

REVIEW

Open Access

Quality of service provision in mobile multimedia - a survey

Hongli Luo^{1*} and Mei-Ling Shyu²

* Correspondence: luoh@ipfw.edu

¹Department of Computer and Electrical Engineering Technology and Information Systems and Technology, Indiana University - Purdue University Fort Wayne, Fort Wayne, IN, USA

Full list of author information is available at the end of the article

Abstract

The prevalence of multimedia applications has drastically increased the amount of multimedia data. With the drop of the hardware cost, more and more mobile devices with higher capacities are now used. The widely deployed wireless LAN and broadband wireless networks provide the ubiquitous network access for multimedia applications. Provision of Quality of Service (QoS) is challenging in mobile ad hoc networks because of the dynamic characteristics of mobile networks and the limited resources of the mobile devices. The wireless network is not reliable due to node mobility, multi-access channel and multi-hop communication. In this paper, we provide a survey of QoS provision in mobile multimedia, addressing the technologies at different network layers and cross-layer design. This paper focuses on the QoS techniques over IEEE 802.11e networks. We also provide some thoughts about the challenges and directions for future research.

Keywords: Quality of Service (QoS), mobile computing, multimedia, 802.11, 802.11e, survey

1. Introduction

There is a rapidly growing demand for real-time multimedia services, such as video streaming, video conferencing, and IPTV. Mobile devices, such as smart phones, PDAs, and laptops, become more and more popular and powerful, and are enabled to access and present rich multimedia contents. Multimedia data is also one of the major factors that drive the development of broadband wireless networks. Broadband wireless networks, such as WiMAX (Worldwide Interoperability for Microwave Access) and 3G (3rd Generation Mobile Telecommunications), are widely used for mobile and wireless Internet access. The heterogeneous and widely deployed wireless networks have made the pervasive and ubiquitous computing possible, which means the access to multimedia data from anywhere at any time. Mobile video streaming applications like mobile TV, mobile gaming, etc. have become the most popular applications on the mobile devices. Multimedia data are also widely used at surveillance, homeland security, transportation, distance learning, health care, etc. The Internet Service Providers (ISPs) are expected to provide multimedia services via multiple wireless networking technologies, such as WLAN (Wireless Local Area Network), 3G, and WiMAX.

Real-time multimedia over the Internet has its quality of service (QoS) requirements, which includes bandwidth, packet loss ratio, delay, and jitter. More sophisticated QoS protocols are typically required for multimedia applications. In particular, the provision

of QoS for multimedia applications in a mobile environment imposes a series of major challenges because of the unreliable wireless channels and the mobility of mobile devices.

- **Unreliable physical channels**

Wireless channels are highly unreliable and have limited bandwidth. Wireless channels have high packet loss rate and bit error rate because of fading and multipath effects. The wireless medium is shared by multiple stations and the bandwidth allocation to one station will be affected by the neighboring stations. Because of the contention characteristics of the channels and MAC layer access methods, it is hard to provide the guaranteed end-to-end delays for the multimedia applications.

- **Node mobility**

Mobile devices are roaming and switching the wireless networks they connect to. To provide a continuous service, the mobile device should be able to connect to the wireless network that is available. For example, a mobile phone may switch from one cell covered by one base station to another cell covered by another based station, or switch from the cellular phone network to a Wireless LAN. The application should be able to provide seamless handoff among different wireless networks and provide an uninterrupted playback of the video with an acceptable QoS.

- **Routing**

Because of the movement of the mobile devices, the topology of the mobile ad hoc networks varies dynamically. The existing routes may either not be available or not be able to support the QoS, which requires the changes of the routings. The selection of routes should be able to accommodate the changes of the topology and provide the QoS.

- **Resource constraints**

There are a number of limited resources on mobile devices, such as limited battery life, screen size, and input methods. The QoS is affected by the limited resources at the mobile devices, so the design of a mobile multimedia system should consider all those factors. The current battery technology is not evolving as fast as the memories and computer hardware. Both the processing and transmission of multimedia data consume power. With the limited power, it requires a power efficient design for both multimedia processing and transmission in mobile environment. The screen size of the mobile device is small, and mobile device is not equipped with full-size keyboards. All these limitations in the input and output pose many challenges in the design of the user interface.

- **Heterogeneity**

The heterogeneity of the mobile devices, access networks, and infrastructure networks makes the end-to-end QoS provision more difficult. The mobile devices have different screen sizes, screen resolutions, and decoder capabilities, so the same multimedia content should be adapted to the capabilities of different mobile devices in a way that is perceptual optimal for the users. There are also multiple wireless networks with different bandwidths used in mobile computing. In a heterogeneous wireless network environment, the mobile devices

should be equipped with multiple wireless interfaces so they can access different networks. Mobile multimedia systems should consider mobility and handoff management when mobile devices are moving among different wireless networks.

- **Evaluation metrics**

In addition to QoS, quality of experience (QoE) that the users experience in multimedia applications is another metric for the performance of a mobile multimedia system. QoE includes video quality, energy saving, and bandwidth efficiency [1]. User experience of mobile video is affected by many factors, such as user profile, interests, context, and content type.

WLAN is widely deployed because of its flexibility and low cost. WLAN provides network connectivity with minimal infrastructure change, and it is easy to set up, configure and manage. The varying and error-prone characteristics of the wireless medium pose challenges in providing QoS for multimedia applications. In an effort to address such challenges, the IEEE LAN/MAN 802 Standards Committee developed and maintained the 802.11 family, a series of over-the-air specifications or modularization techniques, to provide a set of standards for implementing WLAN computer communication by defining the media access control (MAC) and physical (PHY) layers for a LAN with wireless connectivity. In particular, since the traditional 802.11 cannot provide for the QoS, 802.11e was proposed to improve the functionality [2]. In the 802.11 family, 802.11e is the wireless standard that defines a set of Quality of Service (QoS) enhancements for WLAN applications through modifications to the MAC layer. 802.11e adds QoS features and multimedia support to the existing IEEE 802.11 wireless standards with full backward compatibility with those standards.

