

# Improving Videophone Transmission over Multi-rate IEEE 802.11e Networks

Tiantian Guo, Jianfei Cai, Chuan Heng Foh, Yu Zhang  
 Centre of Multimedia and Network Technology  
 School of Computer Engineering  
 Nanyang Technological University  
 Singapore

**Abstract**—In this paper, we propose an adaptive system for improving videophone transmission over EDCA. We consider that a videophone contains a *constant bit rate* (CBR) voice source and a rate-adaptive video source. Two issues are addressed in this research. Firstly, how to solve the AP bottleneck problem, and secondly, how to adjust video source rate to improve the network performance. For the first issue, we propose the adjustment of the *transmission opportunity* (TXOP) to give AP a higher priority in voice transmission in order to eliminate the AC3 transmission bottleneck at the AP. For the second issue, our principle is to guarantee the throughput of voice traffic while transmitting as much video traffic as possible. Moreover, we consider more realistic multi-rate WLANs, where multiple transmission rates are used in the PHY layer depending on the underlying channel conditions.

## I. INTRODUCTION

In order to provide QoS supports for multimedia applications, the IEEE 802.11e standard [1] is developed to enhance the MAC mechanisms. The IEEE 802.11e defines a single coordination function, called the *hybrid coordination function* (HCF), which includes two medium access mechanisms: contention-based channel access and controlled channel access. In particular, the contention-based channel access is referred to as *enhanced distributed channel access* (EDCA), which extends the legacy *distributed coordination function* (DCF) by providing the MAC layer with per-class service differentiation.

To the best of our knowledge, so far the research on videophone or video-conferencing over WLANs is very limited. In [2], purely based on simulations, M. Shimakawa *et al.* assessed the ability of 802.11g in supporting video-conferencing and data applications but without providing any enhancement. As we know, videophone contains both voice and video data. Similar to VoIP, due to the symmetric property in the uplink and downlink traffic of voice and video transmission, there also exists the AP bottleneck problem in videophone. However, the problem is expected to be much severe since the video data volume is typically much larger than the voice data volume. Compared with the video streaming application, videophone has tighter QoS requirements such as the 150 msec end-to-end delay requirement and low delay jitter requirement, which are inherited from VoIP. In addition, unlike video streaming, where videos are often pre-encoded, in videophone the videos are encoded and decoded in real time.

In this paper, we propose an adaptive system for improving videophone transmission over 802.11e EDCA. We consider that a videophone contains a *constant bit rate* (CBR) voice source and a rate-adaptive video source. We mainly address two issues: (1) how to solve the AP bottleneck problem; (2) how to adjust video source rate to improve the network performance. For the first issue, although our previous work in [3] can solve the AP bottleneck problem through assigning AP and stations to different ACs, in this research, we propose to carefully allocate the *transmission opportunity* (TXOP) to give AP higher priority. For the second issue, our principle is to guarantee the throughput of voice traffic to some extent while transmitting as much video traffic as possible. Moreover, we consider more realistic multi-rate WLANs [4], where multiple transmission rates are used in the PHY layer depending on the underlying channel condition. The different transmission rates used in different mobile stations are carefully taken into account in computing TXOP and video source rates in our proposed adaptive system.

The rest of the paper is organized as follows. Section II describes the system setup and the problems of videophone over multi-rate WLANs. Section III introduces our proposed adaptive system in details. Section IV presents the simulation results and finally Section V concludes the paper.

## II. SYSTEM SETUP AND PROBLEM STATEMENT

We consider a common scenario of videophone over a WLAN as shown in Fig. 1, where an AP and many mobile stations form a single 802.11 basic service set (BSS). The BSS is connected to the Internet via the AP. A videophone conversation over EDCA typically involves two users, which are connected through a WLAN access network. We use the 802.11b PHY layer, which supports four different data rates (1 Mbps, 2 Mbps, 5.5 Mbps, 11 Mbps). The particular data rate for a mobile station is determined by its distance to the AP [5].

For the videophone application, voice and video are typically encoded separately and transmitted in different packets [6]. Both voice and video are packetized using RTP/UDP/IP protocols. We assume the voice and video synchronization is handled by either the application layer or the transportation layer through time stamps. Voice traffic is assumed to be generated by a G.711 voice codec with a

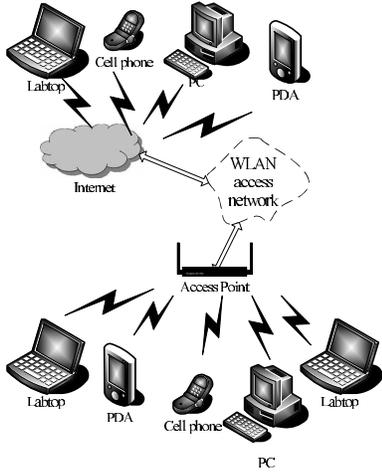


Fig. 1. A common scenario of Videophone over WLANs.

