eigth IEEE international conference on pervasive computing and communications workshops (PERCOM workshops). NY, USA: IEEE; 2010. p. 756–9.

Xianyang F, Feng W. Design and implementation of interference-aware cooperative QoS routing for multi-hop wireless network. In: 2014 the sixth international conference on measuring technology and mechatronics automation (ICMTMA). NY, USA: IEEE; 2014. p. 211–7.

Xiao Y, Li H. Local data control and admission control for QoS support in wireless ad hoc networks. IEEE Trans Veh Technol 2004;53(5):1558–72.

Xiao X, Ni LM. Internet QoS: a big picture. IEEE Netw 1999;13(2):8–18.

Xiao Y. IEEE 802.11e: QoS provisioning at the MAC layer. IEEE Wirel Commun 2004;11(3):72–9.

Xing Y, Mathur CN, Haleem MA, Chandramouli R, Subbalakshmi K. Dynamic spectrum access with QoS and interference temperature constraints. IEEE Trans Mob Comput 2007;6(4):423–33.

Xue Q, Ganz A. Ad hoc QoS on-demand routing (AQOR) in mobile ad hoc networks. J Parallel Distrib Comput 2003;63(2):154–65.

Yang X, Vaidya NH. Priority scheduling in wireless ad hoc networks. In: Proceedings of the third ACM international symposium on mobile ad hoc networking and computing. NY, USA: ACM; 2002. p. 71–9.

Yeh J-Y, Chen C. Support of multimedia services with the IEEE 802.11 MAC protocol. In: IEEE international conference on communications, 2002. ICC 2002, vol. 1. NY, USA: IEEE; 2002. p. 600–4.

Yin S, Xiong Y, Zhang Q, Lin X. Traffic-aware routing for real-time communications in wireless multi-hop networks. Wirel Commun Mob Comput 2006;6(6):825–843.

Yu X, Navaratnam P, Moessner K. Resource reservation schemes for IEEE 802.11-based wireless networks: a survey. IEEE Commun Surv Tutor 2013;15(3):1042–1061.

Zander J, Mahonen P. Riding the data tsunami in the cloud: myths and challenges in future wireless access. IEEE Commun Mag 2013;51(3):145–51.

Zhang S, Franklin DR. Feasibility study on the implementation of IEEE 802.11 on cloud-based radio over fibre architecture. In: IEEE international conference on communications (ICC), 2014. NY, USA: IEEE; 2014. p. 2891–6.

Zhang D, Ionescu D. QoS performance analysis in deployment of Diffserv-aware MPLS traffic engineering. In: The Eighth ACIS international conference on software engineering, artificial intelligence, networking, and parallel/distributed computing, 2007. SNPD 2007, vol. 3. NY, USA: IEEE; 2007. p. 963–7.

Zhang B, Mouftah HT. QoS routing for wireless ad hoc networks: problems, algorithms, and protocols. IEEE Commun Mag 2005;43(10):110–7.

Zhang Q, Zhang Y-Q. Cross-layer design for QoS support in multihop wireless networks. Proc IEEE 2008;96(1):64–76.

Zhang L, Deering S, Estrin D, Shenker S, Zappala D. RSVP: a new resource reservation protocol. IEEE Netw 1993;7(5):8–18.

Zhang Q, Yang F, Zhu W. Cross-layer QoS support for multimedia delivery over wireless internet. EURASIP J Adv Signal Process 2005;2005(2):207–19.

Zhang D-Y, Jiang J, Anani A, Li H-B. QoS-guaranteed packet scheduling in wireless networks. J China Univ Posts Telecommun 2009;16(2):63–7.

Zhao J, Guo Z, Zhang Q, Zhu W. Performance study of MAC for service differentiation in IEEE 802.11. In: IEEE Global telecommunications conference, 2002. GLOBECOM'02, vol. 1. NY, USA: IEEE; 2002. p. 778–82.

Zhao D, Zhu M, Xu M. Leveraging SDN and openflow to mitigate interference in enterprise wlan. J Netw 2014;9(6):1526–33.

Zhou G, Stankovic JA, Son SH. Crowded spectrum in wireless sensor networks. IEEE EmNets 2006;6.

Zhu H, Li M, Chlamtac I, Prabhakaran B. A survey of quality of service in IEEE 802.11 networks. IEEE Wirel Commun 2004;11(4):6–14.

Zhu X, Agrawal P, Pal Singh J, Alpcan T, Girod B. Rate allocation for multi-user video streaming over heterogeneous access networks. In: Proceedings of the 15th international conference on multimedia. NY, USA: ACM; 2007. p. 37–46.

Zhu R. Intelligent rate control for supporting real-time traffic in wlan mesh networks. J Netw Comput Appl 2011;34(5):1449–58.