In addition, the convergence of various wireless networking technologies, such as WLAN, 3G, WiMAX, sensor networks, and RFID, also poses challenges to the QoS of multimedia applications. Mobile devices are heterogeneous in operating systems, CPU, memory, networking capabilities, and battery life. With the increased capacity, mobile devices are used for the entire range of multimedia applications: production, annotation, management, retrieval, sharing, communication, and content analysis, which also affects the way QoS is provided at the mobile devices [3]. The provision of QoS in mobile and ubiquitous multimedia covers multiple research areas, such as the hardware, architecture, protocol, software, and middleware support. Since 802.11 Wireless LAN is widely used for the mobile computing and 802.11e is the new mechanism proposed for QoS, this paper gives a survey on several aspects of research in the QoS provision in mobile multimedia with a focus on the 802.11e networks.

The remainder of this paper is organized as follows. Section 2 gives a general description of research in enabling QoS provision in mobile and ubiquitous multimedia. Then Section 3 presents the current QoS approaches in 802.11e MAC layer. Section 4 reviews the QoS design with cross-layer design. Other important design issues in the QoS provision in mobile multimedia are covered in Section 5, such as power efficient design, heterogeneity and directions for future research. Finally, this paper is concluded in Section 6.

2. Mobile Multimedia QoS Provision Architecture

The provision of QoS for mobile multimedia applications requires the support of the architectures, protocols, and applications so that the mobile devices can access the

multimedia data ubiquitously: anytime and anywhere. Multimedia transmission needs to meet the following requirements, namely high bandwidth, low error rate, low delay, and very small delay variance. As mentioned in [4], the current research effort cannot provide solutions to fulfill all of these requirements for even the wired media. It is thus well-acknowledged that it is even more challenging to meet these requirements for high-quality multimedia transmission over wireless connections. The QoS of multimedia applications are not limited to bandwidth, delay, and jitter. Furthermore, the services provided to the mobile devices should be personalized. Song et al. [5] studied two ways of emphasizing Region of Interest (ROI), zooming in and enhancing the quality to optimize the overall user experience of viewing sports videos on mobile phones. It found out the overall user experience is closely related to the acceptance of video quality and the interest in video content.

There are several methodologies to categorize the QoS research in a mobile environment. QoS support can be provided as a layered model or across several layers. Layered QoS is implemented at one network layer, such as MAC layer, network layer, transport layer, or application layer. There are some commonly used technologies at a particular network layer, such as rate control, admission control, and scheduling. The major techniques used at the MAC layer include admission control and scheduling. The MAC layer schemes of wireless network can be categorized into three types: TDMA, CDMA, and IEEE 802.11. This paper focuses on the scheduling technique at the 802.11e MAC layer.

The majority of QoS research at the network layer has focused on the QoS routing. A multimedia application often has stringent requirements on the delay. QoS routing determines the delivery path for flows taking into account both the availability of network resources and the QoS requirements of the flows. There are active researches in providing QoS aware routing algorithm for mobile multimedia applications. Researches in the recent QoS routing in mobile ad hoc networks have been covered in [6]. One of the major functions at the transport layer is congestion control and the TCP protocol is the dominant protocol at the transport layer. TCP protocol is designed for the wired networks and is not efficient for wireless networks. It reduces the transmission rate when there is a packet loss, which suffers great performance degradation since the wireless channel generates a higher bit error rate. The transport layer protocol should be able to differentiate the packet losses generated by the congestion and by the channel errors [7-9].

Research at the application layer QoS includes scalable video coding [10], transcoding [11], source coding [12], adaptive transmission [13], and rate control [14]. Adaptive transmission exploits the unequal importance of different packets to improve the end-to-end quality of the video. The transmission rate and coding are context-aware, which is adaptive to the network situations, video content, user preferences, and several other factors. Context-aware computing is now a mobile computing paradigm to discover and utilize the contextual information in providing services [15,16]. Middleware is also introduced as an abstract layer to separate the low-level data processing from the high-level applications in the mobile computing environment.

The traditional layered model is designed for the wired networks, and is not efficient for the provision of QoS in mobile networks. Cross-layer design is a promising direction which jointly designs the mechanisms at several layers and achieves the

optimization of the performance. The layered and cross-layer approaches are implemented at the end system, either client side or server side of the multimedia applications. In addition to the adaptations at the end system, QoS can also be provided with the support of the network routers. Furthermore, besides the traditional client-server paradigm for networking applications, peer-to-peer networks have been widely deployed to provide live and on-demand video streaming services in the Internet. A survey of current research on how to provide QoS in peer-to-peer mobile multimedia applications is provided in [17]. In this paper, we focus on the QoS provision at the MAC layer of the IEEE 802.11e WLAN and the QoS provision with the cross-layer design.

3. QoS Provision at the MAC Layer

In this section, the researches related to the QoS provision at the MAC layer of the IEEE 802.11e are discussed. IEEE 802.11 has been widely applied as the technology to provide the mobile and pervasive computing. The MAC layer of the original IEEE 802.11 standard is based on the CSMA/CA mechanism, which does not support QoS of real-time applications. Toward such demands, the IEEE 802.11e standard proposed Hybrid Coordination Function (HCF) to enhance the media access for QoS. HCF is composed of Enhanced Distributed Coordination Access (EDCA) and HCF Controlled Channel Access (HCCA). Many researchers have worked on the tuning of parameters in 802.11e to improve QoS. Here, a review of the scheduling methods used at the EDCA and HCCF is presented separately. Control theoretical approaches used for QoS are also addressed.

3.1 QoS at Enhanced Distributed Coordination Access (EDCA)

EDCA is a contention-based channel access and provides service differentiation in IEEE 802.11e. Differentiated accesses to the wireless medium are provided by prioritizing the traffic categories (TC). There are at most eight prioritized output queues, one for each of the traffic categories. Changing the priorities of the traffic flows can consequently change the QoS received by the traffic flows. Different contention parameters can be tuned adaptively for each access category (AC) to treat the traffic flows differently. EDCA provides differentiated services among different traffic classes, but cannot provide the guaranteed throughputs and bounded delays. It provides a higher QoS to traffic flows with a higher priority while sacrificing the traffic flows with a lower priority, especially when the traffic load is heavy. The performance analysis of IEEE 802.11e EDCA was presented in [18] and [19].