TABLE I  
IEEE 802.11e MAC PROTOCOL SYSTEM PARAMETERS

Access category	AIFSN	$CW_{min}$	$CW_{max}$	Queue length	Maximum retry limit, $r$
AC3	2	7	15	25	8
AC2	2	15	31	25	8
AC1	3	31	1023	25	4
AC0	7	31	1023	25	4

packetization interval of 20 ms and a packet size of 160 bytes. From the view point of MAC layer, the frame payload size must add 12 bytes for RTP, 8 bytes for UDP and 20 bytes for IP, and the total size goes to 200 bytes, which corresponds to a data rate of  $200 \times 8/20 = 80$  kbps.

On the other hand, the video traffic in videophone is typically encoded by simple video codecs such as H.263 or MPEG-4 with a pattern of IPPPP..., i.e. one I frame followed by many P frames, to form a long *group of picture* (GOP) size. In this research, we use the ‘foreman’ sequence as the test sequence, which is encoded by H.263+ at 30 frames/second with a bit rate ranging from 100 kbps to 300 kbps. The GOP size is set to 300 frames, which means that a GOP consists of one I frame followed by 299 P frames. Each frame is encoded and packetized into one packet. At the receiver side, if one frame is lost, all the subsequent P frames are discarded until a new I frame is received.

According to the 802.11e EDCA standard, voice and video traffic are delivered over two different ACs, i.e. AC3 and AC2 respectively, in the MAC layer. Table I shows the IEEE 802.11e system parameters, whose constants follow the standard [1] whenever specified.

Fig. 2 show the voice throughput results using ns-2 simulation for videophone over EDCA, where the video source rate is fixed at 100 kbps. It can be seen that at 11 Mbps, when the number of stations  $N$  is beyond seven, the AP cannot satisfy the voice throughput requirement anymore although the uplink could still support voice throughput till around  $N = 11$ . This is mainly due to the AP bottleneck problem.

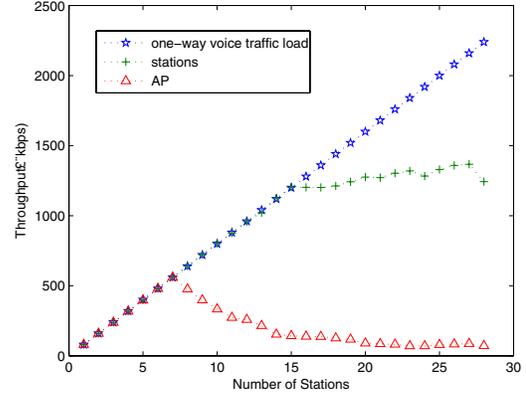


Fig. 2. The voice throughput results using ns-2 simulation for videophone over EDCA with 11 Mbps PHY data rate.

### III. ADAPTIVE SYSTEM DESIGN

In order to solve the problems stated in the previous section, we propose an adaptive system. In the application layer, we implement the video source rate adjustment module in both the AP and mobile stations. The MAC layer consists of a TXOP adaptation module and a channel bandwidth estimation module.

#### A. Bandwidth Estimation

Our design adopts the IdleGap scheme [7] for the bandwidth estimation, which passively measures the network condition and provides the idle period for the period of estimation.

We define a constant time interval  $C$ . Each station and the AP perform periodical bandwidth estimation during the time interval  $C$ . *Idle\_rate* indicates the rate at which the link is idle. During  $C$ , the idle time ( $IT$ ) can be obtained by the following product:

$$IT = C \times Idle\_rate \quad (1)$$

If a node in a WLAN is utilizing the resource, other node(s) should wait for the release of the wireless channel. A node is a receiver if receiving data; when a node does not join the transmission, it is an onlooker; If a node transmits data to another node, it is a sender. During a transmission, a node can be one of the following: sender, receiver or onlooker. We can get the busy time of the channel through observing the transaction activities on one node. The transaction time of node  $i$  can be obtained via the sum of the sending and receiving time to/from node  $i$ , as well as the onlooking time. Therefore, the busy time  $BT$  during  $C$  in any node  $i$  in the network can be deduced as:

$$BT = ST_i + RT_i + OT_i, \quad (2)$$

where  $ST_i$  is the sending time from node  $i$  to  $j$ ,  $RT_i$  is the receiving time from node  $j$  to  $i$ ,  $OT_i$  is the on-looking time at node  $i$ .