Research work at the EDCA mechanism adjusts several parameters to prioritize the traffic and differentiate the service classes, such as Arbitrary Inter-frame Spacing (AIFS), minimum contention windows (CW_{min}), maximum contention window (CW_{max}), and persistence factor (PF). Multiple parameters can be adaptively adjusted according to the conditions of the network, such as congestion window [20] and priorities of the traffic. On the other hand, Adaptive EDCA (A-EDCA) in [21] provides the relative priorities by adjusting the size of the Congestion Window (CW) of each traffic class according to the estimated collision rate. The channel contention is effectively reduced under a high traffic load. Some researches dynamically adjust the priorities to improve the QoS. Li, Zhu and Prabhakaran [22] dynamically re-allocated the flow

priorities evenly to maintain high system performance while providing QoS for individual real-time flows. The overall throughput of the network can be improved by evenly distributing the number of active stations over a set of traffic categories. Han et al. [23] extended EDCA with channel access throttling, which differentiated channel access priorities between member stations by assigning different channel access parameters to different member stations. Patras, Banchs and Serrano [20] adapted the congestion window to the conditions of the WLAN based on the analytical model of its operation. The algorithm is based on the observation that the collision probability in an optimally configured WLAN is approximately constant, independent of the member stations. The collision probability is measured by monitoring the successfully transmitted frames during an inter-beacon period at the AP (Access Point).

In addition to the provision of QoS, fairness in the allocation of resources to different services is also an important issue in WLANs. Ferng and Liao [24] proposed four fair scheduling schemes for the QoS-oriented wireless LAN that take into account priority setting, fairness, and cross-layer interactions. Those schemes target at reducing possible collisions using multiple deficit count to interframe space (IFS) and allowance to IFS mappings for different priorities. Park et al. [25] provided per-class QoS enhancement and per-station fair channel sharing in WLAN access networks. It improves QoS for different service classes by differentiating services with scheduling and queue management. The fair channel sharing is assured by estimating the fair share for each station and dynamically adjusting the service levels of packets.

3.2 QoS at HCF Controlled Channel Access (HCCA)

HCCA provides a centralized polling scheme to allocate guaranteed channel access to traffic flows based on their QoS requirements. The superframe is divided into contention-free period (CFP) and contention period (CP). During the CP, the access to the channel is controlled by EDCA. HCCA is in charge of the contention-free medium access. Hybrid coordinator (HC) can initiate controlled access periods (CAPs) at any time. A wireless station has the right to initiate frame exchange sequences onto the wireless medium for an interval of time, which is called transmission opportunity (TXOP). HC is responsible for allocating TXOPs to each mobile station according to the QoS requirement of the traffic. HCCA provides a reference design which consists of a reference scheduling and an admission control. In the reference design, the scheduler first calculates a common service interval (SI), which is the minimum of the delay bounds of streams. After that, the scheduler calculates the TXOP duration according to the SI and the traffic specification parameters (TSPEC) such as the mean data rate and the mean packet size. Then the admission control is performed and the TXOPs are allocated to each station in a round robin way. The shorter is the SI, the shorter is the scheduling interval. Consequently, more bandwidth is needed to transmit the arrived packets. Therefore, the reference design over-allocates bandwidth to the stations since the allocation is done according to the stringent delay bound of all streams. To overcome this limitation, there are various researches using adaptive scheduling that dynamically tunes the parameters to improve the performance of HCCA.

Different from the reference scheduling which schedules all streams with a common spacing, the equal-SP scheduling proposed by Zhao and Tsang [26] schedules each stream with equal spacing, but at the same time it also schedules different streams

with different spacing. The traffic scheduling algorithm proposed by Skyrianoglou, Passas and Salkintzis [27], which is referred to as adaptive resource reservation over WLANs (ARROW), performs channel allocation based on the actual traffic buffered in the various mobile stations. The adaptive TXOP allocation mechanism proposed in [28] by Arora et al. works in accordance with the channel and traffic conditions and complies with the link adaptation mechanism to ensure long-term fairness among the wireless stations. Ghazizadeh and Fan [29] allocate the channel based on hybrid estimation and error correction according to the actual queuing duration of each mobile station. Cicconetti et al. [30] exploit cross-layer information, such as the application packet generation pattern and application packet generation interval, to design an effective bandwidth sharing and polling strategy.

The information of queue sizes is used widely as feedback information for the calculation of TXOP [31]. Ansel, Ni, and Turletti [32] present an efficient scheduling at the access point, which is based on the measured queue sizes for each traffic stream at each wireless station. The transmission time for each wireless station is assigned based on the queue length and aims at depleting the queue. The scheduling in [33] uses proportional-integral controller to provide a bounded delay for different traffic classes based on the queue length at the mobile station. An adaptive application aware scheduler for HCCA was proposed in [34], which allocates adaptive service intervals, transmission opportunities, and polling order based on the traffic characteristics and instantaneous network conditions. In [35], the allocation of transmission time for TXOP is based on both the queue length and the incoming packet rates of each flow at the wireless station.

3.3 Control-theoretic Approach

There are multiple applications of control theory in the provision of QoS in multimedia applications. Feedback control has been widely used in the design of many aspects of computing [36]. Control theory provides a systematic approach to design feedback systems to improve the performance of a computing system. The goal of control theory is to design a system that is stable to avoid wild oscillation, accurate to provide target response time, and quick to settle to their steady state values. It is used for the packet scheduling and bandwidth allocations in the traditional computer networking applications, such as congestion control [37] and resource management. PI (Proportional and Integral) controller [38,20,39,40] and P (Proportional) controller [41] are widely used for traffic rate control.