The *Idle\_rate* is obtained using the busy time:

$$Idle\_rate = 1 - BT/C \quad (3)$$

## B. Transmission Opportunity Adjustment

The TXOP adaptation module appropriately sets the  $TXOP_{limit}$  parameter in EDCA to promote the channel utilization of the AP. Since the AP handles all downlink voice and video traffic transmitted by remote clients, it is necessary to differentiate the transmission opportunity between the AP and mobile stations. We propose a method to adaptively adjust AC3 TXOP at AP over the time to help solve the AP bottleneck problem, yet ensure the AP not to excessively occupy the common channel. Note that there is no TXOP adaptation in mobile stations since it is not necessary, the multi-rate property of the links is taken into consideration.

Let  $d$  denote the different PHY data rate,  $d \in \{11 \text{ Mbps}, 5.5 \text{ Mbps}, 2 \text{ Mbps}, 1 \text{ Mbps}\}$ ,  $N_d$  is the number of links of every PHY data rate.  $r_A$  is the audio data rate, which is 80 kbps.

The AC3 transmission opportunity with adaptation for AP in each interval is:

$$txop_{voice} = \begin{cases} \sum_d C \cdot r_A \cdot \frac{N_d}{d}, & txop_{voice} < 0.5C \\ 0.5C, & 0.5C \leq txop_{voice} < C \end{cases} \quad (4)$$

## C. Video Source Rate Adjustment

Considering the scenario shown in Fig. 1, we apply the same approach for the video source rate adjustments in stations and the AP. We assume that AP has adjusted video source rates for remote clients using RTP/UDP/IP protocols.

After detecting the channel condition, every node computes how much time has been allocated to the voice traffic on the AP, then it adjusts the video source rate according to the time left in the time interval. This is done automatically and in time when a new interval begins. Each video packet is encoded and packetized into one frame. Each packet is assigned a *relative priority index* (RPI) to indicate the relative importance of a packet. We simply set the RPI value for a packet according to the corresponding frame position in a GOP. Specifically, the I frame in a GoP is given the highest priority (lowest RPI value) and the early P frames have higher priority than the later P frames in a GOP. If an early frame is lost, all the subsequent frames become useless until a new I frame. According to the RPI reported in the video packets, the adaptation have to consider the frame transferring among every time of source rate adaptation.

### 1) The Proposed Algorithm

Our algorithm first reserves necessary bandwidth for voice packets for each station in both the uplink and the downlink. Next we compute the sustainable video source rate for each videophone session, as shown in Table II. Both the stations and the AP use the same approach for the computation. Based on the previous period of transmission, if they are unable to clear their buffer which is likely due to channel congestion, the transmission rate in their previous period will be used as the transmission capability of their next period, where the new source rate is set to the transmission rate in the previous period. On the other hand, if they are able to clear their

TABLE II  
VIDEO SOURCE RATE ADJUSTMENT

---

Let $v_i, \hat{v}_i$ and $v_{min}$ denote the allocated, real, and minimum video source rates for station $i$ . Let $N$ denote the total number of stations. Let $d_i$ denote the PHY data rate for station $i$ .
<b>procedure</b> VideoRateAdapt( $N, d_i$ ) Obtain idle time from IdleGap and store it in $T_{idle}$ $T_{useful} \leftarrow \alpha \cdot T_{idle}$ Calculate the source rate increment $\psi_i$ based on $T_{useful}$ , $N$ and $d_i$
if $\hat{v}_i < v_i$ , then $v_i \leftarrow \hat{v}_i$ if $\hat{v}_i = v_i$ , then $v_i \leftarrow \hat{v}_i + \psi_i$ if $v_i < v_{min}$ , then $v_i \leftarrow 0$ return( $v_i$ )
<b>end</b>

---

buffer indicating under utilization of the channel, the idle time reported by IdleGap is evaluated to decide the new source rate. However, not all idle time reported by IdleGap can be translated into useful transmissions. Overheads are present due to protocol headers, transmission collisions, and IEEE 802.11e backoff idling. We analyze these overheads later and construct a table for a proper translation between idle time and potential useful transmission. This translation is described as  $T_{useful} \leftarrow \alpha T_{idle}$  in Table II, where  $\alpha$  denotes the AC2 utilization ratio.