Recently, control theory is also used to provide QoS in the multimedia applications. For example, feedback control can be used to adjust the scheduling priorities in the MAC layer of WLAN. In [38], PI controller is designed to adjust the priorities of the application to control the end-to-end delay around the required delay level. PI controller for a priority adaptor is determined off-line based on the dynamic input-output pairs via system identification. Then the adaptive controller is implemented to adapt its behavior according to the changing load and network conditions. Huang, Mao and Midkiff [41] use control theory to understand the end-to-end streaming system and develop algorithms for quality control by rate control. Proportional controllers are used to stabilize the received video quality as well as the bottleneck link queue. PI controller is used in [20] to adaptively adjust the CW according to the conditions of the

WLAN. The scheduling algorithm at 802.11e HCCA in [35] is based on optimal control. A quadratic performance index is introduced to obtain an optimal scheduling which minimizes the packet delays at the cost of a small transmission time.

4. QoS Provision with the Cross-Layer Design

Traditional layered design cannot provide QoS for mobile multimedia because of its limited adaptation to the dynamic wireless channels and interaction between layers. The goal of a cross-layer design is to improve the overall performance of the mobile multimedia applications, including the quality of video and power consumption. Cross-layer design jointly adjusts the parameters of different network layers, but the computation is complicated. Cross-layer optimization is very complex since it requires the optimization of multiple parameters across the network layers. One of the challenges of cross-layer design is the difficulty to model the complex cross-layer interactions among the parameters at different network layers. In general, there is a trade-off between the performance and complexity in the cross-layer optimization. A low-complexity cross-layer design is desired.

A cross-layer design enhances the performance of the application by jointly considering the mechanisms at multiple network layers. For example, modulation and coding scheme at the physical layer, scheduling and admission control at the MAC layer, routing at the network layer, congestion control and rate control at the transport layer, and source coding, traffic shaping, scheduling, and rate control at the application layer. Cross-layer QoS mechanisms proposed for 802.11 WLAN can be divided into different categories according to the layers involved.

- Application-PHY layer

Joint source-channel coding at the application layer has been extensively studied [42,43]. Argyriou [42] provided a methodology for joint setting of the parameters of source and channel coding based on an analytical model of the overall system. It employs joint source and application-layer channel coding and rate adaptation at the wireless physical layer.

- Application-Transport layer- MAC/PHY layer

Zhu, Zeng and Li [8] proposed a joint design of source rate control and congestion control for video streaming over the Internet. A virtual network buffer management mechanism was introduced and the QoS of the application was translated into the constraints of the source rate and the sending rate. At the transport layer, a QoS-aware congestion control mechanism was proposed to meet the sending rate requirement derived from the virtual buffer. The joint optimization of parameters in [9] was designed to minimize the expected end-to-end video distortion constrained by a given video playback delay. It includes video coding at the application layer, packet sending rates at the transport layer, and the modulation and coding scheme at the physical layer. A cross-layer design is proposed in [44] that incorporates source rate control at the application layer, congestion control at the transport layer, and wireless loss ratio from the MAC layer.

- Application-MAC layer

Van Der Schaar and Turaga [45] developed a joint application-layer adaptive packetization and prioritized scheduling and MAC-layer retransmission strategy, where the application and the MAC layers jointly decide the optimal packet size and

retransmission limits. Cross-layer design in [46] utilized the data partitioning technique at the application layer and QoS mapping technique at the EDCA-based MAC layer of the 802.11e network. Chilamkurti et al. [47] proposed a cross-layer design for 802.11e which maps video packets at the application layer to the appropriate access categories of 802.11e EDCA at the MAC layer according to the significance of the video data. The approach proposed for IEEE 802.11e HCCA WLAN in [48] consists of admission control and resource allocation at the MAC layer and video adaptation at the application layer.

- Application-MAC-PHY layer

Van Der Schaar, Andreopoulos and Hu [10] proposed an optimization over Application-MAC-PHY layer for scalable video over IEEE 802.11 HCCA. It maximizes the number of admitted stations by creating multiple subflows from one global video flow. Shankar and Van Der Schaar [49] proposed an integrated system view of admission control and scheduling for both content and poll-based access of IEEE 802.11e MAC protocol. The scheme in [1] set parameters at three layers: application, link, and physical layers. It is designed to optimize the video quality of all streams given different power levels and channel conditions of the wireless stations. Wu, Song and Wang [12] proposed a cross-layer optimization framework for delivering video summaries over wireless networks. It jointly optimizes the source coding at the application layer, allowable retransmission at the data link layer, and adaptive modulation and coding at the physical layer within a rate-distortion theoretical framework.

5. Considerations of QoS Provision in Mobile System

There are several factors that need to be considered in the provision of QoS, such as the limited resources on the mobile devices, heterogeneity, and roaming characteristics in the mobile computing environment. Power-efficient design of QoS is the common solution to address the limited battery on the mobile device. Context-aware middleware is used to overcome the heterogeneity issue in the mobile networks and provide context-aware QoS. Handover is essential in mobility management which provides QoS when the mobile devices move from one network to another network. The evolution of new applications and technologies, such as social multimedia and cloud computing, poses many challenges in the provision of QoS.

5.1 Power Efficient Design of QoS

Mobile devices running multimedia applications are limited in every supply. How to prolong the life time of the mobile devices to provide QoS under the energy constraint is important to the QoS provision. Two major operations of wireless multimedia applications that consume most of the energy of the mobile devices are video encoding and data transmission. Minimizing the overall energy consumption at the mobile devices is an active research area. Power-aware design for mobile multimedia considers both video coding and video delivery. Efficient encoding scheme can reduce the data rate of the transmission. At the same time it needs complicated computation, which consumes more power. The mobile device should adaptively adjust its computational complexity and energy consumption according to the contexts, such as network conditions and the contents of the video.

The power-aware multimedia solutions jointly design the video coding parameters and channel parameters to adapt to the video contents and underlying network conditions to minimize the total energy consumption [50]. An efficient system should jointly consider three factors: bit rate, power consumption, and video quality. A balance needs to be achieved between power consumption in computation and communication to provide energy efficient multimedia applications. The goal is to minimize the total power consumption, subject to three constraints: the maximum video distortion to ensure satisfactory video quality, maximum end-to-end delay required by the application, and the maximum computational complexity provided by the mobile multimedia devices. Another goal is to minimize the video distortion, subject to the maximum power consumption allowed, maximum end-to-end delay, and maximum computation complexity. Power-rate-distortion analysis adds a new dimension power to the traditional rate-distortion analysis. The complexity parameters of the video encoding scheme can be dynamically adjusted to maximize the video quality under the energy constraint of the mobile device.