For both the stations and the AP, video source rates are capped at a certain rate. In addition, a minimum video source rate is set such that any video source rate falling under the minimum source rate will be considered as unsatisfactory in viewing quality, and hence the video transmission will be halted in the next period. This transmission pausing will be recovered when sufficient idle time is discovered in the following period. There is a situation where the channel bandwidth is insufficient for all video sessions operating even at the minimum source rate. We recognize that pausing some video sessions may help generate enough bandwidth for other video sessions to continue their video transmissions. Those video sessions of inefficient bandwidth consumption should be knocked out, precisely, those operating at a lower PHY data rate. This selection is equivalent to maximizing the number of video sessions given the bandwidth constraint. The admission control part will be our future work.

### 2) Derivation of AC2 utilization ratio

The AC2 utilization ratio  $\alpha$  is used in our algorithm to translate idle time to the potential useful transmission. We seek existing model with an extension on multi-rate consideration to accurately predict the potential useful transmission when an idle time period is given. Our analysis follows that of [8] for EDCA, and we extend the analysis to include the multi-rate operation specifying in the IEEE 802.11e standard.

Considering only AC2, directly from [8], we have the probability that an AC2 station transmits in a slot

$$\tau = \frac{1 - p^{r+1}}{1 - p} b_{0,0}, \quad (5)$$

the collision probability

$$p = 1 - (1 - \tau)^{(n-1)}, \quad (6)$$

and the stationary distribution of the Markov chain model at the  $\{0,0\}$  state (at the 0-th backoff stage carrying zero in its backoff counter)

$$b_{0,0} = \begin{cases} \frac{\xi}{\iota_r + \kappa}, & r \leq m \\ \frac{\xi}{\iota_m + \kappa + \nu}, & r > m \end{cases} \quad (7)$$

where

$$\begin{aligned} \xi &= 2(1 - 2p)(1 - p) \\ \iota_r &= W_0(1 - (2p)^{r+1})(1 - p) \\ \iota_m &= W_0(1 - (2p)^{m+1})(1 - p) \\ \kappa &= (1 - 2p)(1 - p^{r+1}) \\ \nu &= W_0 2^m p^{m+1}(1 - 2p)(1 - p^{r-m}), \end{aligned} \quad (8)$$

and  $n$  denotes the number of stations,  $r$  and  $m$  are the maximum retry limit and the maximum backoff stage respectively, and  $W_0$  is the minimum backoff window size.

The probability  $P_I$  that a slot is idle is  $(1 - \tau)^n$ , while  $P_B$  that a slot is busy is given by  $1 - P_I$ . The probability  $P_S$  that a slot time contains a successful transmission by a station of class  $AC_2$  is given by  $P_S = n\tau(1 - \tau)^{(n-1)}$ . The probability  $P_C$  that a slot time contains a collision is  $P_C = 1 - P_I - P_S$ .

We consider that an idle time reported by IdleGap may be utilized by a number of video users. In such a case, the idle time period is equivalent to the contention of a number of video users with overheads of collisions and backoff idling, which has been modeled in the above analysis. The quantity  $\alpha$  is the ratio of the potential utilized time for payload to the reported idle period, which can be computed by

$$\alpha = \frac{P_S T_S}{E[ST]} \quad (9)$$

where  $E[ST]$  is the average length of a slot. The value  $E[ST]$  is computed by

$$E[ST] = P_I \sigma + P_S T_S + P_C T_C \quad (10)$$

where  $\sigma$  is the length of a slot time,  $T_S$  is the average time of successful transmission, and  $T_C$  is the average time of collision. Considering the basic access mode we adopt, the values  $T_S$  and  $T_C$  are calculated as

$$\begin{aligned} T_S &= AIFS + T_H + T[P] + SIFS + \delta + ACK + \delta, \\ T_C &= AIFS + T_H + T[P^*] + \delta, \end{aligned} \quad (11)$$

where  $T_H$  is the time for transmitting the packet header at the basic channel data rate,  $T[P]$  is the time for transmitting payload,  $T[P^*]$  is the average time to transmit the longest frame payload involved in a collision [9], and  $\delta$  is the propagation delay. Let  $T_1, T_2, T_3$  and  $T_4$  denote the transmission time for the four types of PHY data rates, and  $\rho_1, \rho_2, \rho_3, \rho_4$  denote the corresponding proportions for the number of stations in each data rate group. We compute  $T[P]$  as

$$T[P] = T_1 \times \rho_1 + T_2 \times \rho_2 + T_3 \times \rho_3 + T_4 \times \rho_4. \quad (12)$$

For calculating  $T[P^*]$ , if all packets have the same time length, obviously  $T[P^*] = T[P]$ . However, in the multi-rate

case, the collision period varies due to different PHY data rates. Thus,  $T[P^*]$  should be the longest time involved in a collision. Let  $t_i$  denote the transmission time of the  $i$ -th station. We derive

$$\begin{aligned} T[P^*] &= E[E[\max(t_1, t_2, \dots, t_k)|k]] \\ &= \sum_{k=2}^n C_n^k \tau^k (1 - \tau)^{n-k} E[\max(t_1, t_2, \dots, t_k)|k] \end{aligned} \quad (13)$$

and

$$\begin{aligned} E[\max(t_1, t_2, \dots, t_k)|k] &= T_1[1 - (1 - \rho_1)^k] \\ &+ T_2[(1 - \rho_1)^k - (1 - \rho_1 - \rho_2)^k] \\ &+ T_3[(1 - \rho_1 - \rho_2)^k - (1 - \rho_1 - \rho_2 - \rho_3)^k] \\ &+ T_4 \cdot \rho_4^k. \end{aligned} \quad (14)$$

With the above analysis,  $\alpha$  is readily obtainable. In our simulation implementation, a table is setup to hold all possible combinations of different number of stations at different PHY data rates. For example, There are eight stations ready to communicate with access point. First, we set all stations to be 11 Mbps, then change the PHY data rate to be 5.5 Mbps one by one. we could calculate  $\alpha$  below:

TABLE III  
EXAMPLE OF  $\alpha$  UNDER MULTIRATE CONSIDERATION

Number of 11Mbps STAs	8	7	6	5	4
Number of 5.5Mbps STAs	0	1	2	3	4
$\alpha$	0.719	0.679	0.663	0.658	0.658

## IV. SIMULATION RESULTS

In this section, we evaluate the performance of our proposed adaptive system using ns-2 simulation. We modify the ns-2 EDCA component to accommodate multi-rate PHY environment and realize the video transmission and adaption. Table IV shows the WLAN system parameters used in the simulation.

TABLE IV  
THE COMMON PARAMETERS USED IN ALL THE EXPERIMENTS.

Parameters	Values
slot	20 $\mu$ s
MAC header	272 bit
PHY header	128 bit
ACK	112bits+PHY header
Channel bit rate	2Mbps
Propagation Delay	1 $\mu$ s
AIFS <sub>min</sub>	2 $\mu$ s
CW <sub>min</sub> (AC2)	15 slot

We first evaluate the system performance with our proposed AC3 TXOP adaptation. The voice traffic is set to 80 kbps and the video source rate is fixed at 100 kbps. Set use  $C = 0.5s$ . From Fig. 3, with the adaptation, the AP can provide a higher capacity up to 12 downlink voice transmissions, while maintaining 10 uplink voice transmissions. In other words, 10 videophone sessions may be achieved. On the other hand, without the adaptation, the AP provides 7 downlink and 12 uplink voice transmissions. This asymmetrical capacity utilization gives support to only seven videophone sessions.

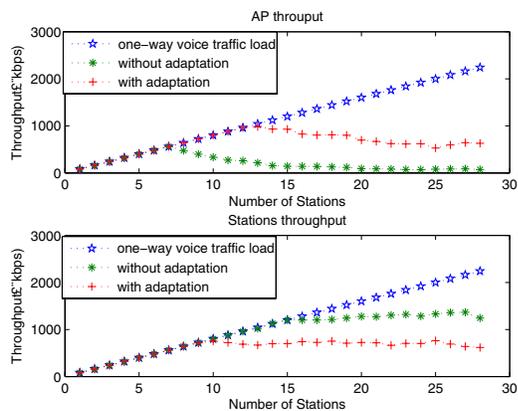


Fig. 3. The comparison of voice throughput in videophone using EDCA and our proposed scheme with a fixed video source rate of 100 kbps and a PHY data rate of 11 Mbps.