Video quality is defined as the mean square error (MSE) between the original video frames and the decoded video frames [51]. Hsu and Hefeeda [52] adopted a power-rate-distortion model to capture the trade-off among the encoding rate, energy consumption of the encoder, and the video distortion. The average video distortion is minimized via the adjustment of multiple link layer parameters. In addition to the power saving at the wireless client stations, power saving can also be considered at the access point of 802.11 networks. IEEE 802.11 standard defines two states for a wireless station, the Awake state and the Doze State. Zhang et al. [53] present IEEE 802.11-based power-saving access point (PSAP) used for solar/battery powered applications. Three different frame design arrangements were introduced for adaptive power saving sleep periods. The beacon broadcast of power-saving access point in [54] carries a network allocation map (NAM) to indicate its temporal operations, which coordinates traffic delivery and power saving at both end stations and the access point (AP). Both power saving and QoS are considered in [55] at the access points. QoS-enabled AP schedules its awakening and sleeping pattern in a way that satisfies the delay and packet loss requirements for the real-time flows.

5.2 Heterogeneity

The provision of QoS in mobile multimedia is challenging because of node mobility, multi-access channel, multi-hop communication, and the limited capabilities of the mobile devices. The delivery of multimedia content should be adapted to the network, user preference, and mobile terminals. The mobile devices have different capabilities such as display size, memory, and computational power. In addition, QoS should also depend on the contexts and adapt to the contexts. Context information includes the network connectivity (such as bandwidth and delay), location, user preferences, time, etc. A context-aware system is able to adapt its behavior according to the current context. Several issues need to be addressed in a context-aware system design. The key issue is how to obtain, store, and represent the context information. Because of the heterogeneity characteristics of the mobile devices, context-aware middleware is one of the common solutions to provide services for pervasive applications.

Upper-layer adaptation works at the high layers of network, such as the application layer. The advantage of the upper-layer adaptation is that it can be widely used on different PHY/MAC hardware and protocols since the adaptation is independent of the PHY/MAC layer protocols. With the help of the middleware, it can work with any physical wireless network.

The adaptation mechanism design in [38] spans the application layer, middleware layer, and network layer. The resources are allocated to multimedia applications under the coordination of three layers to meet the end-to-end requirements. The application layer adopts a requirement adapter to dynamically adjust the requirement levels according to the end-to-end delay measurement and QoS requirements acceptable to the end-users. Feedback control is used in the middleware layer to dynamically adjust the service classes for the applications according to the observed end-to-end delay. Then service differentiation scheduler at the network layer assigns the network resources to different service classes.

Mohapatra et al. [11] proposed a distributed middleware layer to perform joint adaptations at all levels of the system hierarchy for optimized performance and energy benefits on mobile handheld devices. It adopts an intermediate server in close proximity of the mobile device to perform end-to-end adaptations such as energy-aware admission control, network traffic shaping, and video transcoding. The on-device adaptation schemes include dynamic power management techniques for the LCD backlight, the wireless NIC, and the CPU. A context-aware wireless multimedia system proposed in [56] chooses the video content and performs media adaptation according to the changes of various contexts. The middleware components provide the contexts and then perform the reasoning of context information based on an ontology-based context model.

With the support of broadband wireless networks for multimedia applications, how to integrate the various standards of the MAC layer and PHY layer is an interesting topic. It is important that the mobile users can perceive acceptable QoS continuously while they are moving between different access points and networks. Mobile devices are equipped with multiple wireless interfaces and should be able to detect the underlying access network type to adjust the configurations in a heterogeneous network environment. Fernandez et al. [57] proposed an approach that allowed the users to dynamically negotiate QoS profiles with different underlying networks. Bandwidth aggregation over multiple network interfaces of the mobile device mitigates the resource constraints in a wireless network. The mobile users can negotiate their desired service levels and achieve them by using one or more interfaces. Nimmagadda, Kumar and Lu [58] proposed an adaptation of the presentations based on preferences and temporal constraints. The layout is generated by computing the locations, starting times, and durations of the media files.

5.3 Mobility Management

Because of its roaming characteristic, a mobile device needs to switch to different networks to maintain the pervasive and ubiquitous service. Currently, mobile devices have multiple wireless interfaces which enable them to access heterogeneous networks. Handover management determines which access network to switch to when there are multiple wireless networks in the vicinity. Handover can be categorized into horizontal

handover [59] and vertical handover [60]. Horizontal handover refers to the handoff between different access points of the same networking technologies; while vertical handover refers to the handoff across heterogeneous networks, such as WLAN, WiMAX and Cellular network. The handover strategy should not be solely based on the Received Signal Strength (RSS). Several other factors also need to be taken into account, such as the context, quality of service, and user's preference. In the mobile multimedia applications, the handover schemes should consider the quality of the received video or perceived QoE after switching the access network.

Some mobility management systems with the focus on vertical handover have been designed to support QoS of multimedia applications. Fernandes and Karmouch [60] proposed the Context-Aware Mobility Management System (CAMMS) for vertical mobility management. CAMMS is a cross-layer architecture where the handover decision is based on the information from at least two network layers, from the data link layer to the application layer. It considers the context information, power consumption, user preferences, and network conditions. The cross-layer handover scheme proposed in [61] tries to balance the load among different networks. It achieves the maximization of the overall system QoS and user perceived QoE by efficiently utilizing the available communication resources. Wu, Yang and Hwang [62] proposed a handover decision scheme using IEEE 802.21 [63] MIH services in WLAN and WiMAX networks to maintain nearly identical QoS in the handover. The handover decision scheme first used the Analytical Hierarchical Process to calculate the weights of the traffic parameters. Then it applied a typical multiple attributes decision making to calculate the QoS score, and ranked the preferred network according to the score.

5.4 Future challenges and directions

Due to the advancement of new technologies and applications, there are new research areas generated from the new applications. Here, we will discuss some future research directions and challenges.