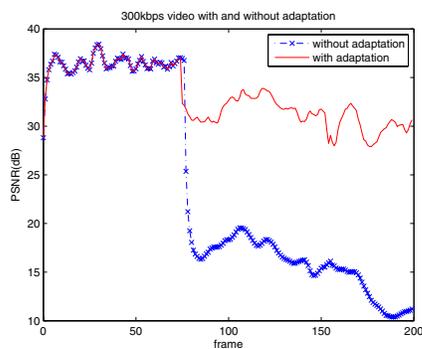


Fig. 4. The PSNR of 11 Mbps uplink.

Then, we conduct an experiment where four 11 Mbps stations start their transmission at the beginning of the simulation. At time 1.8s, we add another four 11 Mbps stations into the network. The simultaneous arrivals of the new stations push the network to operate at near its capacity limit. We measure the PSNR of the uplink received video stream from the first station. The results are shown in Fig. 4.

At the 75th frame transmission time, when new arrivals appear on the channel, the network congestion occurs and the network without adaptation fails to deliver acceptable quality of video stream. The measured PSNR dropped below 20 dB in this case. Whereas, with adaptation, each station recalculates its source rate achieving an overall lower load and avoiding network congestion. As a result, acceptable video quality is observed.

Finally, the PHY data rate of one of the stations transmitting at the beginning of the simulation is adjusted to 5.5 Mbps. The same experiment is conducted and results are shown in Fig. 5. As can be seen, the lower PHY data rate links transmit less traffic than that of the higher PHY data rate links since our algorithm explicitly gives priority to high PHY data rate stations.

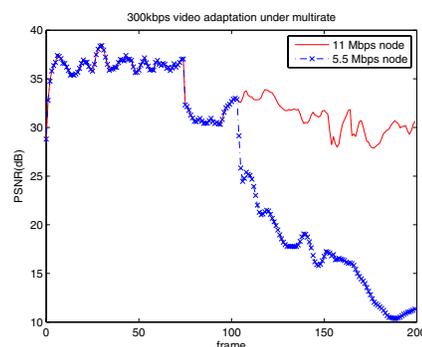


Fig. 5. The PSNR of multirate uplink.

## V. CONCLUSION

In this paper, we have studied the improvement of videophone transmission over multi-rate WLANs. Specifically, we proposed to use TXOP to give AP longer transmission time to solve the AP bottleneck problem of AC3. We also proposed to adjust video source rates to maintain the network at stable conditions. The PHY multi-rate constraints are carefully considered in our design, particularly for the computation of TXOP and video source rates. Simulation results showed that our proposed system can improve the number of videophone sessions from 7 to 10.

## REFERENCES

- [1] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications Amendment 8: Medium Access Control (MAC) Quality of Service Enhancements*, IEEE Std. 802.11e-2005, 2005.
- [2] D. M. Shimakawa and F. A. Tobagi, "Video-conferencing and data traffic over an IEEE 802.11g WLAN using DCF and EDCA," in *Proc. IEEE ICC'05*, vol. 2, May 2005, pp. 1324–1330.
- [3] D. Gao, J. Cai, C. H. Foh, C. T. Lau, and K. N. Ngan, "Improving WLAN VoIP capacity through service differentiation," *IEEE Trans. Circuits and Systems for Video Technology*, Accepted.
- [4] E. Kim and Y.-J. Suh, "Atxop: an adaptive txop based on the data rate to guarantee fairness for IEEE 802.11e wireless LANs," in *Vehicular Technology Conference*, vol. 4, 2004, pp. 2678–2682.
- [5] D. Y. Yang, T. J. Lee, K. Jang, J. B. Chang, and S. Choi, "Performance enhancement of multi-rate IEEE 802.11 WLANs with geographically-scattered stations," in *IEEE Trans. on Mobile Computing (TMC)*, vol. 5, 2006, pp. 906–919.
- [6] C. Liu and S. S. Chen, "Providing unequal reliability for transmitting layered videostreams over wireless networks by multi-arq schemes," in *Proc. IEEE ICIP*, vol. 3, 1999, pp. 100–104.
- [7] H. K. Lee, V. Hall, K. H. Yum, K. I. Kim, and E. Kim, "Bandwidth estimation in wireless LANs for multimedia streaming services," in *Proc. IEEE International Conference on Multimedia and Expo*, 2006.
- [8] Tantra, C. H. Foh, and A. B. Mnaouer, "Throughput and delay analysis of the IEEE 802.11e EDCA saturation," in *Proc. IEEE ICC'05*, June 2005.
- [9] G. Bianchi, "Performance analysis for the IEEE 802.11 distributed coordination function," vol. 18, Mar 2000, pp. 535–547.