1. Because of the popularity of social networking sites, such as Facebook, Twitter, MySpace and LinkedIn, the amount of user-generated multimedia data is increased rapidly. Social multimedia content on the Internet generates a new research area which is called social multimedia computing [64]. The mobile device is now involved in the multiple stages of multimedia applications. It needs a new design to integrate those dynamic contents to the applications, platforms, devices, and services. Effective and efficient multimedia retrieval and delivery in a community-contributed multimedia environment [65] pose new challenges.
2. Mobile cloud media computing [66]: Cloud computing is a new computing paradigm to provide services through the Internet. Computing resources such as processing, memory, and storage are not physically provided by the users [67]. Instead, they are provided and managed by the service providers, while being accessed by the users. Cloud computing technology is promising in providing multimedia services with QoS. How to leverage the media cloud computing platform to address the proliferation of multimedia devices and applications is a new research area.

6. Conclusion

As the popularity of the multimedia applications, mobile devices, and wireless networks increases, the provision of QoS for multimedia services in mobile computing environments becomes a challenging task. IEEE 802.11 wireless LAN is widely deployed because of its low cost and ease of management. This paper presented an overview of the research to provide QoS for mobile multimedia applications with a focus on the 802.11e WLAN. We first reviewed the challenges of the QoS in mobile multimedia environments. We then categorized the researches in the QoS provision at 802.11e from the perspectives of the layered design and the cross-layer design. We also covered several important aspects in the QoS provision of mobile multimedia, such as the power-efficiency and heterogeneity. Finally, we identified some topics and challenges for future research.

Author details

¹Department of Computer and Electrical Engineering Technology and Information Systems and Technology, Indiana University - Purdue University Fort Wayne, Fort Wayne, IN, USA ²Department of Electrical and Computer Engineering, University of Miami, Coral Gables, FL, USA

Authors' contributions

HL drafted the manuscript. M.-LS was involved in drafting the manuscript, and revised the manuscript. Both authors read and approved the final manuscript.

Author's Information

Hongli Luo is currently an assistant professor in the Department of Computer and Electrical Engineering Technology and Information Systems and Technology at Indiana University Purdue University Fort Wayne, Fort Wayne, IN, USA. She received her Ph.D. degree in Electrical and Computer Engineering from University of Miami, Coral Gables, FL, USA in 2006. Her research interests are multimedia networking, wireless networking, video streaming and quality of service. She was a program co-chair of the IEEE International Workshop on Semantic Computing and Multimedia Systems and program vice co-chair of the 2010 International Conference on Multimedia and Ubiquitous Engineering. She also serves on the editorial boards for *Journal of Information Processing Systems* and *International Journal of Smart Home*. Mei-Ling Shyu has been an associate professor at the Department of Electrical and Computer Engineering (ECE), University of Miami (UM) since June 2005. Prior to that, she was an assistant professor in ECE at UM dating from January 2000. She received her PhD degree from the School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana, USA in 1999, and her three master degrees in computer science, electrical engineering, and restaurant, hotel, institutional, and tourism management from Purdue University in 1992, 1995, and 1997. Her research interests include multimedia data mining and information systems. She has co-authored more than 160 technical papers published in prestigious journals, book chapters, and refereed conference/workshop/symposium proceedings. She received the Best Student Paper Award from the Third IEEE International Conference on Semantic Computing in September 2009 and the Johnson A. Edosomwan Scholarly Productivity Award from the College of Engineering at UM in 2007.

Competing interests

The authors declare that they have no competing interests.

Received: 18 October 2011 Accepted: 22 November 2011 Published: 22 November 2011

References

1. Hefeeda M, Hsu C-H (2010) Mobile Video Streaming in Modern Wireless Networks. *Proceedings of the ACM International Conference on Multimedia* 1779–1780
2. IEEE 802.11 WG (2009) IEEE Standard for Information Technology - Telecommunications and Information Exchange between Systems - LAN/MAN Specific requirements, Part 11: Wireless LAN MAC and PHY Specifications. IEEE Draft 802.11e/D0.06
3. Jaimes A, Sebe N, Gatica-Perez D (2006) Human-Centered Computing: A Multimedia Perspective. *Proceedings of the ACM International Conference on Multimedia* 855–864
4. Karmakar G, Dooley LS (2007) Introduction to mobile multimedia communications. *Mobile Multimedia Communications: Concepts, Applications, and Challenges*, IGI Global 45–63
5. Song W, Tjondronegoro DW, Wang SH, Docherty MJ (2010) Impact of Zooming and Enhancing Region of Interests for Optimizing User Experience on Mobile Sports Video. *Proceedings of ACM International Conference on Multimedia* 321–330
6. Hanzo L, Tafazolli R (2007) A Survey of QoS Routing Solutions for Mobile Ad Hoc Networks. *IEEE Communications Survey & Tutorials* 9(2):50–70
7. Chen M, Zakhori A (2006) Multiple TFRC Connections Based Rate Control for Wireless Networks. *IEEE Transactions on Multimedia* 8(5):1045–1062

8. Zhu P, Zeng W, Li C (2007) Joint Design of Source Rate Control and QoS-Aware Congestion Control for Video Streaming Over the Internet. *IEEE Transactions on Multimedia* 9(2):366–376
9. Luo H, Wu D, Ci S, Sharif H, Tang H (2009) TFRC-Based Rate Control for Real-Time Video Streaming over Wireless Multi-Hop Mesh Networks. *Proceedings of IEEE International Conference on Communications* 1–5
10. Van Der Schaar M, Andreopoulos Y, Hu Z (2006) Optimized Scalable Video Streaming over IEEE 802.11 a/e HCCA Wireless Networks under Delay Constraints. *IEEE Transactions on Mobile Computing* 5(6):755–768
11. Mohapatra S, Dutt N, Nicolau A, Venkatasubramanian V (2007) DYNAMO: A Cross-Layer Framework for End-to-End QoS and Energy Optimization in Mobile Handheld Devices. *IEEE Journal on Selected Areas in Communications* 25(4):722–737
12. Wu D, Song C, Wang H (2007) Cross-layer Optimization for Video Summary Transmission over Wireless Networks. *IEEE Journal on Selected Areas in Communications* 25(4):841–850
13. Luo H, Shyu M-L, Chen S-C (2008) Video Streaming over the Internet with Optimal Bandwidth Resource Allocation. *Journal of Multimedia Tools and Applications* 40(1):111–134
14. Luo H, Shyu M-L, Chen S-C (2006) An Optimal Resource Utilization Scheme with End-to-End Congestion Control for Continuous Media Stream Transmission. *Computer Networks* 50(7):921–937
15. Schilit B, Adams N, Want R (1994) Context-Aware Computing Application. *Proceedings of 1994 First Workshop on Mobile Computing Systems and Applications* 85–90
16. Chen G, Kotz D (2000) A Survey of Context-Aware Mobile Computing Research, Technical Report TR2000-381, Dartmouth Computer Science
17. Liu Y, Guo Y, Liang C (2008) A survey on peer-to-peer video streaming systems. *Peer-to-Peer Networking and Applications* 1(1):18–28
18. Hwang IS, Wu JH (2008) Performance Assessment of Service Differentiation in IEEE 802.11e EDCA. *International Journal of Ad Hoc and Ubiquitous Computing* 3(1):21–32
19. Hui J, Devetsikiotis M (2005) A Unified Model for the Performance Analysis of IEEE 802.11e EDCA. *IEEE Transactions on Communications* 53(9):1498–1510
20. Patras P, Banchs A, Serrano P (2009) A Control Theoretic Approach for Throughput Optimization in IEEE 802.11e EDCA WLANs. *Mobile Networks and Applications* 14(6):697–708
21. Romdhani L, Ni Q, Turetli T (2003) Adaptive EDCF: Enhanced Service Differentiation for IEEE 802.11 Wireless Ad-hoc Networks. *Proceedings of IEEE Wireless Communications and Networking Conference (WCNC'03)* 1373–1378
22. Li M, Zhu H, Prabhakaran B (2010) Dynamic Priority re-allocation Scheme for Quality of Service in IEEE 802.11e Wireless Networks. *Wireless Networks* 16(3):759–774
23. Han B, Ji L, Lee S, Miller R, Bhattacharjee B (2009) Channel Access Throttling for Improving WLAN QoS. *Proceedings of the 6th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON '09)* 1–9
24. Ferng H, Liaw H (2009) Design of Fair Scheduling Schemes for the QoS-Oriented Wireless LAN. *IEEE Transactions on Mobile Computing* 8(7):880–894
25. Park E, Kim D, Choi C, So J (2007) Improving Quality of Service and Assuring Fairness in WLAN Access Networks. *IEEE Transactions on Mobile Computing* 6(4):337–350
26. Zhao Q, Tsang DHK (2008) An Equal-Spacing-Based Design for QoS Guarantee in IEEE 802.11e HCCA Wireless Networks. *IEEE Transactions on Mobile Computing* 7(12):1474–1490
27. Skyrianioglu D, Passas N, Salkintzis AK (2006) ARROW: An Efficient Traffic Scheduling Algorithm for IEEE 802.11e HCCA. *IEEE Transactions on Wireless Communications* 5(12):3558–3567
28. Arora A, Yoon S, Choi Y, Bahk S (2010) Adaptive TXOP Allocation Based on Channel Conditions and Traffic Requirements in IEEE 802.11e Networks. *IEEE Transactions on Vehicular Technology* 58(3):1087–1099
29. Ghazizadeh R, Fan P (2008) Adaptive Scheduling for VBR Traffic in Wireless LANs. *Proceedings of the 4th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM '08)* 1–4
30. Cicconetti C, Lenzi L, Mingozzi E, Stea G (2007) An Efficient Cross Layer Scheduler for Multimedia Traffic in Wireless Local Area Networks with IEEE 802.11e HCCA. *ACM Mobile Computing and Communications Review (MC2R)* 11(3):31–46
31. Bheemarjuna Reddy T, John JP, Siva Ram Murthy C (2007) Providing MAC QoS for Multimedia Traffic in 802.11e Based Multi-hop Ad hoc Wireless Networks. *Computer Networks* 51(1):153–176
32. Ansel P, Ni Q, Turetli T (2006) FHCF: A Simple and Efficient Scheduling Scheme for IEEE 802.11e Wireless LAN. *Mobile Networks and Applications* 11(3):391–403
33. Boggia G, Camarda P, Grieco LA, Mascolo S (2007) Feedback-based Control for Providing Real-time Services with the 802.11e MAC. *IEEE/ACM Transactions on Networking* 15(2):323–333
34. Inan I, Keceli F, Ayanoglu E (2006) An Adaptive Multimedia QoS Scheduler for 802.11e Wireless LANs. *Proceedings of IEEE International Conference on Communications* 5263–5270
35. Luo H, Shyu M-L (2009) An Optimized Scheduling Scheme to Provide Quality of Service in 802.11e Wireless LAN. *Proceedings of the Fifth IEEE International Workshop on Multimedia Information Processing and Retrieval (IEEE MIPR-2009)*. San Diego, California, USA pp 651–656
36. Abdelzaher T, Diao Y, Hellerstein JL, Lu C, Singhal S, Zhu X (2009) Tutorial: Recent Advances in the Application of Control to Network and Service Management. *Proceedings of the 11th IFIP/IEEE International Symposium on Integrated Network Management*, New York, NY
37. Hollot CV, Misra V, Towsley D, Gong WB (2001) A Control Theoretic Analysis of RED. *Proceedings of IEEE INFOCOM*. Anchorage, Alaska pp 1510–1519
38. He W, Nahrstedt K, Liu X (2008) End-to-end Delay Control of Multimedia Applications over Multihop Wireless Links. *ACM Transactions on Multimedia Computing, Communications and Applications* 5(2):1–20
39. Hong Y, Yang OWW (2007) Design of Adaptive PI Rate Controller for Best-Effort Traffic in the Internet Based on Phase Margin. *IEEE Transactions on Parallel Distributed System* 18(4):550–561
40. Hu W, Xiao G (2009) Design of Congestion Control Based on Instantaneous Queue Sizes in the Routers. *Proceedings of the IEEE Globecom* 1–6

41. Huang Y, Mao S, Midkiff SF (2009) A Control-theoretic Approach to Rate Control for Streaming Videos. *IEEE Transactions on Multimedia* 11(6):1072–1081
42. Argyriou A (2008) Error-Resilient Video Encoding and Transmission in Multirate Wireless LANs. *IEEE Transactions on Multimedia* 10(5):691–700
43. Zhang Y, Gao W, Lu Y, Huang Q, Zhao D (2007) Joint Source-Channel Rate-Distortion Optimization for H.264 Video Coding Over Error-Prone Networks. *IEEE Transactions on Multimedia* 9(3):445–454
44. Zhu P, Zeng W, Li C (2007) Cross-layer Design of Source Rate Control and Congestion Control for Wireless Video Streaming. *Advances in Multimedia* 1:3–15
45. Van Der Schaar M, Turaga D (2007) Cross-layer Packetization and Retransmission Strategies for Delay-sensitive Wireless Multimedia Transmission. *IEEE Transactions on Multimedia* 9(1):185–197
46. Ksentini A, Naimi M, Gueroui A (2006) Toward an Improvement of H.264 Video Transmission over IEEE 802.11e through a Cross Layer Architecture. *Communications Magazine* 44(1):107–114
47. Chilankurti N, Zeadally S, Soni R, Giambene G (2010) Wireless Multimedia Delivery over 802.11e with Cross-layer Optimization Techniques. *Multimedia Tools and Applications* 47(1):189–205
48. Cai J, Gao D, Wu J (2007) MAC-layer QoS Management for Streaming Rate-adaptive VBR Video over IEEE 802.11e HCCA WLANs. *Advances in Multimedia* 2007:1–11
49. Shankar NS, Van Der Schaar M (2007) Performance Analysis of Video Transmission over IEEE 802.11a/e WLANs. *IEEE Transactions on Vehicular Technology* 56(4):2346–2362
50. Zhang J, Wu D, Ci S, Wang H, Katsaggelos AK (2009) Power-Aware Mobile Multimedia: A Survey (Invited Paper). *Journal of Communications* 4(9):600–613
51. He Z, Liang Y, Chen L, Ahmad I, Wu D (2005) Power-Rate-Distortion Analysis for Wireless Video Communication Under Energy Constraints. *IEEE Transactions on Circuits and Systems for Video Technology* 15(5):645–658
52. Hsu C-H, Hefeeda M (2011) A Framework for Cross-layer Optimization of Video Streaming in Wireless Networks. *ACM Transactions on Multimedia Computing, Communications, and Applications* 7(1):1–28
53. Zhang F, Todd TD, Zhao D, Kezys V (2006) Power Saving Access Points for IEEE 802.11 Wireless Network Infrastructure. *IEEE Transactions on Mobile Computing* 5(2):144–156
54. Li Y, Todd TD, Zhao D (2005) Access Point Power Saving in Solar/Battery Powered IEEE 802.11 ESS Mesh Networks. *Proceedings of the IEEE Second International Conference on Quality of Service in Heterogeneous Wired/Wireless Networks (QSHINE'05)*
55. Kholaf AM, Todd TD, Koutsakis P, Lazaris A (2010) Energy Efficient H.263 Video Transmission in Power Saving Wireless LAN Infrastructure. *IEEE Transactions on Multimedia* 12(2):142–153
56. Luo H, Song C, Wu D, Tang H (2010) Adaptive Wireless Multimedia Communications with Context-Awareness Using Ontology-Based Models. *Proceedings of GLOBECOM* 1–5
57. Fernandez JC, Taleb T, Guizani M, Kato N (2009) Bandwidth Aggregation-Aware Dynamic QoS Negotiation for Real-time Video Streaming in Next Wireless Networks. *IEEE Transactions on Multimedia* 11(6):1082–1093
58. Nimmagadda Y, Kumar K, Lu Y-H (2010) Adaptation of Multimedia Presentations for Different Display Sizes in the Presence of Preferences and Temporal Constraints. *IEEE Transactions on Multimedia* 12(7):650–664
59. Sgora A, Vergados D (2009) Handoff prioritization and decision schemes in wireless cellular networks: a survey. *IEEE Communications Surveys & Tutorials* 11(4):57–77
60. Fernandes S, Karmouch A (2010) Vertical Mobility Management Architectures in Wireless Networks: A Comprehensive Survey and Future Directions. *IEEE Communications Survey & Tutorials* 99:1–19
61. Ciubotaru B, Muntean G-M (2009) SASHA-A Quality-Oriented Handover Algorithm for Multimedia Content Delivery to Mobile Users. *IEEE Transactions on Broadcasting* 55(2):437–450
62. Wu J, Yang S, Hwang B (2009) A terminal-controlled vertical handover decision scheme in IEEE 802.21-enabled heterogeneous wireless networks. *International Journal of Communication Systems* 22(7):819–834
63. IEEE 802.21 (2009) IEEE Standard for Local and metropolitan area networks- Part 21: Media Independent Handover. *IEEE Std 802.21-2008*
64. Tian Y, Srivastava J, Huang T, Contractor N (2010) Social Multimedia Computing. *Computer* 43(8):27–36
65. Mei T, Hsu WH, Luo J (2010) Knowledge Discovery from Community-Contributed Multimedia. *IEEE Multimedia* 17(4):1–17
66. Hua X-S, Hua G, Chen CW (2010) ACM Workshop on Mobile Cloud Media Computing. *Proceedings of the ACM International Conference on Multimedia*
67. Armbrust M., *et al* (2009) Above the Clouds: A Berkeley View of Cloud Computing. Technical Report UCB/EECS-2009-28. EECS Dept., Univ. of California, Berkeley

doi:10.1186/2192-1962-1-5

Cite this article as: Luo and Shyu: Quality of service provision in mobile multimedia - a survey. *Human-centric Computing and Information Sciences* 2011 **1**:5.